Transportation Research Part C 55 (2015) 14–23

Contents lists available at ScienceDirect

Transportation Research Part C

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

Optimization of traffic forecasting: Intelligent surrogate modeling

Eleni I. Vlahogianni

Department of Transportation Planning and Engineering, School of Civil Engineering, National Technical University of Athens, Greece

ARTICLE INFO

Article history:
Received 31 October 2014
Received in revised form 10 March 2015
Accepted 10 March 2015
Available online 1 April 2015

Keywords:
Short-term prediction
Optimization
Surrogate models
Computational intelligence
Traffic flow

ABSTRACT

Transportation modelers are frequently faced with several optimization challenges related to model selection and parameter optimization for forecasting. The concept of surrogate modeling is discussed in order to tackle some limitations related to the practice of developing short-term forecasting algorithms. An automated meta-modeling technique is presented that uses heterogeneous information from multiple types of statistical and computationally intelligent models, along with multi-objective evolutionary strategies to optimize the model and parameter selection. A number of different models from the family of Support Vector Machines, Radial Base Functions and Neural Networks are jointly considered and optimized with the aim to improve the short-term predictability of travel speed. Results are presented and discussed in both a univariate and multivariate framework.

1. Introduction

Forecasting is a process of ultimate importance in the science and practice of transportation for making informed decisions. Forecasting models and algorithms have been considered to be the cornerstone of modern Intelligent Transportation Systems (ITS), because they can boost the efficiency and transferability of ITS services. Forecasting is not a new challenge in transportation and traffic engineering (Adeli, 2001, Vlahogianni et al., 2004, 2014; Song and Li, 2008; Rasouli and Timmermans, 2012; van Lint and van Hinsbergen, 2012; Vlahogianni et al., 2014).

Recently, the advent of novel automated data collection systems – mainly based on mobile sensors and social participatory systems (foursquare, etc.) – that may provide a plethora of data of high resolution and spatial coverage, have marked a new era of computationally intelligent modeling. Computationally intelligent models have been introduced to transportation in the early 80s and are frequently based on advanced optimization techniques, such as evolutionary computation, swarm intelligence and so on, for input and model parameter selection (Vlahogianni et al., 2005, 2008; Teodorović, 2008; Chan et al., 2013; Yao et al., 2014; Feng and Timmermans, 2014; Liu et al., 2014).

Regardless of the many advantages of such approaches – for example their "open" and flexible algorithmic form, the lack of constraints and so on – researchers seem to converge to one and only outcome, that is, the need to develop a holistic manner to select and optimize the specific types of models in order to assure a reliable and computationally less expensive modeling. The scope of the paper is to introduce the concept of surrogate modeling to transportation and propose some modeling solutions to the specific optimization problems that a transportation researcher may face when dealing with forecasting models. An automated modeling technique is discussed and evaluated in the framework of short-term traffic prediction.

E-mail address: elenivl@central.ntua.gr

http://dx.doi.org/10.1016/j.trc.2015.03.016

2. Oopen issues in forecasting

2.1. Model type selection

Some transportation research fields, e.g. short-term forecasting, are considered an excellent testbed for developing and testing complex prediction algorithms, because of the abundance of high temporal resolution data with significant spatial coverage. A critical challenge refers to the type of the model to be used to provide predictions. The general approach followed in most transportation applications is to select the model, which provides the most accurate predictions based on the available dataset. The selection of the proper modeling approach should be largely determined by the properties of the specific dataset and the problem being modeled. However, in most transportation problems, little is known on the underlined dynamics of the phenomenon, leading to assume that there is no single modeling strategy that may well fit the real data, but "reality" may be modeled by a series of different modeling approaches that may act concurrently or complementary (Vlahogianni et al., 2006; Vlahogianni, 2009). This implies that there is a need to develop a heterogeneous prediction strategy to find the optimum set of different types of models, which may provide accurate reliable (in terms of stability in accuracy) predictions.

2.2. Hyperparameter optimization

Assuming that there is a feasible and acceptable solution to the problem of selecting the optimum type of model to produce predictions, the next step is to provide the optimum parameters for the selected model. This demands a great deal of manpower and expertise and is frequently tackled through a trial and error procedure. The greater the number of models considered, the more difficult the optimization problem becomes. In such conditions, an automated manner to optimize the parameters of the different models is required. In the recent literature, selecting model parameters is tackled using optimization approaches, for example genetic algorithms, swarm optimization techniques, simulated annealing and so on (hyperparameter optimization) (Teodorović, 2008; Karlaftis and Vlahogianni, 2011). Very few attempts have dealt with the optimization of multiple different modeling structures and this has been done in a sequential and manual manner.

2.3. Forecasts combination

Comparing both modeling specifications and results are imperative to support the usefulness of a proposed forecasting scheme. Karlaftis and Vlahogianni (2011) discussed the usefulness and efficiency of current comparative studies and suggested that most comparisons conducted are not always fair, particularly when comparing complex nonlinear to simple linear models. Accuracy is of great importance, yet, not the only determinant for selecting the appropriate methodology.

