Time series and support vector machines to predict Powered-Two-Wheeler accident involvement and accident type

Theofilatos, A., Yannis, G., Antoniou, C., Chaziris, A., Serbis, D.

Abstract

Predicting and explaining road accident probability by exploiting high resolution traffic data has been a continuously researched topic in the last years. However, there is no specific focus yet on Powered-Two-Wheelers. Furthermore, urban arterials have not received adequate attention so far, since the vast majority of relevant literature concerns freeways. This study aims to contribute to the current knowledge by investigating the possibility of utilizing support vector machine (SVM) models for predicting Powered-Two-Wheeler (PTW) accident involvement and PTW accident type on urban arterials. The proposed methodology is applied on the basis of both original and transformed time series of real-time traffic data collected from urban arterials in Athens, Greece for the years 2006-2011. Findings suggest that that PTW accident involvement and PTW accident type can be adequately defined by the prevailing traffic conditions. When predicting PTW accident involvement, the original traffic time series performed better than the transformed time series. However, in a few cases, the difference between original time series and transformed time series was negligible. On the other hand, when PTW accident type is investigated, neither of the two approaches clearly outperformed the other. The results of the study indicate that the combination of SVM models and time-series data can be used for road safety purposes especially by utilizing real-time traffic data.

Keywords: Powered-Two-Wheelers, accident type, real-time data, time series, support vector machines

1. Introduction and background

Road accidents are a serious burden to society. Annually, there are 1.25 million fatalities, while half of fatalities on the world's roads are "vulnerable road users": pedestrians, cyclists and motorcyclists (WHO, 2015). Therefore, road safety is a major concern for societies, as accidents impose serious problems to societies in terms of human costs, economic costs, property damage costs and medical costs.

A significant increase in motorcycling activities is observed in many countries worldwide during the last years. Over the last twenty years, the number of mopeds and motorcycles together referred to as Powered-two-wheelers (PTWs) in Europe has almost doubled (Yannis et al., 2010). This shift in mode choice is likely to be attributed to economic, mobility and flexibility benefits offered by PTWs. Furthermore, it should be considered that PTW fatalities accounted for 18% of the total number of road accident fatalities in 2013 in the European Union-23 (EU-23) countries, (ERSO, 2015).

The majority of moped fatalities occurred in urban areas whereas the majority of motorcycle fatalities occurred in rural areas (ERSO, 2015). Per vehicle mile travelled, motorcycle riders have a 34-fold higher risk of death in an accident than the other motor vehicles users (Lin and Kraus, 2009).

Huge efforts have been made by researchers to explain PTW accident risk. Numerous PTW accident-related factors have been identified in international literature. For example, the road environment such as road type, road geometry and roadside installations have been found to have an influence on PTW accident occurrence (Harnen et al., 2003; Wanvik, 2009). Haque et al. (2009) found that several geometrical and environmental factors were linked with non-at fault crashes of motorcyclists. Schneider IV et al. (2012) stated that younger motorcyclists, riders under the influence of alcohol (DUI), riders without insurance or not wearing helmet are more likely to be at-fault in a crash. For a more complete list of PTW relevant risk factors, the reader is encouraged to refer to Vlahogianni et al. (2012) and Theofilatos and Yannis (2015).

The impact of traffic characteristics on PTW safety has not been investigated in a large extent. Various studies have addressed the effect of traffic on vehicle accidents but the literature regarding PTW accidents is limited (Abdul Manan and Várhelyi, 2012; Sharma et al., 2013). The investigation of PTW safety in relation to traffic characteristics on a real-time basis is considered highly important due to the vulnerability of PTWs, but also due to the conflicting interactions with other vehicles on the road, which complicates the understanding PTW drivers' behaviour (Barmpounakis et al., forthcoming). Recently, there is a recent trend in predicting and explaining road accident occurrence with real-time traffic data (Abdel-Aty and Pande, 2005; Abdel-Aty et al., 2007; Ahmed and Abdel-Aty, 2012; Yu and Abdel-Aty, 2013a). However, to the best of our knowledge, there are no relevant studies dedicated to PTWs.

Accident type (referred also as collision type or crash type) is identified as another important parameter with a significant role in road safety, as underlined by a number of studies (Kim et al., 2006; Pande and Abdel-Aty, 2006). However, various types of collisions are generally not distinguished in most studies, possibly because of the difficulties in collecting the necessary data (Christoforou et al., 2011). Christoforou et al. (2011), also stressed that in most of the studies that considered different accident types, a distinction between single- and multi-vehicle accidents is made. The majority of these studies utilize aggregated traffic data (Ceder and Livneh, 1982; Zhou and Sisiopiku, 1997). However, some studies exploit real-time data (Abdel-Aty and Pande, 2005; Lee et al., 2006; Pande and Abdel-Aty, 2006). For example, Abdel-Aty and Pande and (2005), exploited real-time traffic from the I-4 corridor in Orlando and attempted to investigate the factors determining the accident type. The authors found that variation in speed 10 to 15 minutes prior to an accident is among the most significant factors. Golob et al. (2008), stated that congestion had a considerable influence on vehicle involvement. For example, the authors mention that congestion in the left and interior lanes distinguishes single- from multi-vehicle accidents.

