Travel Time Estimation Results with Supervised Non-parametric Machine Learning Algorithms

Ivana Cavar, Zvonko Kavran, Ruder Michael Rapajic
Faculty of transport and traffic sciences
University of Zagreb
Zagreb, Croatia
ivana.cavar@fpz.hr; zvonko.kavran@fpz.hr; miso.rapajic@fpz.hr

***Abstract*— Paper describes urban travel time estimation procedure based on non-parametric machine learning algorithms and three data sources (GPS vehicle tracks, meteorological data and road infrastructure data base). After data fusion and dimensionality reduction, new road classification is defined and for four different time intervals and five different road categories travel time estimation is conducted. For travel time estimation, *k* nearest neighbors (kNN) and IRM-based (Iterative Regression Method) approaches were applied. Best results for two hour forecasting period are achieved for road class with highest traffic flow.**

Keywords-GPS vehicle track; travel time estimation; k-nearest neighbours; iterative regression model; urban traffic.

I. Introduction

Travel time estimate is important information applicable in many intelligent transportation system (ITS) services (route guidance, advanced traveler information system or advanced traffic management system) as well as in transportation planning process. It is one of the largest costs of transportation, and travel time savings are often the primary justification for transportation infrastructure improvements. Travel time information is acceptable and useful to all, decision makers, transport system users, technical and nontechnical staff and often used as relevant when comparing different transportation modes.

Today's accessibility of different data sources and large quantity of data has double effect on travel time estimation. It eases up the procedure by allowing different and up to date quality data to be included in estimation. On the other hand it goes under the 'curse of dimensionality' by making calculation process more demanding if even possible. Therefore statistical methods are not applicable in these cases and advanced analytics and data science mechanisms are required.

In literature, different approaches for travel time estimation can be found. Although there are papers dealing with roads that have different characteristics [1][2], best results are achieved for freeways and free flow traffic conditions [5].Reason for this lies in the smaller number of factors influencing the travel time.

Understanding traffic factors affecting travel time is essential for improving prediction accuracy. Some of traffic factors that affect travel time are free flow travel speed [1], occurrence of incident situations, holidays or other uncommon events [2], congestion level [3] and weather conditions [4][5].

Another important element of travel time estimation is the forecasting period (the greater it is the higher is the prediction error) [6].

When it comes to travel time estimation approaches applied in literature, they can be divided in two basic groups [7]:

- Extrapolation models - mainly based on statistical approaches and historical values as regression models [8] [9], ARIMA (Autoregressive Integrated Moving Average) models [20], STARIMA (Space-Time Autoregressive Integrated Moving Average) [10], Kalman filter [11], ANN (artificial neural networks) [12], SVM Support vector machines [13] and pattern based forecasting [14].
- Explanatory models - based mainly of factor or parameters analyses and traffic flow theory as dynamic traffic assignment [15] [16].

Paper has six sections. After introduction, a description of travel time estimation methodology is given. A case study and result comparison are defined, followed by a conclusion.

II. Description of Travel Time Estimation Methodology

Two non-parametric methods, kNN and IRM, are used for travel time estimation and described in more details.

A. K nearest neighbours

K nearest neighbors is an algorithm that is based on similarity metric described by distance function locates the "best" neighbors (the neighbors that are most likely similar to the target). One of the most popular choices to measure this distance is Euclidean. Other measures include Euclidean squared, City-block, and Chebyshev [17]:

$$D(x,p) = \begin{cases} \sqrt{(x-p)^2} & Euclidean \\ (x-p)^2 & Euclidean\ squared \\ Abs\ (x-p) & Cityblock \\ Max\ (|x-p|) & Chebyshev \end{cases} \tag{1}$$

where x and p are the query point and a case from the examples sample, respectively.

The choice of k can be regarded as one of the most important factors of the model that can strongly influence the quality of predicted outcome. The goal is to find an optimal value for k that achieves the right trade-off between the bias and the variance of the model. For travel time estimation purpose, value of k for each road class was determined by cross validation method.

For regression, a kNN prediction is the average of the k-nearest neighbor outcome.

$$y = \frac{1}{k}\sum_{i=1}^{k} y_i \tag{2}$$

where y_i is the ith case of the examples sample and y is the prediction (outcome) of the query point. For classification problems, like road classification, kNN predictions are based on a voting scheme in which the winner is used to label the query.