Although selecting the "best" model among a set of baseline models through testing and comparisons is of importance, a practical alternative is to provide a model or algorithm or heuristic approach to combine predictions. Combining should be useful in cases where the modeler cannot resort to a single well-specified model, a common case in complex data forecasting (Zheng et al., 2006; Tan et al., 2009; Vlahogianni et al., 2014). There are various approaches to forecasts combinations; interdisciplinary literature has shown that averaging between different baseline prediction may form a simple and viable alternative (Clemen, 1989).

2.4. Multi-objective optimization of forecasting models

A critical question in transportation modeling is how to judge on the performance of a model. This is basically done using a specific error function, such as the mean square error, the mean absolute percent error, the mean relative percent error and so on, as well as a way for estimating the generalization capabilities of a model (cross validation, bootstrap, validation set and so on) (Washington et al., 2010). Evidently, performance is related to error functions and measures for accounting for the generalization power considered. Measures may be broadly divided into absolute (mean absolute error, mean square error, root mean square error, etc.) and relative (mean absolute percent error, relative square error, etc.); some of the most popular error metrics are seen in Table 1.

Table 1
List of measures applied to transportation forecasting.

	Definition
Absolute	
Mean Absolute Error (MAE)	$\frac{\sum_{i=1}^{n}\lvert y_i-\hat{y}_i\rvert}{n}$
Root Mean Square Error (RMSE)	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}$
Relative	
Mean Absolute Percent Error (MAPE)	$\frac{1}{n}\sum_{i=1}^{n}\frac{\lvert y_i-\hat{y}_i\rvert}{y_i}$
Root Relative Square Error (RRSE)	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}$

These measures have their strengths and weaknesses, e.g. the sensitivity to high or low error values, error averaging and so on, making difficult the final decision of which measure to introduce in the objective function (Gorissen et al., 2010b). An efficient approach may be to deal with the above problem as a dynamic multi-objective optimization problem in the hyperparameter space, where each criterion is an objective and a Pareto-optimal set of parameters is selected through ranking procedures (Jin and Sendhoff, 2008).

3. Surrogate modeling

Surrogate modeling is a macro-modeling technique that aims to minimize the time and computational load of developing simulations to replicate input–output relationships (Forrester et al., 2008). Surrogate models are often referred to as approximation, meta- or response surface models. The aim is to produce a faster and simpler approximation of a simulator to make optimization, design space exploration, etc. feasible. To illustrate the above, consider an engineer who wants to develop a model to predict transit demand based on given input data. He/she will have to select a set of parameters in relation to the selected model, which will provide the optimum result in terms of an objective (e.g. accuracy of predictions) for the available dataset. For parameter selection, the engineer will have to conduct a large number of iterations with a significant computational effect on model development in order to cover the largest possible design parameters' space. To assist the parameter selection, powerful optimization techniques are applied (i.e. genetic algorithms, particle swarm optimization, etc.). This procedure may prove to be extremely time-consuming and may not guarantee the optimum solution. Thus, the question may be how to *construct a reliable metamodel that is as accurate as possible over the complete design space of interest using as few data points as possible* (Gorissen et al., 2010a).

Let $f : \Omega \mapsto \mathbb{C}^n$ defined on domain $\Omega \subset \mathbb{R}^d$. Moreover, let $X = \{x_1, \ldots, x_k\} \subset \Omega$ be sample points with function values $f(X) \subset \mathbb{C}^n$. The scope is to find a close approximation $s : \Omega \mapsto C^n \in S$ in space S of function f based on some criterion $\xi = (\Lambda, \varepsilon, \tau)$, where Λ is the generalization estimator, ε the error (or loss) function, and τ is the target value required by the user (global surrogate model generation problem):

$$s^* = \arg\min_{t \in T} \arg\min_{\theta \in \Theta} \Lambda(\varepsilon, s_{t,\theta}, D), \tag{1}$$

where s^* is the best approximation, $s_{t,\theta}$ is an approximation with parametrization $\theta \in \Theta$ of a type of model $t \in T$, with T a set of different types of models.

The proposed framework involves a series of steps and optimization procedures (Gorissen et al., 2010a): based on a sampling strategy, a small initial sample from the dataset is selected and different types of models are developed and optimized. The optimization may be based on one or more measures (multi-objective optimization). During optimization, each model is assigned to a score. The optimization continues until no further improvement is achieved to the performance of the selected models. Following, the models are ranked according to their score and new samples are selected based on the best performing models and the characteristics of the response variable. The entire process repeats itself until one of the following is satisfied: (i) the maximum number of samples has been reached, (ii) the maximum allowed time for model development and optimization is reached, or iii. the required accuracy (set be the modeler) has been met.

