ANALYSIS OF TRAFFIC VELOCITY UNDER DIFFERENT WEATHER AND TEMPORARY CONDITIONS

Héctor Cogollos Adrián
PhD Student, University of Burgos, Spain
Santiago Porras Alfonso
Lecturer, University of Burgos, Spain
Bruno Baruque Zanón
Senior Lecturer, University of Burgos, Spain

ABSTRACT

The density of traffic within urban areas depends on multiple factors, and among those; one that has major impact is the weather. This paper presents a study that empirically analyzes the traffic flow velocity depending on the atmospherically conditions and the day schedule. The objective is to assess systematically to which extent there is a correlation between the vehicle velocity on an urban environment and the weather, depending on the time of day. The presented case of study uses a set of real data, specifically the trips made by taxis in the city of Porto (Portugal). First, the vehicle GPS routes are analyzed to identify departure and arrival points and estimate the route average velocity and adding weather and time conditions. The data is processed with different regressive techniques to obtain the influence of the variables on the velocity. The results show differences between days of the day and weekends, as well as differences in speeds with a favorable climate for driving compared to other more adverse ones.

1. INTRODUCTION

Traffic within cities can be considered as a living being that grows and evolves over time. Cities grow, and it is necessary to designs a good urban planning and road traffic management according to this evolution. Likewise, in recent years with the advancement of technologies and the appearance of concepts such as smart cities, the different municipal corporations invest large amounts of money in putting technologies at the service of their inhabitants to improve their quality of life.

Urban traffic management is a challenging problem nowadays; it is difficult to find a management system or a generic solution to this problem, as each urban area presents a different problem.

This is due to factors as diverse as the available infrastructures, the demography, geographical situation and orography of the city, issues such as the cultural customs of the inhabitants, the economic power or the vehicle fleet, among others.

A determining factor that affects the variation in traffic flow is the weather conditions (Cools, Moons et al. 2010). In the present work, a study is performed to see how different weather conditions affect the way in which users of urban roads move in the city of Porto (Portugal). With this, we want to observe how this behaviour varies at different times and days of the week and depending on the atmospheric climate and analyse if there are significant differences between some conditions and others. With this, it will be possible to make a forecast of the traffic state, which will facilitate its management, being able to avoid traffic jams and accidents (Maze, Agarwal et al. 2006).

The rest of the work is divided in, related work, dataset explanation methodology employed, experiment results, conclusions and future work.

2. RELATED WORK

Traffic flows within urban areas have been widely studied in the literature and encompass different problems that all affect the traffic of vehicles in cities. Examples of them can be found in traffic light management (Gupta, Kumar et al. 2017), accident prevention (Aldegheishem, Yasmeen et al. 2018), contaminant emissions reduction (Anjum, Noor et al. 2019), or traffic flow predictions to avoid traffic congestion (Hu, Wang et al. 2018).

In our specific case, we want to analyse the state and variation of traffic flows velocity depending on the atmospheric conditions due to it is one of the factors that most influence the behaviour of drivers.(Kilpeläinen and Summala 2007).

In this field one of the pioneering works in this field was (Tanner 1952) in which they analyse the correlation between the use of different vehicles under the effect of rain, distinguishing between working days and holidays. As technology advances, more scientific articles related to traffic and weather have appeared.

Golob and Recker (2003) demonstrate with multivariate statistical methods such as weather, traffic flow and lighting influence on accidents on the highways of California.

The effect of the snow or precipitations in the traffic flow congestion using video recordings is studied in (Asamer and Van Zuylen 2011). Tsapakis, Cheng et al. (2013), show the variation in the speed of London drivers under various conditions of snow, rain, and temperature.

Paying attention to snow and how it affects traffic we can find the case of Beijing (Weng, Liu et al. 2013) . Authors highlight the reduction in speed and capacity of urban roads, as well as the increase in travel times; in addition, these effects are noticeable in the following days due to the ice although no longer snow.

In (Stamos, Salanova Grau et al. 2016) the effect of heavy rains and how it affects the speed of vehicles in Thessaloniki (Greece) is analyze from data of 1,200-taxi fleet that contained GPS coordinates and speed information. under certain atmospheric conditions (Yasanthi and Mehran 2020), in this case traffic data is crossed with atmospheric data, including data such as temperature and the condition of the pavement or wind speed.