Summing up, Powered-Two-Wheeler safety has not received significant attention when real-time traffic data are used. Moreover, there are a few more gaps of knowledge in terms of data. First, the vast majority of relevant international literature exploited real-time traffic data from freeways and not from urban roads. Secondly, time series of real-time traffic data were not extensively utilized in investigating road accidents so far.

In terms of methodology, traffic time series could be exploited by means of predictive models such as k-NN, Support Vector Machines (SVMs) and so on (Zhao et al., 2012). Support Vector Machines in specific, is a relatively new statistical method used for predictive purposes in road safety (Li et al. 2008; Yu and Abdel-Aty, 2013b). Moreover, as Li et al. (2008) suggest, the assessment of SVM models' performance when only traffic flow is considered, should be implemented

Consequently, the research presented in this paper aims to add to the current knowledge by investigating the possibility of utilizing support vector machine (SVM) models for predicting Powered-Two-Wheeler (PTW) accident involvement and PTW accident type on urban arterials, by exploiting real-time traffic data. It is noted that the term "predicting PTW accident involvement", refers to the prediction of whether a PTW is involved or not in an accident that has occurred.

2. Methodology

In order to predict PTW involvement and PTW accident type a combined methodological approach was followed. More specifically, a time series classification was performed by applying SVMs (Support Vector Machines), which is a very powerful and advanced machine learning technique. Firstly, the SVMs were applied by utilizing the original time series data and secondly by utilizing the DWT (Discrete Wavelet Transform) transformed data. Lastly, the results are compared.

2.1. Time series approach

2.1.1. Original time series data

Time series classification is used when it is desired to build a classification model based on labelled time series and then use the model to predict the label of unlabelled time series. Classification of unlabelled time series to existing classes is a further traditional data mining task. By "labelled time series", it means that a training dataset with correctly classified observation is used, and then the built models are used to predict the labels of a test dataset (Kleist, 2015).

It is possible to extract new features from time series in order to potentially improve the performance of classification models. There are various such techniques for feature extraction such as the Singular Value Decomposition (SVD), Discrete Fourier Transform (DFT), Discrete Wavelet Transform (DWT), Piecewise Aggregate Approximation (PAA), Perpetually Important Points (PIP), Piecewise Linear Representation and Symbolic Representation (Zhao, 2012).

In this approach, the original time series data are used, namely data which have been sampled at equispaced points in time, without applying any techniques for feature extraction.

2.1.2. Wavelet transform

The Wavelet transform provides a multi-resolution representation using wavelets. In this chapter, a Discrete Wavelet Transform (DWT) is used to extract features from time series and then build a classification model (Burrus et al., 1998). The time series and its transform can be considered to be two representations of the same mathematical entity.

The very name wavelet originates from the requirement that they should integrate to zero, “waving" above and below the x-axis (Vidakovic and Mueller, 1991). A DWT is any wavelet transform for which the wavelets are discretely sampled and is an orthornormal transform. As with other wavelet transforms, a key advantage it has over Fourier transforms is temporal resolution: it captures both frequency and location information (location in time).

McLeod et al., (2012) provide a very good description of Discrete Wavelet Transform. A time series of dyadic length is considered z_t,t= 1,......n, where n = 2^J. The discrete wavelet transformation (DWT) decomposes the time series into J wavelet coefficients vectors, $W_{i,j}$= 0,......,J-1 each of length n_j= 2^{J-j}, j= 1,......,J plus a scaling coefficient V_J. Each wavelet coefficient is constructed as a difference of two weighted averages each of length $\lambda^j = 2^{j-1}$. Similarly to Discrete Fourier Transformation, the DWT provides an orthonormal decomposition, W= WZ, where W′= (W′$_1$,…,W′$_{J-1}$, V′$_{J-1}$), Z= (z_1,......,z_n)′.

There are two functions that play a primary role in wavelet analysis, the scaling function (father wavelet) and the wavelet (mother wavelet). The simplest wavelet analysis is based on Haar scaling function, (Haar Wavelet Transform), (Struzik and Siebes, 1999). The Haar wavelet is a sequence of rescaled "square-shaped" functions which together form a wavelet family or basis. The Haar sequence is recognised as the first known wavelet basis and it was proposed in 1909 by Alfréd Haar (Haar, 1910).

The Haar scaling function φ(x) is defined as:

$$\varphi(x) = \begin{cases} 1, & 0 \le x < 1 \\ 0, & otherwise \end{cases} \qquad \text{(Eq. 1)}$$

The Haar Wavelet’s mother function is then defined as ψ(x)= φ(2x)-φ(2x-1)

$$\psi(x) = \begin{cases} 1, & 0 \le x < 1/2 \\ -1, & 1/2 \le x < 1 \\ 0, & otherwise. \end{cases} \qquad \text{(Eq. 2)}$$

Figure 1 that follows, illustrates an example of a graphical representation of a Haar Wavelet Transform, which is the simplest DWT (Zhao, 2012). Figure 2 shows another graphical example of a Haar Wavelet extracted from Vidakovic and Mueller, (1991).