3.1. Multi-objective heterogeneous genetically optimized surrogate models

In the above modeling context, several modeling issues may arise, such as the types of models to consider, the sampling strategy, the optimization approach to follow and so on. One solution, also popular to transportation problems, is to use evolutionary strategies to optimize both the model selection process and the optimization of the parameters of the models (hyperparameter optimization) (Jin, 2011). In the present work, the island model is considered (Whitley et al., 1999); different sub-populations M_i are created for each model type $t \in T$, called *demes*. Each deme is allowed to evolve based on an elitist GA and have its own GA optimization problem representation. Parents are selected according a selection algorithm (e.g., tournament selection) and offspring go through either crossover with probability p_c or mutation with probability $1 - p_c$. The quality of the model is evaluated based on the fitness function according to specific criteria. The current deme population is, then, replaced with its offspring together with k elite individuals. For each generation, migration between individuals is allowed to occur based on three parameters: (i) the migration interval m_i, (ii) the migration fraction m_f,and (iii) the migration direction m_d. Migration allows solutions to mature in semi-isolation without being forced to consistently compete with others. Selection and recombination are restricted per deme, so as each deme can evolve toward different regions on the search space.

Each solution has a certain type of models and a certain number of models from each type. Since, there exist models that may perform poorly with limited data, the above algorithm is enhanced with an extinction prevention mechanism, which guarantees that a model type – initially participating in the surrogate approach – can never extinct from the results. Based on this, a type of model may be set not to fall below a certain threshold in the candidate models. This is accomplished by letting the latest vanished models of a specific type to replace the worst models of other types that have sufficient number of candidates in the solution. Moreover, to tackle the inevitable mix of different types of models, an ensemble approach is proposed; if two models of different types are selected to recombine, an ensemble is created with the models' members. Thus, as soon as migration occurs, model types start mixing, and ensemble models arise as a result. These are treated as a distinct model type. In the specific approach, ensemble members are forced to differ as predefined percent in their response. In the specific study, a flexible simple average ensemble is implemented. Finally, one should take into consideration that the approximation task may involve multiple conflicting criteria leading to assume that the best model type can vary per criteria. For this, a multi-objective approach is proposed to select the best modeling type. GAs are set to optimize the model type and hyperparameter selection based on these criteria.

4. Application to short-term traffic forecasting

4.1. The data

Data come from a 10 km section of Attica Tollway (Athens, Greece), serving a high daily demand corridor. The study area with the detectors' locations are seen in Fig. 1. The available data are from one week on January 2008. Observations of volume and speed per lane are available in 5 min intervals from detectors positioned every 500 m. Fig. 2 shows the time series of volume and speed in a single section for a typical day. The joint consideration of volume and speed time series depict a typical traffic behavior with shifts between high and low demand, as well as shifts from unconstrained to constrained traffic and vice versa.

Three different modeling approaches are evaluated:

- Short-term speed prediction S_t using time delayed information: ($S_t = f(\mathbf{S}_{\tau(1-m)})$, where τ and m is the time delay and the dimension of the look back time window.
- Short-term speed prediction enhanced with volume information (exogenous variable): $S_t = f(\mathbf{S}_{\tau(1-m)}, \mathbf{V}_{\tau(1-m)})$, where $\mathbf{V}_{\tau(1-m)}$is the volume time series.
- Short-term speed prediction enhanced with speed data from previous locations (exogenous variables): $S_t^x = f(\mathbf{S}^x_{\tau(1-m)})$, where $x = 1, \ldots, n$.

The first is the simple approach to short-term traffic forecasting where speeds are predicted at time t based on data at time $\tau(1-m)$. The second and third model will assess whether information on volume at time $\tau(1-m)$ may improve the predictability of speeds at time t. The last model will incorporate the effect of spatial patterns of speed to the predictability of speed in the location of interest. The first model is univariate, whereas the rest of the models introduced are multivariate.

To set the parameters τ and m, two non-linear methods previously used in short-term traffic forecasting are implemented; the mutual information and the false nearest neighbors criterion (details on the implementation may be found in Vlahogianni et al., 2006). The optimum time delay τ corresponds to the first local minimum in the relationship of the mutual information versus time delay, whereas, for a specific time delay τ, the global minimum of the phase nearest neighbors gives the optimum dimension m. As seen in Table 2, for the case of univariate speed prediction, mutual information resulted in $\tau = 1$, whereas, based on the false nearest neighbors criterion, m equals to 3, meaning 10 min speed data in the past will be used to predict speed 5 min in the future.