Finally, congestion is studied in Porto, (Portugal) with diagrams and probabilistic models (Silva, d'Orey et al. 2018), for this they use data obtained from taxis and the weather, but only focusing on a small central area.

3. DATA SET DESCRIPTION

The data set used contains GPS routes of taxis in the city of Porto (Kaggle 2018). The TRAVISANA software has been used to process it (Cogollos, Porras et al. 2020). In it the routes are pre-processed, filtered, and speeds are calculated. In Figure 1 the application pipeline process is shown.

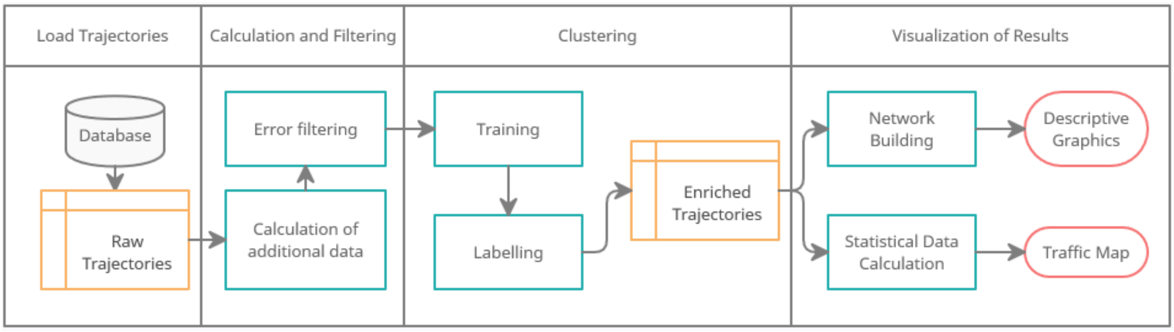

Figure 1: TRAVISANA Pipeline.

In this specific case, the atmospheric data, extracted from(Ltd 2020), are also added to the routes. All the data is combined in the application to obtain the complete data set with the speeds of each vehicle in each trip and the atmospheric conditions of that moment.

Specifically, the set consists of the following variables:

- Id: Route identifier.
- Velocity: Average speed of the vehicle during the trip in kilometers per hour.
- Day: Day of the week from 1 Monday to 7 Sunday.
- Temperature: Temperature at that moment in Celsius degrees.
- Precipitation: Amount of rain in milliliters.
- Visibility: Range of visible space in kilometers.
- Humidity: Grams of water for each cubic meter of air.

Finally, Weather: categorical variable, with 20 values: clear, cloudy, fog, heavy rain, heavy rain at times, light drizzle, light rain, light rain shower, mist, moderate or heavy rain shower, moderate or heavy rain with thunder, moderate rain, moderate rain at times, overcast, partly cloudy, patchy light drizzle, patchy light rain, patchy rain possible, sunny and torrential rain shower.

4. METHODOLOGY

To see the influence of each variable on t the vehicles velocity, the following regression techniques are used.

- Multiple regression: This method explains the variable under study as far as possible, using a group of variables that, directly and indirectly, participate in its value.
- Principal component regression (PCR): This method differs from the classical multiple linear regression in that first a principal component analysis is performed on the dependent variables and later the multiple regression is applied using the obtained factors.
- Partial least squares regression (PLS Regression): This technique combines the multiple regression ant the principal component analysis but is focused on reduce the multicolinearity and obtain more accurate predictions, as difference PCR is more focused on reduce the number of dependent variables.
- Regularized regression: Also known as regulated or constrained regression, specifically the elastic net regression has been used in this paper. This family of regression methods adds a penalty term to the best fit trying to obtain less variance achieve and restricting the influence the dependent variables over the independent one by compressing their coefficients.
- Multiple adaptative regression splines (MARS): It is an algorithm that automatically creates a piecewise linear model which provides an intuitive stepping block into nonlinearity. In order to identify the nonlinear relationships MARS assesses each data point for each predictor as a knot and creates a linear regression model with the candidate features.
- Bagged MARS: Bagging it is a machine learning technique that raises the stability of models by reducing the variance and improving the accuracy, avoiding overfitting.
- Bagged MARS with gCV Pruning: In this version, a generalized cross-validation is added to a Bagging MARS. In addition, a pruning step is executed in every iteration, removing the term in the model that gives the smallest increase in the sum of squared error.