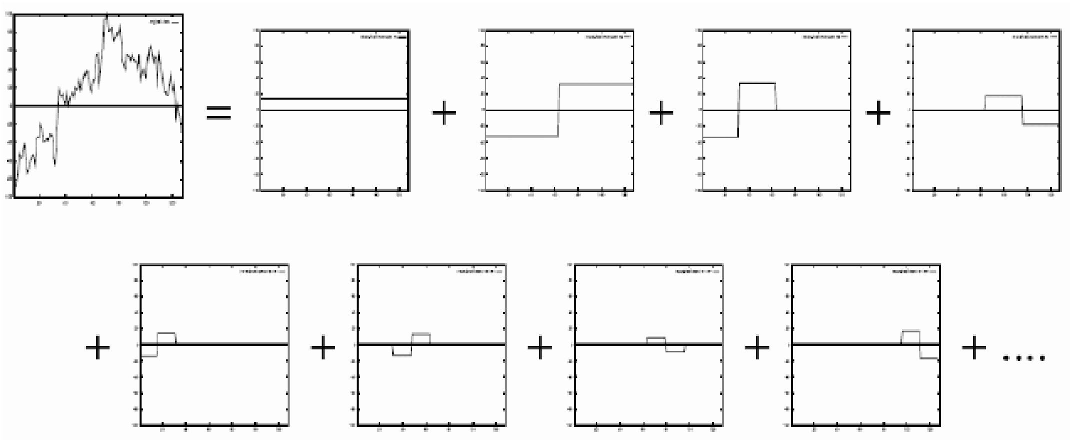

Figure 1: A graphical representation of a simple Haar Wavelet Transform. Source: Zhao, 2012.

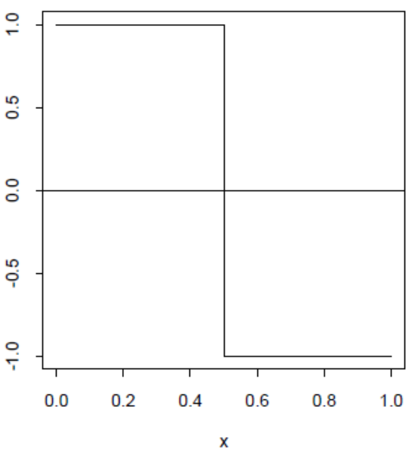

Figure 2: A Haar Wavelet. Vidakovic and Mueller, 1991.

2.2 Support vector machines

Traditional statistical modelling has been widely used for transportation data analysis. However, such approach contains some limitations, for example modelling assumptions that may not always be true. Non-parametric and artificial intelligent methods could then be applied to overcome such limitations.

Support Vector Machines (SVMs), constitute a relatively new modelling technique, which is useful for classification problems (Keckman, 2005). In transportation science, the studies having used SVMs are relatively rare (Li et al., 2008; Li et al., 2012), especially in real-time crash risk evaluation (Yu and Abdel-Aty, 2013b and 2014).

SVMs have originated from statistical learning theory (Vapnik, 1998), and have been developed by Cortes and Vapnik (1995) mainly for binary classification. Basically, when building a SVM model, the aim is the optimal separating hyperplane between two classes by maximizing the margin between the classes' closest points (Meyer, 2001). Therefore, different classes are separated by the hyperplane:

$$\langle w, \Phi(x)\rangle + b = 0 \qquad \text{(Eq. 3)}$$

which corresponds to the decision function

$$f(x) = \ sign(\langle \Phi(x_i\,), w\rangle + b) \qquad \text{(Eq. 4)}$$

The points lying on the boundaries are the support vectors, while the middle of the margin is the optimum separating hyperplane. Figure 3 provides a graphical illustration of a linear separable example of SVMs (Meyer, 2001).

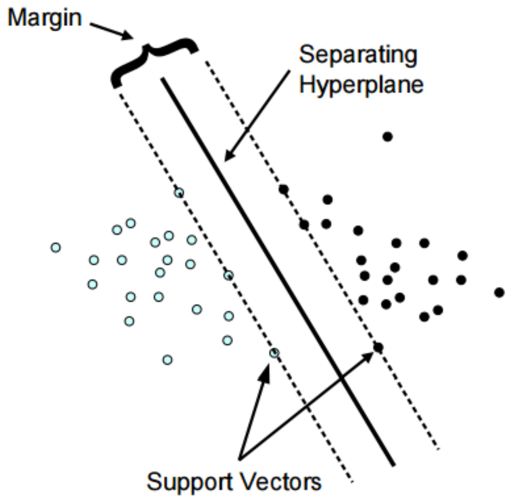

Figure 3: Graphical illustration of SVM classification (linear separable example). Source: Meyer, 2001.

In addition, SVMs can be enhanced to tackle nonlinear classification problems, regression and outlier detection. The major limitation of SVMs is that the models cannot be directly used to identify the relationships between the dependent and the independent variables. Therefore, SVMs can be considered as a "black box" technique. The reader is encouraged to refer to Karatzoglou et al. (2006), where a very detailed description of SVM model formulation is provided for R language.