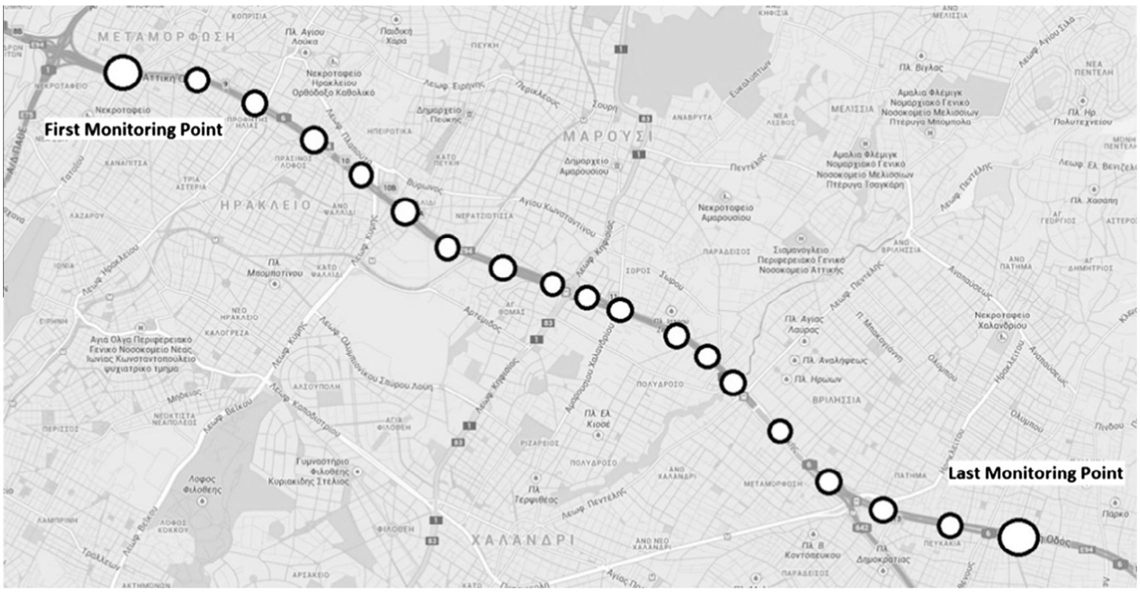


Fig. 1. Study area map.

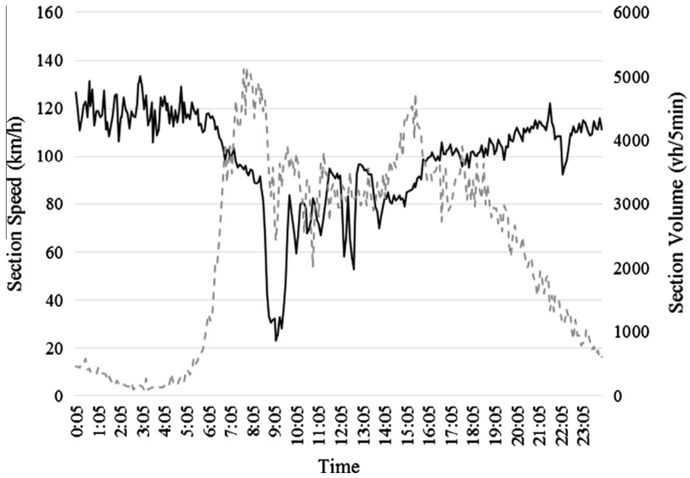


Fig. 2. Time series of section travel speed and volume every 5 min for a typical day.

Table 2
Models and spatio-temporal dependencies specifications.

Models	Endogenous information τ, m	Exogenous spatio-temporal dependencies τ, m
Univariate	1, 3	–
Multivariate (speed, volume)	1, 2	1, 2
Multivariate ($\text{speed}_{x,\,x-1}$)	1, 2	1, 2

To set the relationship between speed and volume time series, as well as the spatial dependencies for the speeds from previous locations, a cross-mutual information analysis is implemented, which revealed that speeds at the location of interest at time t is correlated to volume at $t-1$ and speeds at the upstream location (Table 2). The above representation resembles to an autoregressive time series model with exogenous information, for the case of models enhanced with volume and spatial information on speeds (Tselentis et al., 2015).

4.2. Candidate models for multi-objective optimization

Artificial Neural Networks (ANN), Least Squares Support Vector Machines (LS-SVMs) and Radial Base Function (RBF) models are considered in the surrogate models. Each modeling type has different parameters to be optimized using genetic algorithms. The result of a heterogeneous recombination will be an averaged model. The above model types will be competing to approximate the data. The ANN is a simple Multilayer Perceptron (MLP) with one hidden layer trained with Levenberg–Marquardt back-propagation with Bayesian regularization. The structural (number of hidden units) and learning parameters are optimized by the GA. For the LS-SVM models, the kernel type is chosen to be a RBF, whereas the support vector parameters are optimized by the GA. Finally, for the RBF models, the regression function, correlation function, and correlation parameters are GA optimized. Three RBF functions are considered: Gaussian, Multi-quadric and Exponential.

Data is sampled using the Local Linear adaptive sampling (LOLA) algorithm (Crombecq et al., 2009). LOLA identifies new sample locations by making a tradeoff between covering the design space evenly and concentrating on regions, where the actual response is nonlinear. LOLA is independent of the type of model and is able to automatically identify non-linear regions and sample these more densely compared to more linear, 'flatter' regions. This property of LOLA is critical to the efficiency of surrogate models, as the trained model is independent of the sample selection settings and any difference in performance between models cannot be due to differences in sample distribution. In the specific application, the starting sample size is 20, augmented with corner points. In each iteration 36 samples are added (equivalent to 3 h data) up to the maximum sample size, which is set to 288 (equivalent to 1 day data).