B. *Iterative regression method*

IRM is iterative multivariate linear regression approach where estimation for dependent continuous variable is based on determination of unknown number of independent variables from set of all proposed predictors.

Estimation procedure is described through following steps:

1. Defining dependent variable y and $z_1, \ldots, z_n$ independent variables.
2. Calculate correlation coefficient between y and each of independent variables $z_1, \ldots, z_n$.
3. Select independent variable with the most significant correlation coefficient.
4. Mark selected independent variable with x_1 and calculate parameters a_1 and b_1 as first iteration by linear regression function:
$$y_1 = a_1 + b_1 x_1 \tag{3}$$
5. Determine the residual r:
$$r_1 = y_1 - (a_1 + b_1 x_1) \tag{4}$$
6. Calculate correlation coefficient between residual (r_1) and each of remaining independent variables.
7. Select independent variable with the most significant correlation coefficient.
8. Mark selected independent variable with x_2 and calculate parameters a_2 and b_2 as second iteration by linear regression function:
$$r_2 = a_2 + b_2 x_2 \tag{5}$$
9. Regression functions from first and second iteration give more precise estimation of dependent variable:
$$y_1 = (a_1 + a_2) + b_1 x_1 + b_2 x_2 \tag{6}$$
10. Repeat steps 5 to 9 until residual r becomes small enough (stops effecting 0,001% of predicted value for y in previous iteration)

C. *Applied travel time estimation procedure*

Methodology proposed for travel time estimation procedure incorporates multiple data analytics techniques. Overall overview of used methods and procedures is given on Figure 1.

First phase incorporates different data collection methods, data fusion and cleansing, as well as data interpolation (when needed). Product of these procedures is input for time series data analysis that results with detection of characteristic time intervals. This is done based on the analysis of speed records. Characteristic time intervals are considered to be time periods in which the vehicle throughput capacity on the observed part of the road is significant. For example, significantly low (to identify the morning and afternoon "peak" periods) or significantly high (free flow conditions).

After this phase, gathered data are used for revision of official road classification. Reason for this lays in the fact that official road classifications are mainly used for general purposes and roads are classified based on infrastructural characteristics. For deep traffic analysis like travel time estimation this classification is not sufficient and new classification is developed. In this paper, classification methodology is used to classify road segments (continuous part of the road from one intersection to the sequent one) based on real traffic flow data collected from GPS tracks as

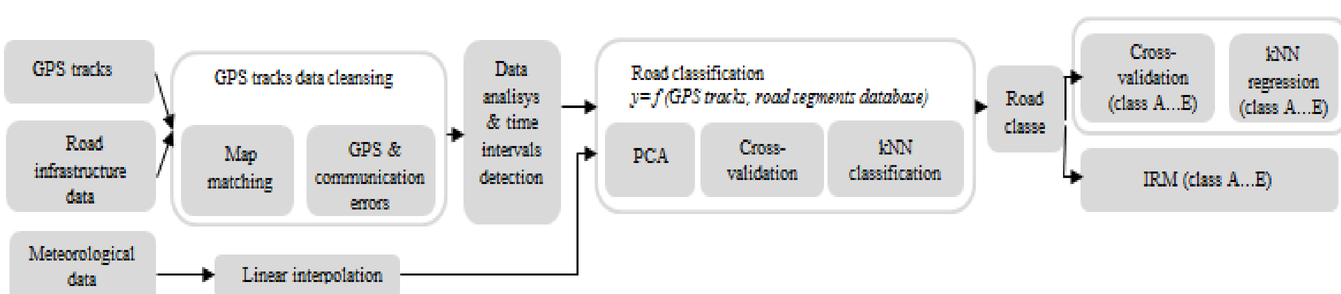

Figure 1. Travel time estimation procedure.

well as road segments infrastructure data. The software used for this procedure is StatSoft Statistica Data Miner. Priori the application of classification technique called kNN principle component analysis (PCA) was applied. The purpose of this was to reduce dimensionality of collected data and to select variables relevant for road classification.