From all the available kernel-based algorithms (kernels) (e.g. linear, polynomial, Gaussian Radial-basis function and sigmoid), the Gaussian Radial-basis function kernel was considered in this paper (Karatzoglou et al., 2005):

Radial-basis function kernel (RBF):

$$K(x_i, x_j)= \exp(-\gamma)||x_i, x_j||^2), \gamma>0 \qquad \text{(Eq. 5)}$$

where, γ is the kernel parameter.

Moreover, with the Gaussian Radial-basis kernel function, the SVM model has two parameters (C, γ) which need to be determined. The cost parameter C controls the penalty for misclassifying a training point and consequently the complexity of the prediction function (Karatzoglou et al., 2006). A high cost value C, will result in a complex prediction function in order to misclassify as few training cases as possible. On the other hand, a low cost parameter C, results in simpler prediction functions. Thus, this type of SVM model is called C-SVM (Karatzoglou et al., 2006).

Karatzoglou et al. (2006), provide the primal form of the bound constraint C-SVM formulation:

minimize $t(w, \xi) = (\frac{1}{2})||w||^2 + (\frac{1}{2})\beta^2 + (\frac{C}{m}) \sum_{i=1}^{m} \xi_i$ (Eq. 6)

subject to $y_i(\langle \Phi(x_i), w \rangle + b) \geq 1 - \xi_i$

where, i= 1,……m

and $\xi_i \geq 0$, where i= 1,……m.

The dual form of the bound constraint C-SVM formulation (Karatzoglou et al., 2006) is:

maximize $W(\alpha) = \sum_{i=1}^{m} a_i - \frac{1}{2} \sum_{i,j=1}^{m} a_i a_j (y_i y_j + k(x_i, x_j))$ (Eq. 7)

subject to $0 \leq a_i \leq \frac{C}{m}$, where i= 1,......m

and $\sum_{i=1}^{m} a_i y_i = 0$.

3. Data preparation

The urban roads that were chosen were the Kifisias and Mesogeion avenues in Athens, Greece, mainly due to the fact that they had very similar characteristics. The required accident data were collected from the Greek accident database SANTRA, which is provided by the Department of Transportation Planning and Engineering of the National Technical University of Athens. A 6-year period was considered for the analyses of the present thesis, namely 2006-2011.

Traffic data were extracted from the Traffic Management Centre (TMC) of Athens, which operates on a daily basis from July 2004 covering various major arterials in the city of Athens. In order to apply the SVM models, a number of different datasets had to be prepared. Having known the time and location for each accident, the 3-hour time series of traffic flow, occupancy and speed (in 5-minutes intervals ending at the time of the accident) from the closest upstream as well as the closest downstream loop detector were utilized. For example, if an accident occurred in Kifisias Avenue on Wednesday 12 August 2009 at 13:00, then traffic data from Wednesday 12 August 2009 10:00 to 13:00 are extracted from the closest upstream and downstream loop detector measured in 5-min intervals. There were rare cases when loop detectors suffered from problems that might have resulted in unreasonable values for speed, volume, and occupancy. Such unrealistic values (e.g. occupancy>100%, speed>200 km/h or speed>0 along with flow=0) were discarded from the database. Accidents with traffic data unavailability were also discarded.

Thus, each dataset contains one of the following time series: traffic flow upstream, traffic flow downstream, speed upstream, speed downstream, occupancy upstream, occupancy downstream. Then, by following the Discrete Wavelet Transform (DWT) procedure described earlier all datasets are then transformed. Consequently, for each dataset containing the original time series data a duplicate dataset is created with the transformed time series data.

Then for each dataset, the SVM models were applied in order to predict:

a) PTW accident involvement and
b) PTW accident type.

In order to enhance the classification performance of SVMs, PTW accident type was classified as a binary outcome, namely single and multi-vehicle accidents, transforming the classification problem to a two-category classification. This approach was followed, because literature indicates that typical multi-classification problems have been very commonly observed when methods such as SVMs, Artificial Neural Networks (ANN) or classification trees are applied (Li et al., 2012; Delen et al., 2006). For example, Li et al. (2012), developed SVMs to model injury severity and the SVM model ignored severity categories with small proportions (namely fatal and incapacitating injuries) to improve the overall classification accuracy.

The final datasets consist of 527 accidents, where the PTWs were involved in 326 of them (61.9%). Regarding PTW accident type, PTWs were involved in 107 single-vehicle accidents (32.8%) and 219 multi-vehicle accidents (67.1%).

4. Results

The DWT procedure was conducted with the use of the package *wavelets* (Aldrich, 2010) in R software (R, 2015), and all SVM models were developed through the package *e1071* (Dimitriadou et al., 2011) in R software. It is noted that when building a SVM model, a training and a testing set have to be defined. For each dependent variable; PTW accident involvement and PTW accident type, respectively.

