Traffic flow prediction model and probability of congestion

To cite this article: M Pilamunga Poveda *et al* 2020 *IOP Conf. Ser.: Mater. Sci. Eng.* **844** 012008

IOP Conf. Series: Materials Science and Engineering **844** (2020) 012008 doi:10.1088/1757-899X/844/1/012008

Traffic flow prediction model and probability of congestion

M Pilamunga Poveda[1*], J M Lavin[2], M Aldás Sánchez[3], D Ríos Insua[4]

[1] Design and Architecture School, Universidad Técnica de Ambato, Av. Los Chasquis and Rio Payamino, 180207, Ambato, Ecuador
[2] Cesine Centro Universitario, Calle Jose Simón Cabarga 6, 39005, Santander, España.
[3] Civil Engineering School, Universidad Técnica de Ambato, Av. Los Chasquis and Rio Payamino, 180207, Ambato, Ecuador
[4] CSIC, Serrano 117, 28040 Madrid, España

[*] Email: em.pilamunga@uta.edu.ec

Abstract. Investment in the management of integrated activities for better community development determines the progress of a city. Generate actions directed to change and evolution can promote greater satisfaction among its inhabitants. One of those activities is the one concerning transit, which is one of the most transcendental factors in the quality of life of a modern city. Thus, good road planning is of vital importance for their development. To do this, studies must be conducted that allow us to have a diagnosis of reality and thus make decisions would help to minimize or avoid problems that are occurring or that may appear in the future. However, to make a study of this type whenever it is required is expensive in terms of time and resources. This work proposes a model that allows predictions of traffic flows in different relevant sections of a city and, predict issues such as traffic saturation in them, thus determining possible solutions to road congestion, it can be presented at a given moment. These possible alternative solutions could consequently would reduce costs in time, fuel, pollution among others.

1. Introduction

Urban traffic is one of the problems that most influence the quality of life of residents in cities. This traffic has been accentuated in the last decades, due to the development of three simultaneous phenomena:

- The increase in the use of private cars to the detriment of public transport, due to the fact that it is insufficient to cover all sectors of the city;
- The generalization of the diffuse city model [1], which has reached its current development after a decentralization process; and,
- Economic growth.

Ambato is characterized for being a node of great commercial activity in the national context, existing an intense commerce. For this reason, in recent decades, the city has experienced a considerable increase in population and buildings, added to the geography of the city with many slopes and the growing presence of cars, the concentration of public and private entities in certain areas of the city, the good commercial activity of the marketplaces and industries, the great concentration around the educational centers, among others, complicate the traffic in several sectors of the city. For these reasons, it is very important to generate tools that help manage traffic and allow to determine a better road planning in a

IOP Conf. Series: Materials Science and Engineering **844** (2020) 012008 doi:10.1088/1757-899X/844/1/012008

technical way, reducing costs depending on distances and waste of time, fuel expense and environmental degradation.

2. Related works

First, it is necessary to expose some definitions and basic principles of networks. A road network is a set of streets and avenues that intersect at intersections, which may occur at a level or slope. In general, the stretches of streets or avenues have a certain sense of circulation and, therefore, are designated as directed. Therefore, the basic model will use graphos. A grapho is a set, not empty, of objects called vertices (or nodes) and a selection of pairs of vertices, called edges, that can be oriented or no [2].
Mobility studies are based on the analysis and optimization of a traffic network through characterizing the infrastructure and its capacity include detailed field investigations necessary to know the origin / destination relationships and their main characteristics. This allows to adjust the functional characteristics of the different elements of the system, including a detailed analysis of the intersections and their different methods of regulation, Fernandez y Valenzuela [3].
Below are some existing vehicle traffic models that help to better understand the advantages of the proposed model. Each user defines how to reach a destination, choosing the one that considers the best route, considering factors such as cost, time, security and comfort. The estimated flows can serve both to describe traffic and to predict or recommend a pattern of vehicular flow in a network. For this, for example, can be used the models of Castillo et al. [4] or Tebaldi and West [5], which represent the macroscopic behavior of traffic in large urban areas or entire cities. Through these models it is possible to estimate the vehicular flows in a network where there is a certain demand for trips and the effects of congestion make the travel times in the arcs depend on the flows.
The urban and environmental impacts associated with traffic go, to a large extent, directly associated with the degree of saturation ($\boldsymbol{x}$) of any road device (road, intersection, whereabouts, network). The degree of saturation is defined as the ratio between the flow used by the device (q) and its capacity (Q), that is, $x = q/Q$. Congestion is a consequence of the degree of saturation and usually manifests in:

- Additional delays to people beyond those required to make a trip at a safe and constant speed.
- Queues of vehicles or people, that are generated in sections of track, crossings or terminals of public or private transport like bus stops or parking lots.
- Involuntary stops of vehicles or users, which occur in the same places as mentioned before, due to the advance and stagnation of a queue (start-stop condition).

Different studies such as Akcelik [6], IHT [7], indicate that the manifestations of congestion begin to occur for values of ($\boldsymbol{x}$) greater than 0,7 ó 0,8. During the last decades, traffic flow prediction models with different degrees of complexity have been proposed. The simplest are the combination of random paths, but only work well in specific situations [8].
A more complex approach [9], is based on time series models ARIMA, seasonal models ARIMA [8] and [10], Kalman filters [11], simulation [12] and [13], local regression models [14], combination of Kohonen maps and ARIMA models [15], and and Markov chains see Yu et al. [16]. Here dynamic linear models are used as a variation. Introductions to these models can be seen in West and Harrison [17] and Petris et al. [18], which include numerous applications in the economy and the business world. In comparison with the aforementioned proposal, the sample approaches that of Okutani and Stephanedes [11], but it has the advantage of allowing a priori information to be incorporated. Compared with the rest, it has the advantage of being a dynamic methodology, allowing predictions in real time.

3. Proposed work

To develop the study, it was necessary to analyze the information collected by the Municipality of Ambato, which contained the vehicular flows in different road intersections called gauging nodes, considered as nodes of high vehicular inflow.

IOP Conf. Series: Materials Science and Engineering **844** (2020) 012008 doi:10.1088/1757-899X/844/1/012008

To carry out an analysis in greater depth, we took as an example part of the Ambato road network, in the selected sector most of the traffic problems that exist in the rest of the city are presented, the gauging nodes are marked with orange color, and the red arrows represent the groups that took the data in the gauging node and the direction in which the vehicles were directed at the moment of counting, as can be seen in Figure 1.
Because, originally, the data was collected in the form of concentration of vehicles of each node, several considerations were taken into account for the allocation of data to each of the arcs of the network.

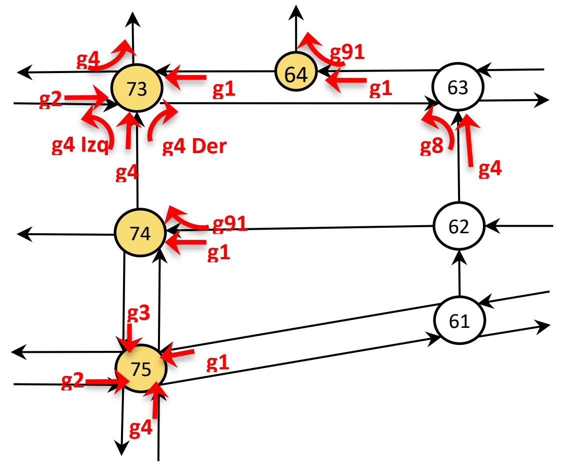
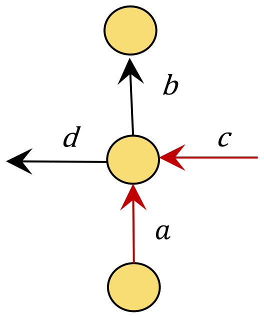


Figure 1. Scheme of the example network with the registered data groups



$$b = pa + qc \qquad (1)$$

where p is the flow fraction a that goes in the direction of b; and q is the fraction of c that goes in the direction of b.