5. EXPERIMENTS RESULTS AND DISCUSSION

To apply the methods to the data set, the caret package from R has been used(Kuhn 2008). All the techniques have been applied with a general cross validation process, with *k=10*. The root mean square error (RMSE) after the general cross validation process (GVC) and the root square sum of the errors (RSS) are used as performing metrics.

The aim of the study it is investigate the relationship between the velocity and the weather conditions in different days of the week.

Initially, all the data have been processes for a first approximation using MARS. Results can be observed, in Table 1 and Figure 2. According with them we can distinguish that the factor with more influence in the velocity are the sunny weather, humidity, and the weekend’s days. Therefore, we can conclude that the daily days and weekends has an influence on vehicle velocity in weekends the velocity is higher than in daily days. For a further exploration and more accuracy study of weather influence, the data has been split in daily days and weekends.

Variable	GVC	RSS
WeatherSunny	100.0	100.0
Humidity	100.0	100.0
Day7 (Sunday)	68.0	68.2
Day6 (Saturday)	55.1	55.3
Temperature	48.0	48.3
Precipitation	37.9	38.2
WeatherCloudy	37.4	37.7
WeatherPartly cloudy	32.6	32.9
WeatherOvercast	27.2	27.4

Table 1: GVC and RSS Values for all data.

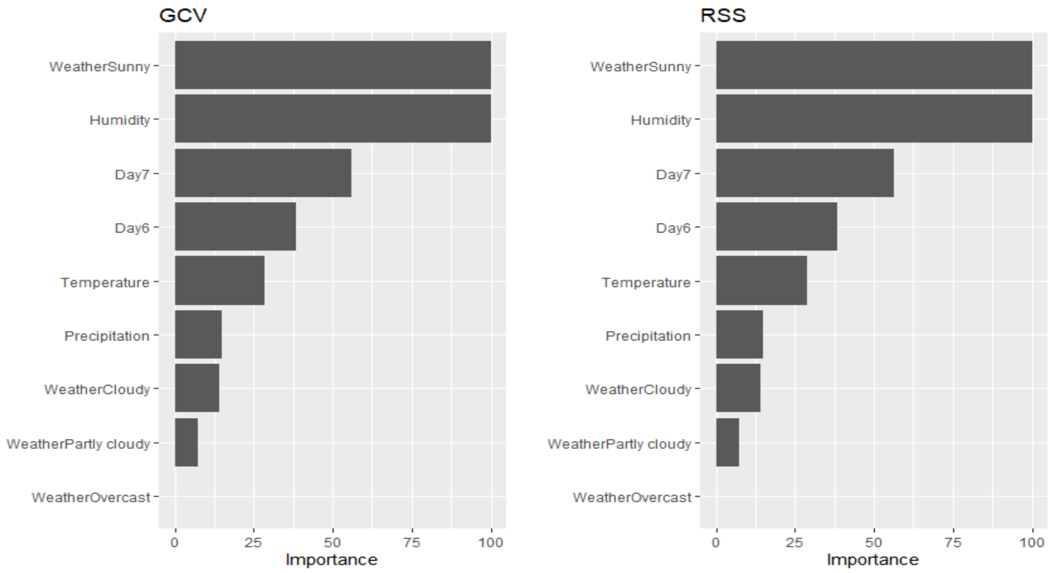

Figure 2: Variable importance according to the contribution to GCV and RSS values as predictors for all data.

In Table 2 the RMSE of predicted velocity for all methods are presented, according with them, the methods who best fit are MARS for daily days and PLS regression for weekends.

	RMSE	
Method	**Daily**	**Weekend**
Multiple regression	12.64511	12.77780
PCR regression	12.64583	12.94490
PLS regression	12.64511	**12.77776**
Regularized regression	12.64522	12.77780
Bagged MARS	12.67633	12.92724
Bagged MARS with gCV Pruning	12.67451	12.90767
MARS	**12.55634**	12.79040

Table 2: RMSE values for Daily days and weekends.