When building a SVM model, a training and a testing dataset have to be defined. The models are calibrated on the training set and then are used for prediction of the dependent variable on the testing set. Two different training and testing sets have to be prepared in order to further compare the results and to reduce the bias that arises when the accident database is randomly separated as Li et al. (2008) propose. In this paper, a 10-fold cross validation technique was applied on each dataset, in order to have a measure of the overall classification performance of the SVM models. Generally, in k-fold cross-validation, the original sample is randomly divided into k equal sized subsamples. Of the k subsamples, a single subsample is used as the validation dataset for testing the prediction performance of the model while the remaining k−1 subsamples are used as training data to calibrate the model. The cross-validation process is then repeated k times (the folds), with each of the k subsamples used exactly once as the validation data (Kohavi, 1995). Consequently, 10 subsamples were created. The total accuracy of SVM models was calculated with this simple equation:

$$\text{Accuracy} = \frac{number\ of\ correctly\ predicted\ cases}{total\ number\ of\ cases} \qquad \text{(Eq. 8)}$$

Moreover, different parameters of C and γ were tested each time in order to improve the performance of SVMs. The need for an investigation of SVMs performance by applying different values of C and γ was emphasized by Yu and Abdel-Aty (2013b), being another step further made by this paper.

4.1 PTW accident involvement prediction

In general, the SVM models of this paper showed relatively good classification accuracy compared to other similar studies (Li et al., 2012). However, it is noted that this is the first attempt to incorporate real-time traffic time series in SVMs. The development of the models showed that by modifying the two parameters, C and γ accordingly, the classification accuracy can be substantially influenced. A detailed presentation of SVM models' performance follows. Table 1 illustrates the total classification accuracy of SVM models regarding PTW accident involvement. In general, the original time series approach produced better results than the DWT approach. Overall, in regard with the parameters C and γ, the models performed better when the Cost parameter C had taken the value of 10 or 100, while the parameter γ values ranged from 0.001 to 0.01.

Total SVM performance	Speed downstream	Speed upstream	Flow downstream	Flow upstream	Occupancy downstream	Occupancy upstream
Original Time series						
10-fold cross validation mean accurracy %	62.17%	65.64%	61.90%	64.15%	60.72%	65.38%
(C, γ)	(100, 0.002)	(10, 0.01)	(100, 0.001)	(100, 0.001)	(100, 0.0001)	(100, 0.001)
DWT Time series						
10-fold cross validation mean accurracy %	60.65%	58.38%	60.82%	61.39%	59.20%	61.54%
(C, γ)	(10, 0.001)	(100, 0.01)	(100, 0.002)	(100, 0.0001)	(100, 0.01)	(100, 0.0001)

Table 1: Total classification accuracy of SVM models to predict PTW accident involvement.

When the original time series are considered, the best performance is consistently higher than 60%. The best accuracy is achieved when the speed upstream (65.64%) and occupancy upstream (65.38%) are considered. On the other hand, the time series of occupancy downstream of the accident location, were the worse predictors of PTW accident involvement (60.72%). When the DWT time series are considered, the total classification accuracy was generally around 60%. Speed upstream and occupancy downstream had the lowest accuracies (58.38% and 59.20% respectively).

The lower difference between original and DWT time series accuracy, was observed in speed downstream (1.52%) and occupancy downstream (1.52%). On the contrary, the highest difference was observed in speed upstream time series (7.26%). In all aforementioned situations, the DWT performed worse. The lower prediction performance of the transformed time series in predicting PTW accident involvement, may imply that there may be not necessary to extract features from time series but use the original time series instead. Alternatively, a different transformation could be applied as mentioned earlier.

4.2 PTW accident type prediction

Table 2 illustrates the total classification accuracy of SVM models regarding PTW accident type. As a first remark, neither of the two approaches (original or DWT time series) clearly outperformed the other, as the classification accuracies were similar. In regard with the parameters C and γ, the models performed better when the Cost parameter C had taken the value of 10 or 100, while the parameter γ values ranged from 0.000001 to 0.01.

Total SVM performance	Speed downstream	Speed upstream	Flow downstream	Flow upstream	Occupancy downstream	Occupancy upstream
Original Time series						
10-fold cross validation mean accurracy %	65.20%	66.76%	65.64%	66.57%	66.56%	66.67%
(C, γ)	(100, 0.01)	(100, 0.0001)	(10, 0.01)	(100, 0.0001)	(100, 0.0001)	(100, 0.0001)
Wavelet transformation						
10-fold cross validation mean accurracy %	63.63%	66.46%	66.26%	66.67%	65.31%	65.77%
(C, γ)	(100, 0.01)	(10, 0.001)	(100, 0.0001)	(10, 0.000001)	(10, 0.001)	(10, 0.001)

Table 2: Total classification accuracy of SVM models using to predict PTW accident type.

Overall, the models perform better in accident type prediction than in PTW accident involvement. When the original time series are considered, the best performance is consistently higher than about 65%. The best accuracy is achieved when the speed upstream (66.74%) is utilized. On the other hand, the utilization of time series of traffic flow downstream of the accident location, provided the lowest prediction of PTW accident involvement (65.64%). When the DWT time series are considered, the total classification accuracy was generally around 63-66%. Speed downstream had the lowest accuracy (63.63%), whilst flow upstream had the best accuracy (66.67%).