Since kNN algorithm is highly sensitive to the choice of *k*, after the reduction of dimensionality a *v* fold cross-validation procedure is applied to determine k value. Distance measure used in kNN classification is Euclidian distance for standardized data. Data are standardized to transform all values (regardless of their distributions and original units of measurement) to compatible units from a distribution with a mean of 0 and a standard deviation of 1. This makes the distributions of values easy to compare across variables and/or subsets and makes the results entirely independent of the ranges of values or the units of measurements. Results of this phase are road segments separated in different categories. For each of these categories, in next step, different travel time estimation procedure is carried out.

At this point, whole data set is divided into two subsets, one for training and one for model development. Travel time estimation is based on two approaches kNN regression in combination with v fold cross-validation and IRM approach.This is applied separately for each road class and each characteristic time interval.

D. *Advantages and areas of practice*

Developed travel time estimation procedure has multiple advantages for practical application as:

- Supports application of multiple data sources (different sensors);
- Allows automated data fusion and dimensionality reduction;
- Improves estimation procedure by basing estimation steps on real traffic performance data (road classification is based on traffic data, not only on road infrastructure data);
- Once defined values of *k* for each road class (in cross-validation step) doesn't need to be calculated again. Due to the fact that this is the most computational demanding step in travel time estimation procedure, future application of this procedure is much less time and computational expensive and therefore suitable for real time travel estimation:
- Completely automated travel time estimation that can give results in real time.

Travel time is one of most important values in traffic planning and allows decision maker to base its decisions on traffic performance on different routes/modes choice etc. Travel time estimation information can be delivered to drivers to aim them in route selection choice or to be a part of automated route guidance system as well as to be incorporated in different ITS services.

III. TRAVEL TIME ESTIMATION (CASE STUDY: CITY OF ZAGREB)

The travel time estimation procedure we propose in this paper is applied for case study in City of Zagreb and described through description of data collection process, road classification, travel time estimation and evaluation of results.

A. *Data collection process*

Data used for travel time estimation were collected from three sources including 51 835 560 GPS tracks, road segments data and meteorological data from official Meteorological and Hydrological Service. Geographical area covered with data collection process was urban area of City of Zagreb, Croatia and time window for data collection was thirteen months.

B. *Data on the traffic conditions*

Traffic conditions are analyzed based on the GPS vehicle tracks. The data were recorded by the device and sent via mobile network. Data were sent to the local server and stored in the database (Table 1).

TABLE 1. TRAFFIC DATA

Recorded data	**Variable description**
Log time	time of recording expressed in UTC (Universal Time Coordinated) standard
Vehicle ID	identifier of the observed vehicle / GPS device
X coordinate	x coordinate of the GPS record (WGS84 – World Geodetic System 1984)
Y coordinate	y coordinate of the GPS record (WGS84)
Speed	current speed in [km/h]
Course	angle at which the vehicle is travelling with reference to the North
GPS status	indicates the accuracy of the record. GPS status 3 indicates that the data have been collected from 4 or more satellites and GPS status 1 that data have been collected from fewer than 2 satellites.
Engine status	shows whether the vehicle engine was running or was turned off while making the recording.

C. *Data on the road infrastructure*

Road infrastructure data describe road segments and are stored in the form of digital map database (direction, blocked turns, length, marked pedestrian zones, etc.). The elements included in this database are presented in Table 2.

TABLE 2. ROAD INFRASTRACTURE DATA

Record	Variable description
Segment *ID*	identifier of the road segment
Type	numeric code of the road type according to official classification defined in Spatial plan of City of Zagreb
Direction	code representing road direction
Start x	x coordinate of the beginning of segment (WGS84)
Start y	y coordinate of the beginning of segment (WGS84)
End x	x coordinate of the end of segment (WGS84)
End y	y coordinate of the end of segment (WGS84)
Length	length of the segment in [m]
Name	name of road which contains the respective segment

Presentation of road segment data in a form of a digital map is given in Figure 2.

Figure 2. Part of a digital map representing geographic area of City of Zagreb, Croatia.

D. *Meteorological data*

Meteorological data are received from Meteorological and Hydrological Service. This database includes variables described in Table 3.

TABLE 3. METEOROLOGICAL DATA

Meteorological data
Minimum and maximum daily air temperature and their difference;
Daily air temperature
Wet ground temperature and ground condition
Air pressure and air steam pressure
Sun shining period;
Precipitation and relative humidity;
Snow cover thickness and thickness of new snow cover;
Horizontal visibility
Direction and speed of wind
Ground temperature at -2cm, -5cm, -10cm, -20cm, -30cm, -50cm, -100cm
Observatory diary data

More on data collection and cleansing procedure can be found in literature [18].