Daily days study: Figure 2 and Table 3 shows the variable importance. According with them it can be observed that sunny weather and humidity have a great influence over velocity as expected according to the initial analysis, but it is observed that between days the temperature and that the weather is cloudy contributes to a higher speed. Analysing the combinations of the variables in Table 4, it is noticeable that the speed increases with temperatures above 9 degrees, and humidity levels below 57 either with sunny or cloudy weather. On the other hand, when precipitation appears, speeds decrease.

Variable	GCV	RSS
WeatherSunny	100.0	100.0
Humidity	100.0	100.0
Temperature	54.7	55.1
WeatherCloudy	43.7	44.2
Precipitation	37.4	37.9

Table 3: GVC and RSS Values for daily data.

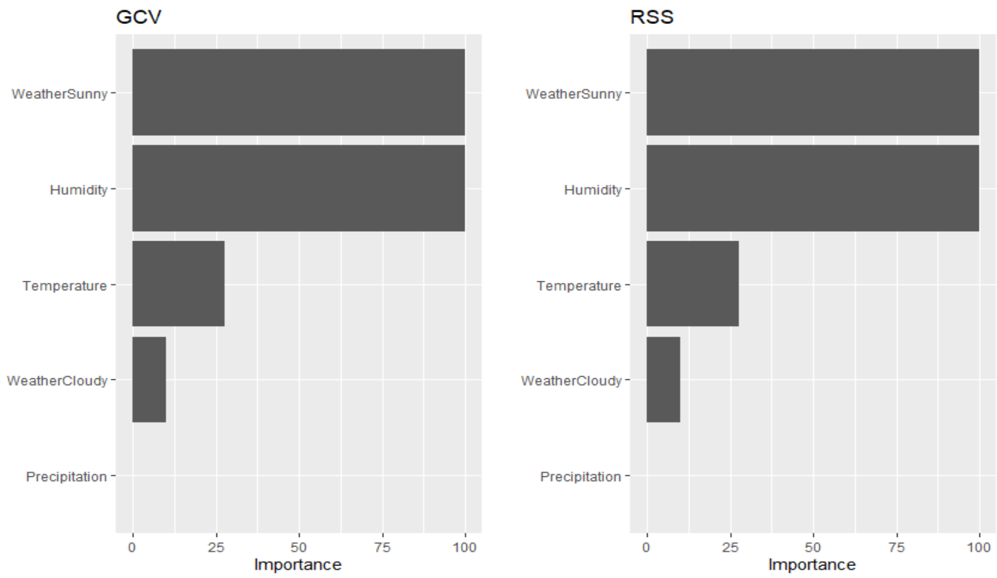

Figure 3: Variable importance according to the contribution to GCV and RSS values as predictors for daily days.

Variable	Coefficent
(Intercept)	22.95928
WeatherSunny	0.82939
WeatherSunny*h(Humidity-57)	0.55996
h(Temperature-9)	0.17940
h(9-Temperature)	12.73014
WeatherCloudy*h(Temperature-9)	1.42678
h(9-Temperature)*h(87-Humidity)	0.11849
h(9-Temperature)*h(Precipitation-0.8)	-5.37981
h(9-Temperature)*h(0.8-Precipitation)	-15.80329
WeatherCloudy*h(Temperature-9)*h(Humidity-97)	-1.34813

Table 4: Mars coefficients for Daily Days.

Weekend study: We used PLS as it is the method with less RMSE. In Table 5 can be seen as more atmospheric states influence the speed, as in the previous case, the most influential factors are sunny weather and humidity. However, visibility and rainfall are important; this can be related to rainy climates. Our PLS result only gives us one level and we do not see the iterations between the variables, but by analyzing the coefficients, we can say that the velocity is higher with sunny and cloudy weather and decreases with light rain or fog. The positive weight of rainfall suggests that speed increases with rainfall and the negative of visibility that decreases the more visibility there is. These factors seem contradictory and would require a more complete study.