Moreover, when predicting PTW accident type, it is observed that both approaches provided similar results. Original time series performed better that DWT, when speed downstream and occupancy (both in upstream and downstream) are utilized. On the other hand, DWT performed slightly better for flow downstream. For speed upstream and for flow upstream, the SVM models showed similar results for original and DWT time series. Therefore, it may be suggested that both approaches are appropriately adequate for predicting PTW accident type.

The next two figures (Figures 4 and 5), demonstrate the relative performance of the SVM models when predicting PTW accident involvement and PTW accident type, by utilizing original and DWT time series.

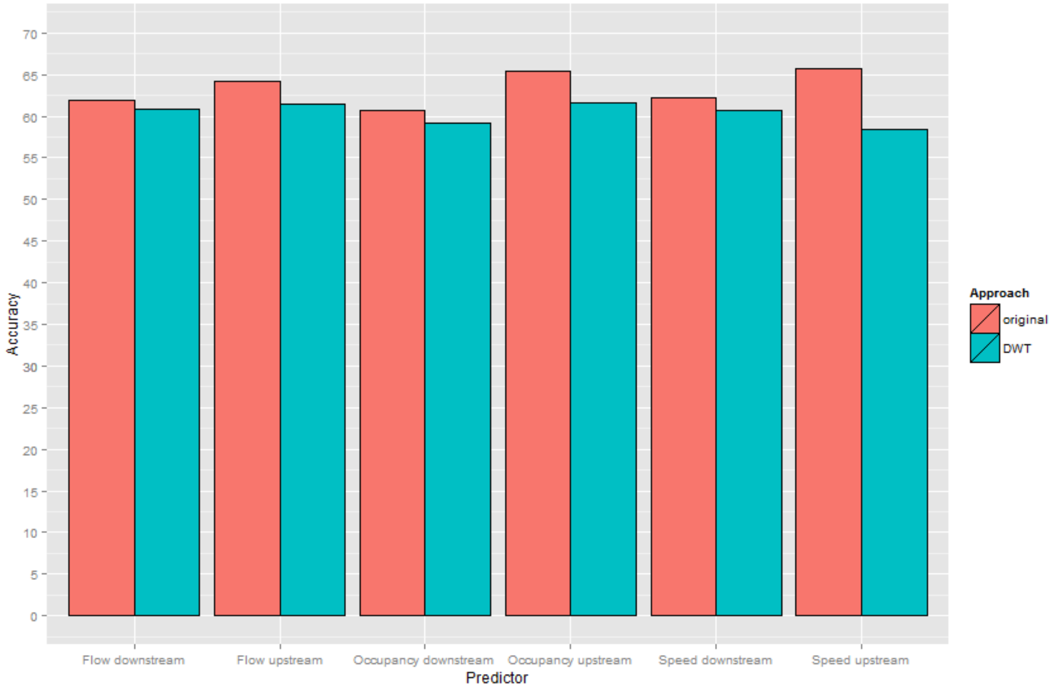

Figure 4: Graphical illustration of SVM models performance when analysing PTW accident involvement.

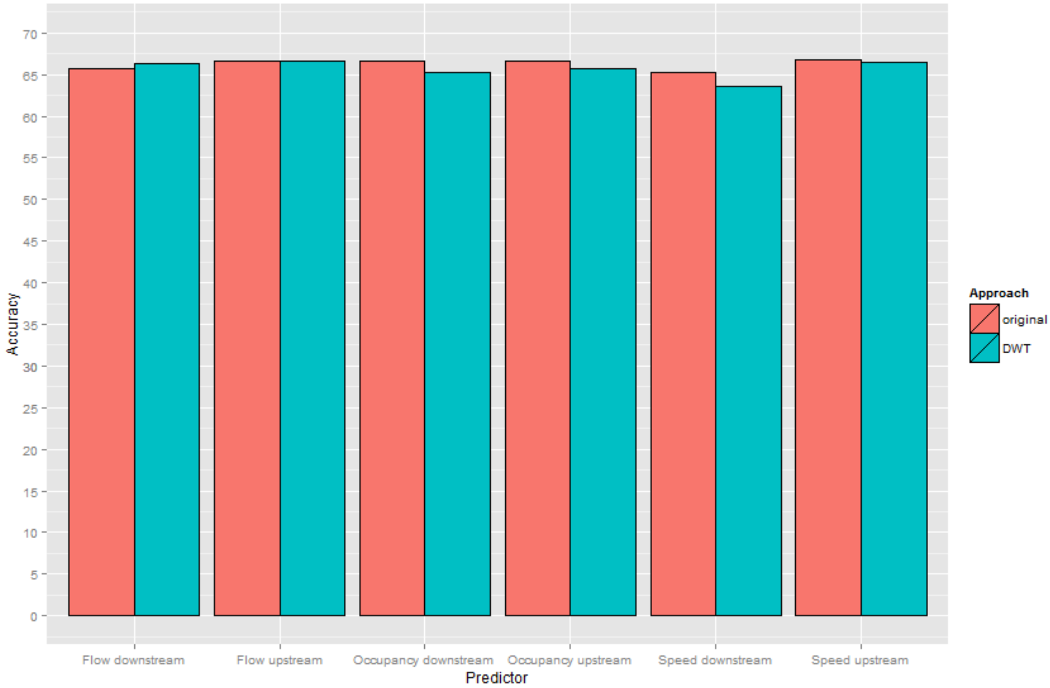

Figure 5: Graphical illustration of SVM models performance when analysing PTW accident type.