E. *Road classification*

After the collection and fusion of gathered data and priori the development of travel time estimation, road classification has been revised (official road classification in City of Zagreb is defined by Spatial plan of City of Zagreb [19]).

As a result of the procedure proposed on Figure 1, in Road classification box, a set of four distinctive variables is selected:

- mean average speed on road segment,
- speed's standard deviation on road segment,
- length of road segment,
- vehicle count of each road segment.

Based on this, for road classification 28NN model [5] is used (*k*-nearest neighbors model that for segment classification takes into account neighborhood of 28 most similar road segments). Results of classification procedure separates roads in five classes marked as class A, class B, class C, class D and class E. Compared to official road classification that contains also five road classes (highways, fast roads, state, county and local roads) that are divided based on their construction characteristics and the level of authority in charge of their maintenance this way achieved results are distinct in 47.22 % cases.

More on extracted road classes is given in Table 4, where number and share of segments included in each class, overall length of segments and their share in total road network length, number of records/segment length, mean average of number of records on each segment, vehicle speed mean average value, vehicle speed standard deviation and average segment length for each class are presented. Based on this data we can see that class E includes largest number of segments (around 64 %) and larges share (although somewhat lower than segment count share). By comparison of length and segments shares it is visible that class A contains smallest amount of segments but they have lower number of intersections and higher lengths then segments in others classes (length share/segments share for class A equals 3.07). Also, 2.96% of all segments belong to the class A, 11.76% to classes A and B and 30.36% to classes A, B and C. These classes have highest traffic flow and 94.7% of all records were recorded there. Based on average number of records for each class we can see that number of records per segment length decreases from road class A to road class E in non-linear manner. Mean average of number of records on each segment, vehicle mean average speed and its standard deviation as well as average length of road segment decreases from road class A to road class E. And while average speed for class A represents double average speed of class E deviation is increased. We can say that road class A while having highest traffic flow and average vehicle speed also has largest discontinuity in vehicle movements. Respectively, class E has lowest vehicle speed deviation and most continuous traffic flow.

More on road classification based on hybrid approach can be found in literature [5].

 ISBN: 978-1-61208-242-4

TABLE 4. ROAD SEGMENTS DATA

Raoad class	Number of segments	Segments share [%]	Length of road segments in class [km]	Length share [%]	Number of record / segment length [m]	Mean average of number of records on each segment	Vehicle speed mean average value	Vehicle speed standard deviation	Average segment length [m]
Class A	392	2.96	114 210	9.08	26	7664	62.33	15.91	292.1
Class B	1 166	8.8	175 613	13.96	15.7	2357	49.68	14.18	150.6
Class C	2 463	18.6	314 139	24.97	8.8	1123	42.19	12.28	127.5
Class D	714	5.39	67 151	5.34	4.1	386	41.52	11.49	94.1
Class E	8 508	64.25	586 902	46.65	3.5	240	31	10.01	69

IV. TRAVEL TIME ESTIMATION

The travel time estimation procedure is described through definition of characteristic time intervals and application of kNN and IRM methods.

A. *Definition of characteristic time intervals for travel time estimation*

For the purpose of travel time estimation characteristic, time intervals should be identified. A presentation of one such analysis is given in Figure 3 for segment Slavonska Avenue, direction East-to-West. Figure 3 clearly shows the morning congestion which is most expressed in the interval between 06:00 h and 08:00 h and the afternoon congestion which is a bit longer, taking from 15:30 h -18:30 h.

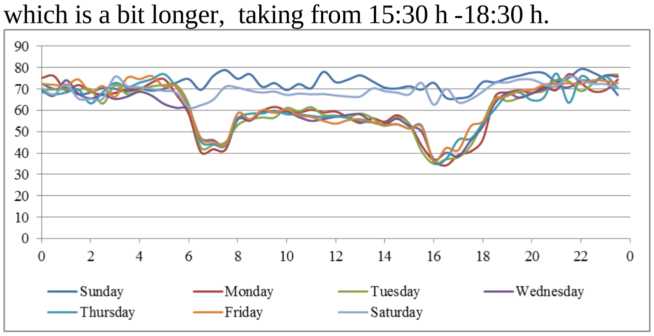

Figure 3. Segment speed analysis

Based on such analysis characteristic time intervals for travel time estimation are defined as follows:

- Morning peak interval (06:00-08:00 h);
- Noon free flow interval (12:00-13:00 h);
- Afternoon peak interval (15:30-18:30 h);
- Evening free flow interval (21:00 – 22:00 h).