Variable	GCV	RSS	Coefficent
WeatherSunny	100	100	3.05085
Humidity	27,9	27,9	0.19905
WeatherLight rain	25,8	25,8	-0.17142
Visibility	24,9	24,9	-0.06000
Precipitation	21,5	21,5	0.36494
WeatherOvercast	21,1	21,1	0.91770
Temperature	18,4	18,4	-0.72814
WeatherPartly cloudy	17,7	17,7	1.19975
WeatherModerate rain	15,4	15,4	0.63879
WeatherModerate or heavy rain shower	11,7	11,7	0.06692
WeatherTorrential rain shower	11,6	11,6	-0.22982
WeatherLight rain shower	5,1	5,1	0.728437
WeatherLight drizzle	2,6	2,6	-0.08700
WeatherMist	1,3	1,3	-0.12495
WeatherCloudy	1,29	1,29	0.43920
WeatherPatchy rain possible	0	0	0.67629

Table 5: PLS regression coefficients for Daily Days.

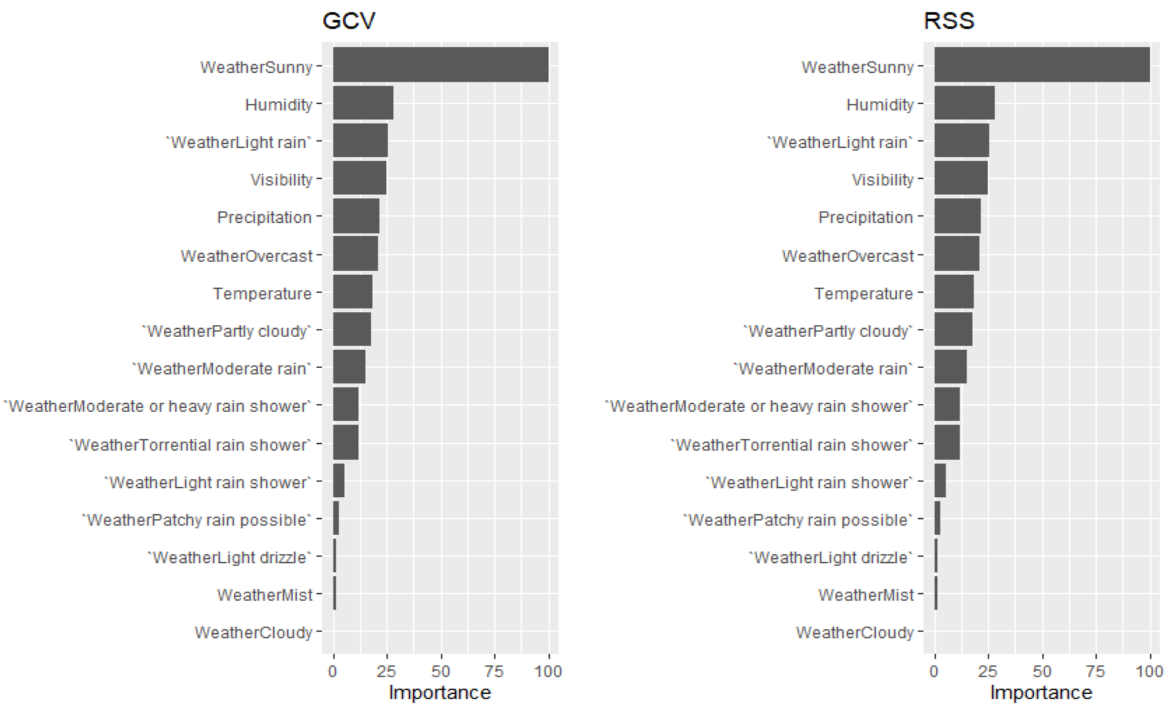

Figure 4: Variable importance according to the contribution to GCV and RSS values as predictors for weekends.

6. CONCLUSIONS AND FUTURE WORK

As seen in the results section, atmospheric conditions and the day of the week significantly influence the average speed at which Porto taxis circulation. It can be concluded that under sunny or cloudy days without rainfall, taxis travel at a higher speed, as conditions worsen the speed is reduced. For daily analysis, the speed is higher on weekends than on daily days, this could be due to a lower volume of circulation, regardless of the weather. We must highlight humidity importance; we assume that it will be due to the geographical situation of Porto.

This study has its limitations, evidently factors such as the volume of vehicles, the type of road, the speed limit, the lighting or the time of day also affect the flow of traffic. Collecting these data and obtaining data from different types of vehicles to build a better model is one future line of work in order achieve more precise results.