5. Conclusions

This paper has presented the prediction results from the Support Vector Machine (SVM) models, which were applied on the original time series and on the Discrete Wavelet Transformed (DWT) time series of flow, speed and occupancy upstream and downstream of the accident location. This methodological approach, was a first attempt to incorporate time-series data when analysing road safety with real-time traffic data. Moreover, the opportunity of applying of a relatively new and scientifically strong classification technique such as Support Vector Machines, in road safety with real-time traffic data was only recently been explored.

The prediction of whether a PTW is involved in an accident, given that the accident has occurred, is two-category classification problem. Overall, the original time series data performed better than the DWT time series. However, in a few cases, the difference between original time series and DWT time series was very low. Consequently, original time series are preferred for predicting PTW accident involvement.

This combined approach proposed by this paper, is considered promising when investigating the accident type of Powered-Two-Wheelers (PTW) that occur in urban roads. The dependent variable (single and multi-vehicle accident) is considered as a two-category classification problem as well. In this case, both original time series and DWT time series performed well. Neither of the two approaches clearly outperformed the other. Therefore, both type of time series could be utilized and produce good prediction.

In general, it is observed that this combined approach provides better results when accident type is aimed to be predicted. Moreover, when original time series are utilized, upstream speed had consistently the best accuracy both for PTW accident involvement and PTW accident type.

Summing up, this methodological approach showed promising results, having produced a number of adequately high correct classification percentages in some cases. The main conclusion of this paper is that the combination of SVM models and time-series data can be used for road safety purposes especially by utilizing high resolution traffic data. Clearly, this direction has to be exploited further. It is interesting though, that despite the fact that Zhao (2012) suggests to extract features from time-series when perform time series data-mining, the performance of SVMs on the DWT data did not generally outperformed original time series. This means that the transformation applied in the paper is not always necessary, but alternatively, a different transformation could be applied, such as Singular Value Decomposition (SVD), Discrete Fourier Transform, Piecewise Aggregate Approximation, Perpetually Important Points and so on.

6. References

Abdel-Aty, M., Pande, A. (2005). Identifying crash propensity using specific speed conditions. Journal of Safety Research, 36, 97-108.

Abdel-Aty, M., Pande, A., Lee, C., Gayah, V., Dos Santos, C. (2007). Crash risk assessment using intelligent transportation systems data and real-time intervention strategies to improve safety on freeways. Journal of Intelligent Transportation Systems: Technology, Planning, and Operations, 11(3), 107-120.

Abdul Manan, M.M., Várhelyi, A. (2012). Motorcycle fatalities in Malaysia. IATSS Research, 36, 30-39.

Ahmed, M., Abdel-Aty, M. (2012). The viability of using Automatic Vehicle Identification data for real-time crash prediction. IEEE Transactions of intelligent transportation systems, 13(2), 459-468.

Aldrich, E. (2010) Wavelets: A package of functions for computing wavelet filters, wavelet transforms and multiresolution analyses. Retrieved from: http://cran.rproject.org/web/packages/wavelets/index.html.

Barmpounakis, E., Vlahogianni, E., Golias, J. (forthcoming). Intelligent Transportation Systems and Powered Two Wheelers Traffic. IEEE Transactions on Intelligent Transportation Systems, in press.

Burrus, C.S., Gopinath, R.A., Guo, H. (1998). Introduction to wavelets and wavelet transforms: A Primer. Prentice-Hall, Inc.

Ceder, A., Livneh, M. (1982). Relationship between road accidents and hourly traffic flow–I. Accident Analysis and Prevention, 14, 19-34.

Christoforou, Z., Cohen, S., Karlaftis, M. (2011). Identifying crash type propensity using real-time traffic data on freeways. Journal of Safety Research, 42, 43-50.

Cortes, C., Vapnik, V. (1995). Support-vector network. Machine Learning, 20, 1-25.

Delen, D., Sharda, R., Bessonov, M. (2006). Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. Accident Analysis and Prevention, 38(3), 434-444.

Dimitriadou, E., Hornik, K., Leisch, F., Meyer, D., Weingessel, A. (2011). e1071: Misc functions of the department of statistics (E1071), TU Wien.

European Road Safety Observatory, ERSO. (2015). Traffic safety basic facts "Motorcycles and Mopeds". European Commission, Traffic Safety Basic Facts on Motorways, European Commission, Directorate General for Transport.

Golob, T., Recker, W., Pavlis, Y. (2008). Probabilistic models of freeway safety performance using traffic flow data as predictors. Safety Science, 46, 1306-1333.