B. *Urban travel time estimation procedure*

At this point, whole data set is divided into two subsets. Data collected during first eleven months represents training set for model development. Data collected during last two months are removed from model development procedure and will serve for the travel time estimation error testing.

C. *KNN based approach*

For each road class one road, belonging to that class, is selected to serve as demonstrative one for results of travel time estimation procedure for selected time intervals. These roads are defined in Table 5.

TABLE 5. SHOW UP ROADS FOR EACH ROADS CLASS

Road class	Demonstrative road
A	Slavonska avenue
B	Selska road
C	Savska road
D	Prisavlje
E	Jordanovac

Results of *v* fold cross-validation are presented on Figures 4, 5, 6, 7 and 8, as well as in Table 5.

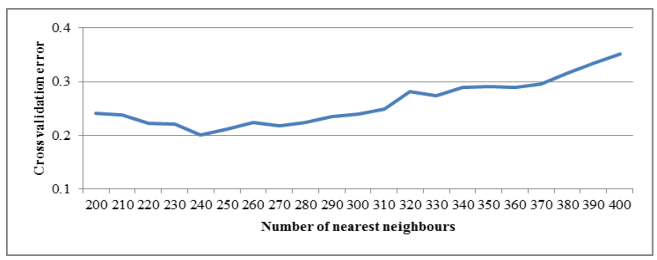

Figure 4. Results of v fold cross validation for road class A

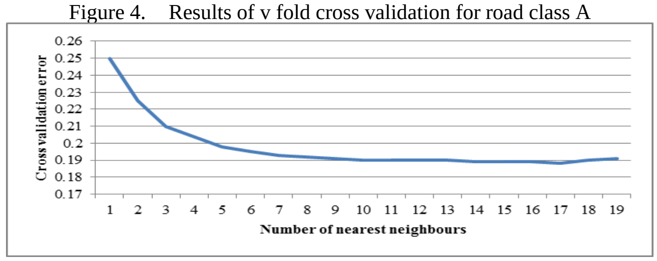

Figure 5. Results of v fold cross validation for road class B

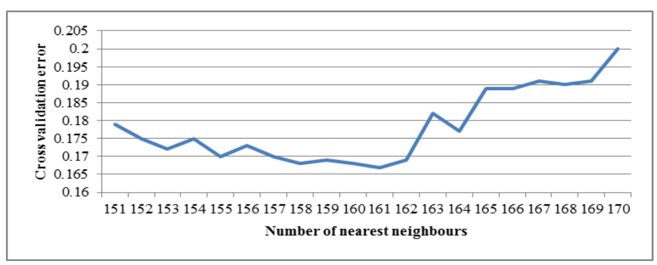

Figure 6. Results of v fold cross validation for road class C

 ISBN: 978-1-61208-242-4

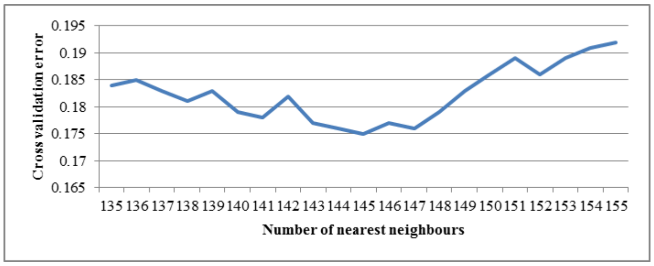

Figure 7. Results of v fold cross validation for road class D

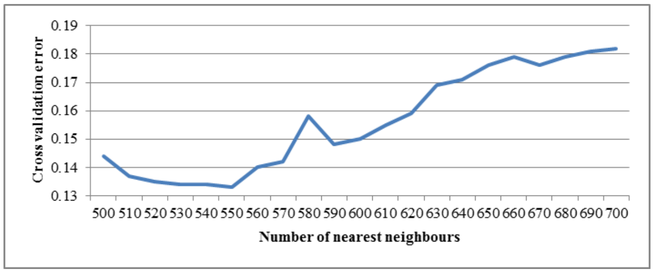

Figure 8. Results of v fold cross validation for road class E

On Figures 4-8, for each road class, computed errors are shown for number of nearest neighbor around optimal value for *k* (optimal in a cross-validation sense). Search for optimal value of *k* was done in phases and figures 4-8 are displaying only results of phase in which optimal *k* was selected for each road class.