A maximum of 10 generations for each sampling interval is considered; the GA optimization terminates if either the maximum number of generations is reached, or 3 generations with no improvement. The size of each deme is set to 3, meaning initially, 9 different models will be constructed, 3 for each modeling paradigm. The migration interval *mi* is set to 4, the migration fraction *mf* to 0.1 and the migration direction is *forward* (copies of the *mf* best individuals from island *i* replace the worst individuals in islands $i+1$). A stochastic uniform selection function is used. The fitness of an individual is defined based on the root mean square error and the mean relative percent error in the cross-validation set (20% of data). Extinction prevention is taken into consideration. The surrogate modeling framework parameter specifications are seen in Table 3.

Table 3
Evolutionary surrogate model parameter specification.

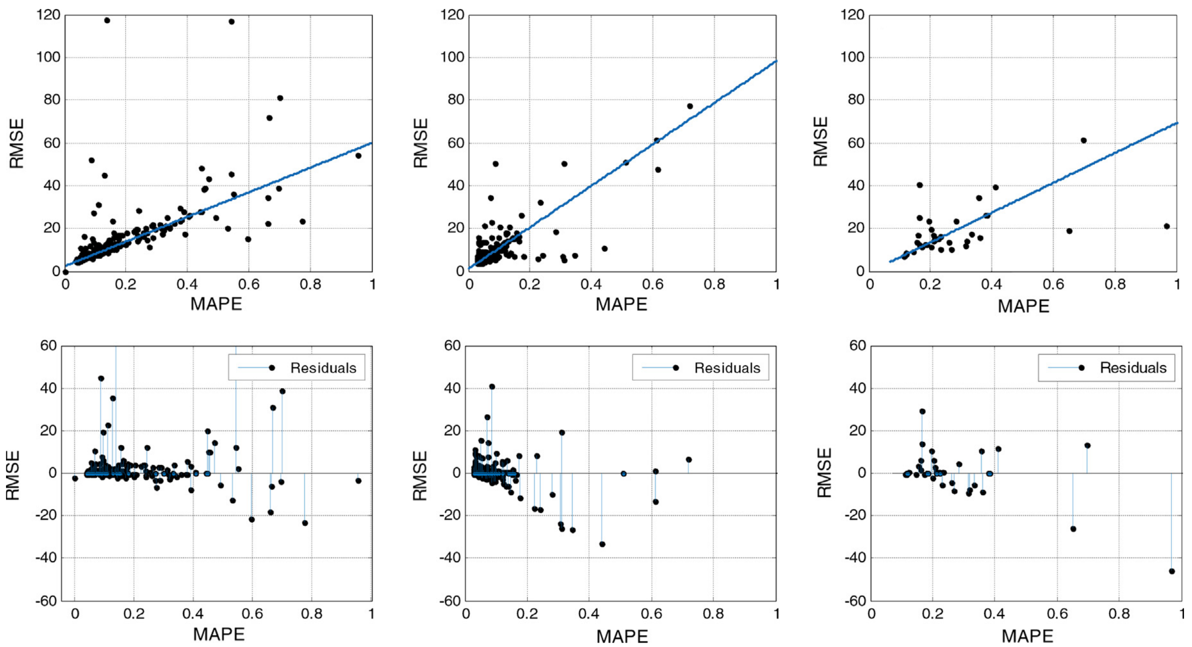
Parameter	Value
Generations	10
Deme size	3
Selection	Uniform
Crossover probability	0.7
Mutation Probability	0.3
Migration interval	4
Migration fraction	0.1
Migration direction	Forward
Fitness (cross-validation)	RRSE, MARE
Maximum ensemble size	4
k elite individuals	1
Ensemble: members percent difference	0.05
Ensemble: replace probability of a model	0.8
Ensemble: threshold value	2

4.3. Results

In the univariate framework, surrogate models are developed for one step ahead speed forecasting using as input past information on speed at $(t-2)$. In the multivariate framework, first, information on volume at time $t-1$ is added to the input space which, along with speed at $t-1$, is used to predict travel speed at time t. In the second stage, information of speed in the adjacent upstream location is used at time $t-1$ which, along with speed at time $t-1$, is used to predict speed at t. The above models are optimized with the aim to minimize the RMSE and the MAPE.

During optimization numerous models are constructed and evaluated. Fig. 3 depicts the scatter plots of the MAPE versus RMSE of the models developed for each one of the surrogate models. A linear relationship between the two different error measures may be observed.

The structure and accuracy of the different surrogate models developed are seen in Table 4. All models considered provide predictions of high accuracy, an evidence of the robustness of the proposed modeling framework and the optimization strategies; it seems that by adding exogenous information (either volume or speed from upstream location) the accuracy of the models improves. Interestingly, for all three modeling attempts, the resulting surrogate structure is the same with respect to the number of models from each category included in the optimized final surrogate model. The only difference



Adjusted R-square: 0.99, RMSE: 0.50 *(a)* *Adjusted R-square: 0.99, RMSE: 0.22* *(b)* *Adjusted R-square: 0.95, RMSE: 0.26* *(c)*

Fig. 3. Search space of surrogate models. (a) Speed. (b) Speed with volume information. (c) Speed with spatial information.

Table 4
Structure and accuracy (cross-validation) of the trained surrogate models for the univariate and multivariate framework.