Haar, A. (1910). Zur zheorie der orthogonalen funktionensysteme. Mathematische Annalen, 69(3), 331-371.

Haque, M.M., Chin, H.C., Huang, H. (2009). Modelling fault among motorcyclists involved in crashes. Accident Analysis and Prevention, 41, 327-335.

Harnen, S., Wong, S.V., Radin Umar, R.S., & Wan Hashim, W.I. (2003). Motorcycle crash prediction model for non-signalized intersections. IATSS Research, 27(2), 58-65.

Karatzoglou, A., Meyer, D., Hornik, K. (2006). Support Vector Machines in R. Journal of Statistical Software, 15(9), 1-28.

Karatzoglou, A., Smola, A., Hornik, K., Zelies, A. (2005). Kernlab-Kernel Methods. R package, Version 0.6-2. Retrieved from: http://CRAN.R-project.org/.

Keckman, V. (2005). Support vector machines-an introduction. In: Wang, L. (eds). Support Vector Machines: Theory and Applications, 1-48. Springer-Verlag, Berlin, Heidelberg, New York.

Kim, D., Washington, S., Oh, J. (2006). Modeling crash types: new insights into the effects of covariates on crashes at rural intersections. Journal of Transportation Engineering, 132, 282-292.

Kleist, C. (2015). Time series data mining methods: A review. Master Paper, Humboldt-Universitat zu Berlin, School of Business and Economics.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the 14th international joint conference on Artificial intelligence, 2, 1137-1143.

Li, X., Lord, D., Z., Y., Xie, Y. (2008). Predicting motor vehicle crashes using Support Vector Machine models. Accident Analysis and Prevention, 40, 1611-1618.

Li, Z., Liu, P., Wang, W., Xu, C. (2012). Using support vector machine models for crash injury severity analysis. Accident Analysis and Prevention, 45, 478-486.

Lin, M.R., Kraus, J.F. (2009). A review of factors and patterns of motorcycle injuries. Accident Analysis and Prevention, 41, 710-722.

McLeod, A. I., Yu, H., Mahdi, E. (2012). Time series analysis with R. In Time Series Analysis: Methods and Applications, Chapter 23 (pp. 661-712) in Handbook in Statistics, Elsevier, Volume 30, Edited by T. S. Rao, S. S. Rao and C. R. Rao.

Meyer, D. (2001). Support vector machines. R News, 1(3), 23-26.

Pande, A., Abdel-Aty, M. (2006). Assessment of freeway traffic parameters leading to lane-change related collisions. Accident Analysis and Prevention, 38, 936-948.

R Development Core Team. (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from: http://www.R-project.org.

Schneider IV, W.H, Savolainen, P., Van Boxel, D., & Beverley, R. (2012). Examination of factors determining fault in two-vehicle motorcycle crashes. Accident Analysis and Prevention, 45, 669-676.

Sharma, A.K., Landge, V.S., Deshpande, N.V. (2013). Modeling Motorcycle Accident on Rural Highway. International Journal of Chemical, Environmental and Biological Sciences, 1(2), 313-317.

Struzik, Z.R., Siebes, A. (1999). The haar wavelet transform in the time series similarity paradigm. In Principles of Data Mining and Knowledge Discovery, pages 12-22. Springer.

Theofilatos, A., Yannis, G. (2015). A review of powered-two-wheeler behaviour and safety. International Journal of Injury Control and Safety Promotion, 22(4), 284-307.

Vapnik, V. (1998). Statistical learning theory. Wiley, New York, U.S.

Vidaković B., Müller, P. (1991). Wavelets for Kids – A Tutorial Introduction, Duke University, 1991, unpublished. Retrieved from: http://www.isye.gatech.edu/~brani/wp/kidsA.pdf.

Vlahogianni, E., Yannis G., Golias, J. (2012). Overview of critical risk factors in Power-Two-Wheeler safety. Accident Analysis and Prevention, 49, 12-22.

Wanvik, P.O. (2009). Effects of road lighting: An analysis based on Dutch accident statistics 1987-2006. Accident Analysis and Prevention, 41, 123-128.

World Health Organization, WHO. (2015). Road Traffic Injuries. Retrieved from: http://www.who.int/mediacentre/factsheets/fs358/en/.

Yannis, G., Vlahogianni, E., Golias, J., Saleh, P. (2010). Road infrastructure and safety of powered-two wheelers. In Proceedings of the 12th World Conference on Transport Research, July 11-15 2010, Lisbon.

Zhou, M., Sisiopiku, V. (1997). Relationship between volume-to-capacity ratios and accident rates. Transportation Research Record, 1581, 47-52.

Yu, R., Abdel-Aty, M., (2013a). Investigating the different characteristics of weekday and weekend crashes. Journal of Safety Research 46, 91-97.

Yu, R., Abdel-Aty, M., (2013b). Utilizing support vector machine in real-time crash risk evaluation. Accident Analysis and Prevention, 51, 252-259.

Yu, R., Abdel-Aty, M., (2014). Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. Safety Science, 63, 50-56.