D. IRM based approach

Based on previously described procedure, for each road class IRM function is determined and given below:

TABLE 6. IRM FUNCTIONS FOR EACH ROAD CLASS

Class A
$y= 12.85423 + 0.99152*x_1 + 0.09618 * x_2 - 0.0603* x_3 -0.2758*x_4 - 0.0702*x_5 -0.3456*x_6 -0.005*x_7+0.00136*x_8 +0.000074*x_8 +0.02841*x_9-0.0001*x_{10}-0.00004*x_{11}-6.697*x_{12}--0.0404*x_{13}- 0.0453*x_{14}+0.0292*x_{15}-0.0009*x_{16}+0.10499*x_{17}+0.04875*x_{18}$
Class B
$y= 438.20584 + 1.0033*x_1 + 0.00382 * x_2 - 25.58* x_3 -1.442*x_4 - 0.0516*x_5 -0.0519*x_6 -0.0006*x_7+0.63512*x_8 +0.0.12485*x_9 - 0.00258*x_{10}$
Class C
$y= -914.62584 + 1.0033*x_1 + 0.0897 * x_2 - 0.0508* x_3 -0.0216*x_4 +0.21456*x_5 -0.0663*x_6 -0.0498*x_7-0.1783*x_8 -0.04844*x_9 - 0.4439*x_{10} +0.0434*x_{11}$
Class D
$y= 33441.848 + 0.99755*x_1 + 0.01721 * x_2 - 0.0272* x_3 -730.4*x_4 +0.19136*x_5 -0.4659*x_6 -0.0384*x_7$
Class E
$y= 59963.37591 -1405*x_1 + 0.71646* x_2 - 817.8* x_3 -0.55613*x_4 - 0.2512*x_5 -0.00008*x_6 +0.02573*x_7-0.0171*x_8$

Table 6 incorporate variables from all three data sources proving this way their justification as model inputs.

V. URBAN TRAVEL TIME ESTIMATION RESULTS

Based on defined steps, different kNN and IRM regression procedures were applied on data set for each road class. From testing data set a number of vehicle tracks for each time interval is selected and compared with achieved results. Based on this data MAPE (mean absolute percentage error) is calculated for each road class and each time interval defined:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A_i - F_i}{A_i}\right| \tag{7}$$

where:

n – number of observations,

A_i – the actual value of recorded travel time for observation *i*,

F_i – the forecast value of travel time for observation *i*.

Described results are presented on Figure 9 for morning peak period, Figure 10 for noon time period, Figure 11 for afternoon peak period, and Figure 12 for evening period. For a morning peak period best results are achieved for road class E and IRM method, while for noon time interval best results are achieved for road class A and kNN method. For afternoon time interval best results are achieved for road class D with IRM method, while the same method results in highest MAPE for road class A. Results of evening time interval show that the best results are scored by kNN method for a road class D and the worst ones for road class C by IRM.

Figure 9. Results of morning peak period

Figure 10. Results of noon period

Figure 11. Results of afternoon peak period

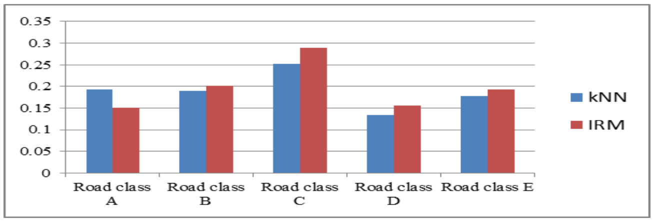

Figure 12. Results of evening period

Method with better results for defined road class and time interval is used for real time travel time estimation and results are presented in Table 7.