Surrogate model	Model	Number	MAPE (%)	RMSE	Global score
Univariate	SVM	2	0.040	4.220	2.213
	Ensemble[a] (SVM, MLP)	2			
	RBF	2			
	MLP[b]	3			
Multivariate (speed volume)	SVM	2	0.031	3.386	1.709
	Ensemble[a] (SVM, RBF)	2			
	RBF	2			
	MLP[b]	3			
Multivariate (speed spatial information)	SVM[b]	2	0.031	4.046	2.039
	Ensemble[a] (MLP)	2			
	RBF	3			
	MLP	2			

[a] Average of the performance of the included models.
[b] Best model.

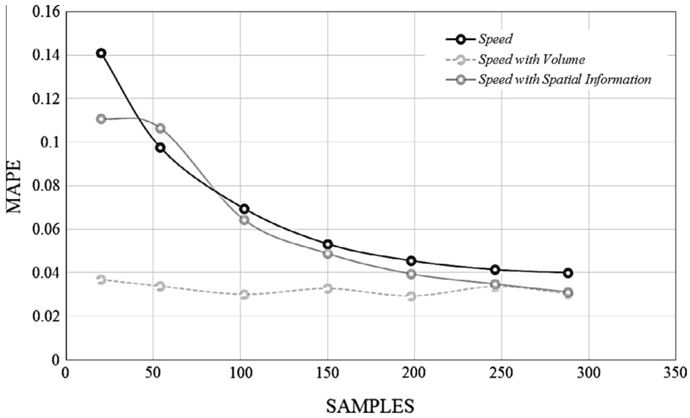


Fig. 4. MAPE evolution with sample size for the three surrogate models.

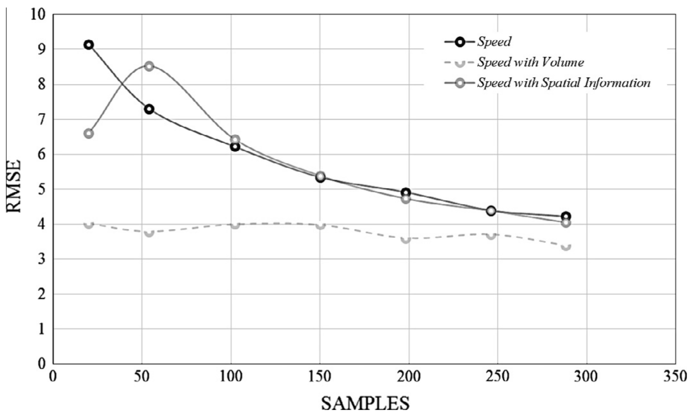


Fig. 5. RMSE evolution with sample size for the three surrogate models.

is in the accuracy of the final model, as well as in the structure of the different Ensemble models. In the surrogate model, where only speed from the location of interest is used, the Ensemble model encompasses a LS-SVM model and a MLP. In the multivariate framework, when volume is taken into consideration, the Ensemble consists of one LS-SVM model and a RBF model, whereas in the case of spatio-temporal speed information the Ensemble consists of different MLP models.

Figs. 4 and 5 depict the evolution of the MAPE and the RMSE with the sample size for the univariate and multivariate short-term speed forecasting framework respectively. Results show that a 5% accuracy may be achieved with approximately

Table 5
Summary results of the best fitted model for the three different surrogate frameworks considered.

	Univariate MLP	Multivariate (*speed volume*) MLP	Multivariate (*speed spatial information*) LSSVM
MAE	2.740	2.186	2.060
max AE	17.683	11.913	26.507
MSE	14.733	8.603	14.612
RMSE	3.838	2.933	3.823
MAPE	0.034	0.022	0.028
max APE	0.412	0.222	0.617
RRSE	0.240	0.172	0.257

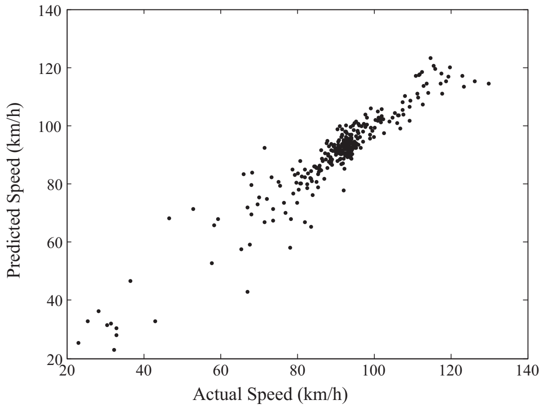


Fig. 6. Actual versus predicted travel speed values for the best fitted model (Ensemble model) for the univariate framework.

150 samples that, in the specific application, are drawn from approximately 3 sampling attempts. The surrogate model with volume as exogenous input variable has a distinct behavior; an average of less than 5% accuracy is reached with very limited data as opposed to the rest of the models. The training and optimization of models is below 20 min for one day data sampling every 3 h. Evidently, the computational time for optimization and learning will significantly reduce, in the case of optimization using a single measure (e.g. MAPE) in the objective function.