Figure 2. Scheme for the imputations formulation

The imputation of values for the arcs were made according to the format in Figure 2. For this reason, percentages have been assigned based on the knowledge of the sector and on the experience when traveling through such intersections, trying to bring them closer to the reality and keep relationship with the initial data. Thus, if the flows a and c are available and we want to predict the flow in b, we will use a prediction of the style $b = pa + qc$.
Basically, general rules of network flows are applied, such as the one that indicates that the sum of inflows to a node must be equal to the sum of outflows. Thus, according to the scheme of Figure 2., if information is available in the arcs a, c and d, it is easy to calculate the values for the remaining arc with the formula $b = a + c - d$.
In general, it has to estimate fractions p of traffic that adopt certain routes. For this, proceed as follows. We have data $((x_1, n_1), \ldots, (x_m, n_m))$ to the hours $h_1, \ldots, h_m$, where n_i is the total number of vehicles and x_i the number of vehicles by way of interest. A priori, it is believed that $p \sim \beta e(\alpha, \beta)$. It is easily verified that, under the model $x_i \sim Bin\,(n_i, p)$, it turns out that distribution posteriori is:

$$p|data \sim \beta e\left(\alpha + \sum x_i \ , \qquad \beta + \sum (n_i - x_i)\right) \qquad (2)$$

If we need to summarize the information, we can do it through

$$\hat{p}|data = \frac{\alpha + \sum x_i}{\alpha + \beta + \sum n_i}. \qquad (3)$$

Under conditions of scarce information, we can do,

$$\hat{p}|data = \frac{1 + \sum x_i}{2 + \sum n_i}. \qquad (4)$$

Next, it is shown in Figure 3, the arcs where it was possible to impute the data in the example network marked with red color, according to the format shown in Figure 2 and the exposed on the beta-binomial model, see French and Ríos Insua [19].

IOP Conf. Series: Materials Science and Engineering **844** (2020) 012008 doi:10.1088/1757-899X/844/1/012008

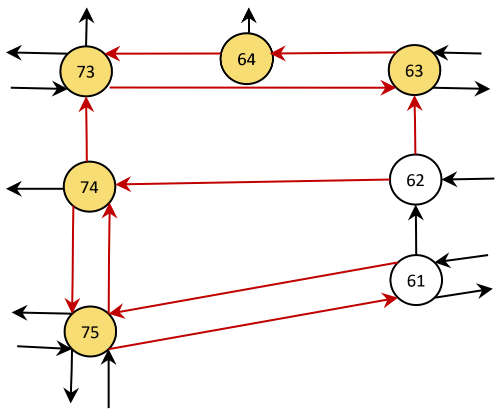


Figure 3. Scheme of the example network with imputed data

Once the data have been obtained in most of the arcs through the indicated method, the methods used in the problem are illustrated.

3.1. Principal Components Analysis in the Network Example

The Principal Components Analysis (PCA) [20] is a technique of exploratory data analysis, whose objective is the synthesis of information, or reduction of the number of variables. A main component is a linear combination of the original variables. In order to determine the nodes of greater influence within the example network, an PCA was made on the available data of the same. In this way, you can establish the most influential nodes in this network section and establish them as main nodes (components).