TABLE 7. SIMULATION RESULTS FOR EACH ROAD CLASS

Road class	MAPE
Class A	0.043197
Class B	0.188525
Class C	0.089705
Class D	0.052738
Class E	0.141764

VI. CONCLUSIONS

Based on results achieved from case study in City of Zagreb, we can see that GPS tracks, meteorological data and data on road infrastructure present good starting point for urban travel time estimation.

In a case of large data sets application of traditional statistical approach is not advised due to the computationally demanding process of travel time estimation. Dimensionality reduction applied in this paper resulted in reduction of 58 input variables into the average of 21 of them mainly reducing variables with high correlations while keeping maximum data variability contained in original data set. For example, it was evident that data on snow cover thickness and thickness of new snow cover are highly correlated. Data on new snow cover had higher influence on speed deviations and therefore remained in travel time estimation data set after the dimensionality reduction step. Also, data on direction and speed of wind had poor influence on travel time estimation and were mainly removed during dimensionality reduction step. Relative humidity and horizontal visibility were correlated in significant manner together with *y* coordinate of the record during autumn/winter period. This can be explained through influence of river Sava and often fogs in surrounding areas during the cold weather. Since *y* coordinate yield significant correlation with travel time only in this context it was often omitted during dimensionality reduction phase together with horizontal visibility variable while *x* coordinate proved to have high influence on the value of travel time. Beside *x* coordinate, road segment length, wet ground temperature, average speed, speed deviation, log time and daily air temperature remained in data set after dimensionality reduction step for each road class. As expected, values that identified road/segment as road name, segment ID, beginning and end coordinates had very high correlation and were mainly presented by segment ID variable after dimensionality reduction step. Also, it should be noted that vehicle ID remained as significant variable in data set proving the influence of driving characteristics of each driver to the travel time value.

For reduced data sets, kNN and IRM methods results were compared. Based on the achieved results, real time forecasting procedure is applied with forecasting interval of 2 hours. Best results are achieved for road class A (mainly freeways) with MAPE 0.043197, followed with road class D and C. Lowest MAPE is achieved for road class B and it is 0.188525.

Comparison of results to others research is very difficult due to the fact that there is no research that uses same road classification technique or definition of characteristic estimation time interval. Therefore comparison will be made based on most similar values for characteristic time interval and/or road class.

When compared to the results achieved in literature [20] for afternoon peak hour, road length of 2.9 km, desired speed of 64 km/h and without consideration of a long queue that based on this data would correspond to the road class A from our model the error achieved by model proposed in this paper is twice smaller (4.3%) than average error from literature (8.3%).

When results are compared with Google maps [21] travel time estimate for a road class C in length of 2.317 km and afternoon peak period, the achieved result by Google maps is 6 minutes (360 sec) while model proposed in this paper gave estimated travel time for this input data to be 477 seconds. Travel time value measured by GPS record was 476 seconds.

Compared to results for morning peak period MAPE values of three algorithms from literature [22], point-detection-based algorithm (10.6 %), probe-based algorithm with 5% probe rate (10.8%) and adaptive Kalman filter algorithm with 5% probe rate (7.6%), are higher than average MAPE value for all road classes achieved by model proposed in this paper. Average MAPE for all road classes achieved by proposed model equals 5.02% for morning peak period.

ACKNOWLEDGMENT

Authors are grateful to Mireo d.o.o. for GPS vehicle tracks and digital map provision.

Authors are grateful to Meteorological and Hydrological Service for data provision.

REFERENCES

[1] Lum, K.M., Fan, H. S.L., Lam, S.H., and Olszewski, P. (1998) Speed-Flow Modeling of Arterial Roads in Singapore, Journal of Transportation Engineering, Vol. 124, no. 6, pp. 213-222. doi: 10.1061/(ASCE)0733-947X(1998)124:3(213) [retrieved: July, 2012]

[2] J.W.C. van Lint. (2004) Reliable Travel Time Prediction for Freeways. Research School for Transport, Infrastructure and Logistics, Netherlands, 2004.

[3] Richardson, A. J. and Taylor, M.A.P. (1978) Travel time variability on commuter journies, High Speed Ground Transportation Journal, Vol. 12, No.1, pp 77-99.