Table 5 shows the results for the best fitted models. In the case where only speed or speed and volume information is used, the best fitted model is the MLP with 8 and 9 hidden units respectively, whereas in the case of spatial speed information the best fitted model is the LSSVM with RBF kerner. High performance solutions are achieved in all three models, a finding also supported by the scatter plots of actual versus predicted speeds for the three surrogate models developed (Figs. 6 and 7).

Although the approach followed based on surrogate modeling may result in high accuracy models using a small set of data (one day traffic data) and acceptable computational time, there is no guarantee that the presented modeling framework is uniquely identified and there does not exist a different solution – in terms of the models included and the optimization of their parameters – that may produce similar results. This is an inherent feature of the models and optimization techniques utilized when compared to the traditional modeling approaches.

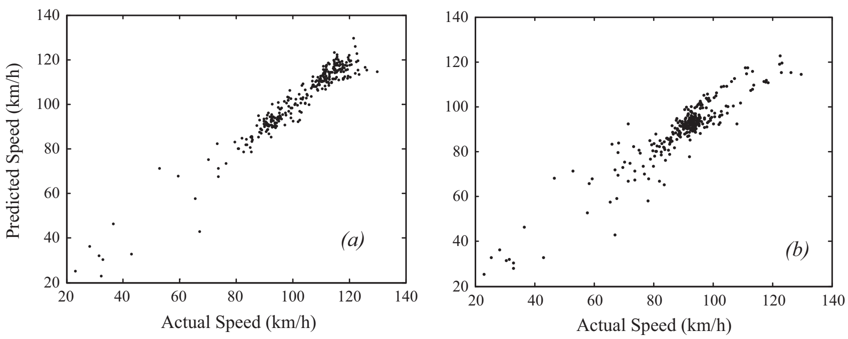


Fig. 7. Actual versus predicted travel speed values for the best fitted model (MLP model) for the multivariate framework with (a) volume and (b) spatial information.

Moreover, another issue that should be stressed is the interplay between problem representation and the complexity of optimization. Traffic flow is a complex phenomenon, especially when studied in very short intervals (shorter than one hour). According to literature, the proper representation of the phenomenon should include spatio-temporal multivariate data, as well as exogenous information on influential factors (e.g. weather) (Vlahogianni et al., 2014). The extent to which surrogate models that encompass full information on the above factors can be optimized and converge to a result is an issue that should be further researched. Nevertheless, the application presented demonstrates the usefulness of surrogate modeling and sampling techniques to the development of parsimonious and efficient transportation forecasting models.

5. Conclusions

In view of the new "data driven" world, there is a need to address the forecasting problem through a more efficient and computationally reasonable manner. To this end, the proposed approach attempts to provide an automated tool for model selection and hyperparameter optimization based on computationally intelligent approaches. The methodology is based on surrogate modeling and jointly considers different modeling types that are heterogeneously developed and optimized using genetic algorithms with the aim to minimize two different error measures (MAPE and RMSE). Moreover, a sampling approach is implemented to reduce the computational burden of model selection and multi-objective optimization.

The approach is implemented on time series of freeway travel speeds. Different models are developed to assess the effectiveness of the approach under a univariate (only speed) and multivariate (volume and speed from different locations) framework. Results show that surrogate modeling can provide accurate predictions in both a univariate and multivariate framework in significantly reduced time using one week data. Improvement over 20% in prediction accuracy is achieved in the multivariate framework with volume information; this may indicate that a more thorough investigation of the spatio-temporal traffic patterns may significantly improve predictions using surrogate modeling.

From a conceptual standpoint, the approach may be implemented to various transportation problems and provide predictions introducing exogenous variables and complex spatio-temporal dependencies. From a methodological standpoint, the proposed approach provides the ability to automatically perform model selection and hyperparameter optimization by avoiding multiple exhaustive runs based on a given dataset. Moreover, the approach is flexible and generic in that, it does not depend on data characteristics and may "naturally" incorporate various types of models, optimization techniques and sampling methods in a multi-objective framework. Nevertheless, as the optimization is based on a nature inspired stochastic approach, full determinism in the solution and convergence stability is not guaranteed. The proposed approach may be extended to a computationally inexpensive holistic decision making tool for the modelers to assess the predictability of transportation phenomena.

Acknowledgements

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) – Research Funding Program: THALES. Investing in knowledge society through the European Social Fund. The data used in this paper are a courtesy of Attica Tollway Operations Authority.