In addition, it is also limited by the characteristics and climate of Porto, as future work we plan to compare it with other cities with different climates and orography.

REFERENCES

ALDEGHEISHEM, A., H. YASMEEN, H. MARYAM, M. A. SHAH, A. MEHMOOD, N. ALRAJEH AND H. SONG (2018). "Smart road traffic accidents reduction strategy based on intelligent transportation systems (tars)." Sensors 18(7): 1983.

ANJUM, S. S., R. M. NOOR, N. AGHAMOHAMMADI, I. AHMEDY, L. M. KIAH, N. HUSSIN, M. H. ANISI AND M. A. QURESHI (2019). "Modeling traffic congestion based on air quality for greener environment: an empirical study." IEEE Access 7: 57100-57119.

ASAMER, J. AND H. J. VAN ZUYLEN (2011). "Saturation flow under adverse weather conditions." Transportation research record 2258(1): 103-109.

COGOLLOS, H., S. PORRAS AND B. BARUQUE (2020). TRAVISANA (TRAFFIC VISUAL ANALYSER): Traffic analysis application using geopositioning routes. Provincial office of the general registry of intellectual property in Burgos. B. University. Spain. BU-121-20.

COOLS, M., E. MOONS AND G. WETS (2010). "Assessing the impact of weather on traffic intensity." Weather, Climate, and Society 2(1): 60-68.

DATLA, S., P. SAHU, H.-J. ROH AND S. SHARMA (2013). "A comprehensive analysis of the association of highway traffic with winter weather conditions." Procedia-social and behavioral sciences 104: 497-506.

GOLOB, T. F. AND W. W. RECKER (2003). "Relationships among urban freeway accidents, traffic flow, weather, and lighting conditions." Journal of transportation engineering 129(4): 342-353.

GUPTA, V., R. KUMAR, K. S. REDDY AND B. PANIGRAHI (2017). Intelligent traffic light control for congestion management for smart city development. 2017 IEEE region 10 symposium (TENSYMP), IEEE.

HU, W., H. WANG, Z. QIU, L. YAN, C. NIE AND B. DU (2018). "An urban traffic simulation model for traffic congestion predicting and avoiding." Neural Computing and Applications 30(6): 1769-1781.

KAGGLE. (2018, 2018). "Taxi Trajectory Data." 2021, from https://www.kaggle.com/crailtap/taxi-trajectory.

KILPELÄINEN, M. AND H. SUMMALA (2007). "Effects of weather and weather forecasts on driver behaviour." Transportation research part F: traffic psychology and behaviour 10(4): 288-299.

KUHN, M. (2008). "Building predictive models in R using the caret package."

LTD, Z. (2020). "World Weather Online." from https://www.worldweatheronline.com.

MAZE, T. H., M. AGARWAL AND G. BURCHETT (2006). "Whether Weather Matters to Traffic Demand, Traffic Safety, and Traffic Operations and Flow." Transportation Research Record 1948(1): 170-176.

SILVA, C., P. M. D'OREY AND A. AGUIAR (2018). Interpreting traffic congestion using fundamental diagrams and probabilistic graphical modeling. 2018 IEEE International Conference on Data Mining Workshops (ICDMW), IEEE.

STAMOS, I., J. M. SALANOVA GRAU, E. MITSAKIS AND G. AIFADOPOULOU (2016). "Modeling effects of precipitation on vehicle speed: floating car data approach." Transportation Research Record 2551(1): 100-110.

TANNER, J. (1952). "Effect of weather on traffic flow." Nature 169(4290): 107-107.

TSAPAKIS, I., T. CHENG AND A. BOLBOL (2013). "Impact of weather conditions on macroscopic urban travel times." Journal of Transport Geography 28: 204-211.

WENG, J., L. LIU AND J. RONG (2013). "Impacts of snowy weather conditions on expressway traffic flow characteristics." Discrete dynamics in nature and society 2013.

YASANTHI, R. G. AND B. MEHRAN (2020). "Modeling free-flow speed variations under adverse road-weather conditions: case of cold region highways." Case studies on transport policy 8(1): 22-30.