[4] Chien, S.I-J. and Kuchipudi, C. M. (2003) Dynamic travel time prediction with real-time and historic data, Journal of Transportation Engineering, Vol. 129, No. 6, pp. 608-616., doi: 10.1061/(ASCE)0733-947X(2003)129:6(608) [retrieved: July, 2012]

[5] Cavar, I., Kavran, Z. , and Petrovic, M. (2011) Hybrid approach for urban roads classification based on GPS tracks and road subsegments data, Promet – Traffic&Transportation, 23 (4): pp. 289-296.

[6] Kisgyorgy, L. and Rilett, L.R. (2002) Travel Time Prediction by Advanced Neural Network., Periodica Polytechnica Civil Engineering, Vol. 46, No. 1, pp. 15-32.

[7] Versteegt, H. H. and Tampere, C. M. J. (2005) PredicTime state of the art and functional architecture, TNO Inro report 2003-07, No. 03-7N-024-73196.

[8] Sun, H., Liu, H.X.,Xiao, H. , and Ran,B., (2003) Short term traffic Forecast Using the Local Linear Regression Model. TRB paper no. 03-3580. Transportation Research Board, 82th annual meeting, January 2003., electronic proceedings.

[9] Smith, B.L., Williams, B.M., and Oswald, R.K. (2002) Comparison of parametric and nonparametric models for traffic flow forecasting. Transportation Research, Part C, Vol. 10 (2002), pp. 303-321. doi: 10.1016/S0968-090X(02)00009-8 [retrieved: July, 2012]

[10] Kamarianakis, Y. and Prastacos, P. (2003) Forecasting traffic flow conditions in an urban network: Comparison of multivariate and univariate approaches, Transportation research record, vol. 1857, pp. 74-84.

[11] Zhou,C. and Nelson, P.C. (2002) Predicting Traffic Congestion Using Recurrent Neural Networks, 9th World Congress on ITS. Chicago, electronic proceedings.

[12] Tao, Y. and Ran, B. (2006) An Application of Neural Network in Corridor Travel Time Prediction in the Presence of Traffic Incidents, Applications of Advanced Technology in Transportation, 2006., Chicago, Illinois, USA, pp. 455-460, doi 10.1061/40799(213)72) [retrieved: July, 2012]

[13] Vanajakshi, L. and Rilett, L. R. (2007) Travel Time Prediction Using Support Vector Machine Technique, IEEE Intelligent Vehicles Symposium, Istanbul, Turky, pp. 600 - 605.

[14] Wild, D. (1994) Pattern-based Forecasting, Proceedings of the Second DRIVE-II Workshop on Short-Term traffic forecasting, Delft, pp 49-63.

[15] Ben-Akiva, M.E:, Bierlaire, M., Didier, B., Koutsopoulos, H.N., and Rabi, M., (2006) Simulated-based tools for dynamic traffic assignment: DynaMIT and applications, ITS America 10th annual Meeting, Boston, USA, electronic proceedings.

[16] Ben-Akiva, M.E:, Bierlaire, M., Didier, B., Koutsopoulos, H.N., Rabi, M., (2006) DynaMIT: a simulation-based system for traffic prediction, DACCORD Short-term forecasting workshop, Delft, The The Netherland, pp. 19-23.

[17] Webb, A.R.: Statistical Pattern Recognition, John Wiley & Sons, Ltd., 2002.

[18] Ćavar, I., Marković, H., and Gold, H., 2006. GPS vehicles tracks data cleaning methodology, [CD. ROM] Proceedings of the xth ICTS 10th International Conference on Traffic Science.

[19] Spatial plan of City of Zagreb, Prostorni plan grada Zagreba, Službeni glasnik, Zagreb 2009

[20] Liu, H.C., Ma, W., Wu, X., and Hu, H., (2009) Real-Time Estimation of Arterial Travel Time under Congested Conditions, Transportmetrica, Volume 8, Issue 2, 2012., doi: 10.1080/18128600903502298 [retrieved: July, 2012]

[21] http://maps.google.com/ [retrieved: june, 2012]

[22] Chu, L., Oh, J.-S., and Recker, W., Adaptive Kalman filter based freeway travel time estimation, 84th TRB Annual Meeting, Washington D.C., January 9-13 2005. Transportation Research Board. http://www.www.clr-analytics.com/Files/Adaptive%20Kalman%20Filter%20Based%20Freeway%20Travel%20time%20.pdf [retrieved: june, 2012]