References

Adeli, H., 2001. Neural networks in civil engineering: 1989–2000. Comput.-Aid. Civ. Infrastruct. Eng. 16, 126–142.
Chan, K.Y., Dillon, T., Chang, E., Singh, J., 2013. Prediction of short-term traffic variables using intelligent swarm-based neural networks. IEEE Trans. Control Syst. Technol. 21 (1), 263–274.
Clemen, R.T., 1989. Combining forecasts: a review and annotated bibliography. Int. J. Forecast. 5, 559–583.
Crombecq, K., De Tommasi, L., Gorissen, D., Dhaene, T., 2009. A novel sequential design strategy for global surrogate modeling. In: Winter Simulation Conference. Winter Simulation Conference. pp. 731–742 (December)
Feng, T., Timmermans, H.J., 2014. Trade-offs between mobility and equity maximization under environmental capacity constraints: a case study of an integrated multi-objective model. Transport. Res. Part C: Emerg. Technol.
Forrester, A., Sobester, A., Keane, A., 2008. Engineering Design via Surrogate Modelling: A Practical Guide. John Wiley & Sons.
Gorissen, D., Couckuyt, I., Demeester, P., Dhaene, T., Crombecq, K., 2010a. A surrogate modeling and adaptive sampling toolbox for computer based design. J. Mach. Learn. Res. 11, 2051–2055.
Gorissen, D., Couckuyt, I., Laermans, E., Dhaene, T., 2010b. Multiobjective global surrogate modeling, dealing with the 5-percent problem. Eng. Comput. 26 (1), 81–98.
Jin, Y., 2011. Surrogate-assisted evolutionary computation: recent advances and future challenges. Swarm Evol. Computat. 1 (2), 61–70.
Jin, Y., Sendhoff, B., 2008. Pareto-based multiobjective machine learning: an overview and case studies. IEEE Trans. Syst. Man Cybern. C Appl. Rev. 38 (3), 397–415.
Karlaftis, M.G., Vlahogianni, E.I., 2011. Statistics versus neural networks in transportation research: differences, similarities and some insights. Transport. Res. Part C: Emerg. Technol. 19 (3), 387–399.
Liu, M., Singh, H.K., Ray, T., 2014. Application specific instance generator and a memetic algorithm for capacitated arc routing problems. Transport. Res. Part C: Emerg. Technol.
Rasouli, S., Timmermans, H., 2012. Uncertainty in travel demand forecasting models: literature review and research agenda. Transport. Lett. 4 (1), 55–73.
Song, H., Li, G., 2008. Tourism demand modelling and forecasting—a review of recent research. Tourism Manage. 29 (2), 203–220.
Tan, M.-C., Wong, S.C., Xu, J.-M., Guan, Z.-R., Zhang, P., 2009. An aggregation approach to short-term traffic flow prediction. IEEE Trans. Intell. Transp. Syst. 10 (1), 60–69.
Teodorović, D., 2008. Swarm intelligence systems for transportation engineering: principles and applications. Transport. Res. Part C: Emerg. Technol. 16 (6), 651–667.

Tselentis, D.I., Vlahogianni, E.I., Karlaftis, M.G., 2015. Improving short-term traffic forecasts: to combine models or not to combine? IET Intel. Transport Syst. 9 (2), 193–201.
van Lint, J.W.C., van Hinsbergen, C.P.I.J., 2012. Short term traffic and travel time prediction models. In: Sadek, S., Chowdhury, R. (Eds.), Artificial Intelligence Applications to Critical Transportation Issues, Transportation Research Circular, Number E-C168, November. National Academies Press, Washington, DC.
Vlahogianni, E.I., 2009. Enhancing predictions in signalized arterials with information on short-term traffic flow dynamics. J. Intell. Transport. Syst.: Technol., Plann., Operat. 13 (2), 73–84.
Vlahogianni, E.I., Golias, J.C., Karlaftis, M.G., 2004. Short-term traffic forecasting: overview of objectives and methods. Transport. Rev. 24 (5), 533–557.
Vlahogianni, E.I., Karlaftis, M.G., Golias, J.C., 2005. Optimized and meta-optimized neural networks for short-term traffic flow prediction: a genetic approach. Transport. Res. Part C: Emerg. Technol. 13 (3), 211–234.
Vlahogianni, E.I., Karlaftis, M.G., Golias, J.C., 2006. Statistical methods for detecting nonlinearity and non-stationarity in univariate short-term time-series of traffic volume. Transport. Res. Part C: Emerg. Technol. 14 (5), 351–367.
Vlahogianni, E.I., Karlaftis, M.G., Golias, J.C., 2008. Temporal evolution of short-term urban traffic flow: a non-linear dynamics approach. Comput.-Aid. Civ. Infrastruct. Eng. 22 (5), 317–325.
Vlahogianni, E.I., Karlaftis, M.G., Golias, J.C., 2014. Short-term traffic forecasting: where we are and where we're going. Transport. Res. Part C: Emerg. Technol.
Washington, S., Karlaftis, M.G., Mannering, F.L., 2010. Statistical and Econometric Methods for Transportation Data Analysis, second ed. Chapman & Hall/CRC Press, Boca Raton, FL.
Whitley, D., Rana, S., Heckendorn, R.B., 1999. The island model genetic algorithm: on separability, population size and convergence. J. Comput. Inform. Technol. 7 (1), 33–47.
Yao, B., Hu, P., Lu, X., Gao, J., Zhan, M., 2014. Transit network design based on travel time reliability. Transport. Res. Part C: Emerg. Technol.
Zheng, W., Lee, D.-H., Shi, Q., 2006. Short-term freeway traffic flow prediction: Bayesian combined neural network approach. J. Transport. Eng. 132 (2), 114–121.