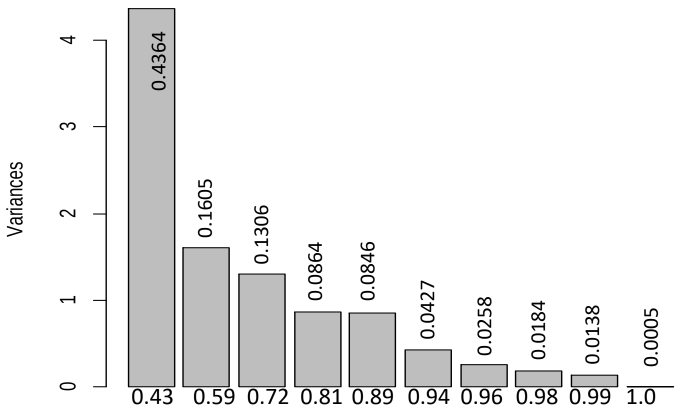


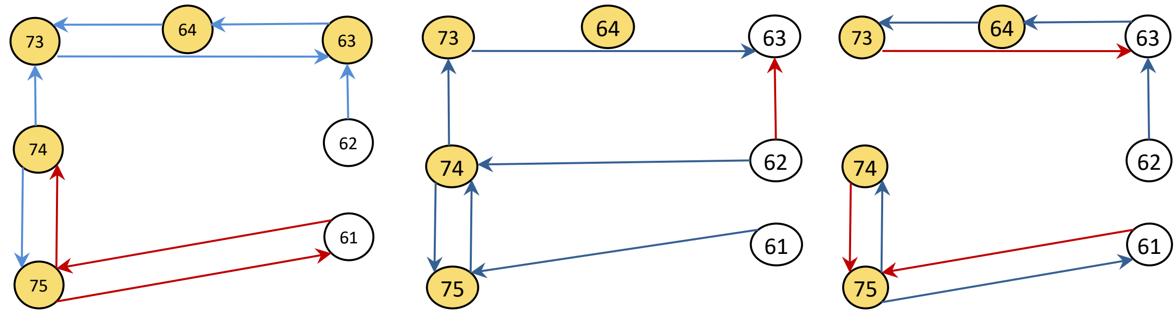
Figure 4. Bar chart of the variances explained by main components

Table 1. Matrix of results of the selected main components

Arco	PC1	PC2	PC3
X61.75	0,24	0,50	-0,11
X75.61	0,28	0,08	0,62
X75.74	0,39	0,18	0,31
X74.75	-0,32	0,24	-0,11
X62.74	-0,08	0,72	0,03
X74.73	-0,22	0,31	0,09
X73.63	-0,31	0,17	-0,33
X64.73	-0,40	0,01	0,44
X63.64	-0,41	-0,01	0,42
X62.63	-0,35	-0,10	0,12

IOP Conf. Series: Materials Science and Engineering **844** (2020) 012008 doi:10.1088/1757-899X/844/1/012008

Figure 4 contains the results of the PCA, which includes proportions of variance explained by the main components. In this case, it seems reasonable to keep the first three main components, since they explain 72.75% of the variance, the rest of the components contribute less than 10% each. The resulting charge matrix for the chosen components is shown in Table 1.
In order to be able to interpret them better, the following graph of the example network with blue arrows for negative charges, red for positive charges and those approximated to zero are not plotted, from the data matrix of **¡Error! No se encuentra el origen de la referencia.**.



a) Component PC1 b) Component PC2 c) Component PC3

Figure 5. Graph of positive and negative charges of each component

As shown in Figure 5., the first component PC1 has a positive relationship with the arcs of node 75, which would mean that there is a high vehicular demand in that node, 3 of the 4 arcs have high vehicular circulation that influences the network and the remaining arc only acts as network output. In addition, it maintains a high negative relationship with the rest of the nodes, which would suggest that the vehicles that circulate through the rest of the nodes come from other inputs to the network and that node 75 acts as an exit node.
For its part, the second main component, according to the results matrix, has a high positive relationship with eight of the ten arcs. As far as negative correlations are concerned, they are values close to 0, as well as a pair of positive values, so they have not been taken into account when analysing the component. It could be said then that most of it has a positive relationship, especially with the arcs of entry and exit of nodes 74 and 75, indicating that there is a high traffic flow in these arcs, and potential trouble spots to consider. Finally, the third component has a high correlation with the data referring to movements towards node 73.
It could be concluded that the example network has been reduced to 3 variables that interpret what is happening within it, especially to and from nodes 75 and 73, which would be the most relevants from the point of view of traffic variability, suggesting that they are the areas that require more attention because they have greater influence within the network.

3.2. Prediction models

3.2.1. Time Series: The main problem in this article is the short-term forecast of traffic on the streets of Ambato. Figure 6 shows the trend and behavior lines of each of the arcs of the example network in the three-time bands, which has provided valuable information to help select the appropriate prediction models. For this reason, it has been considered convenient to apply dynamic linear models (DLM) to create a prediction model of vehicular flow in the arcs of the network under study. West and Harrison [17] and Petris et al. [18] provide introductions to the DLMs.

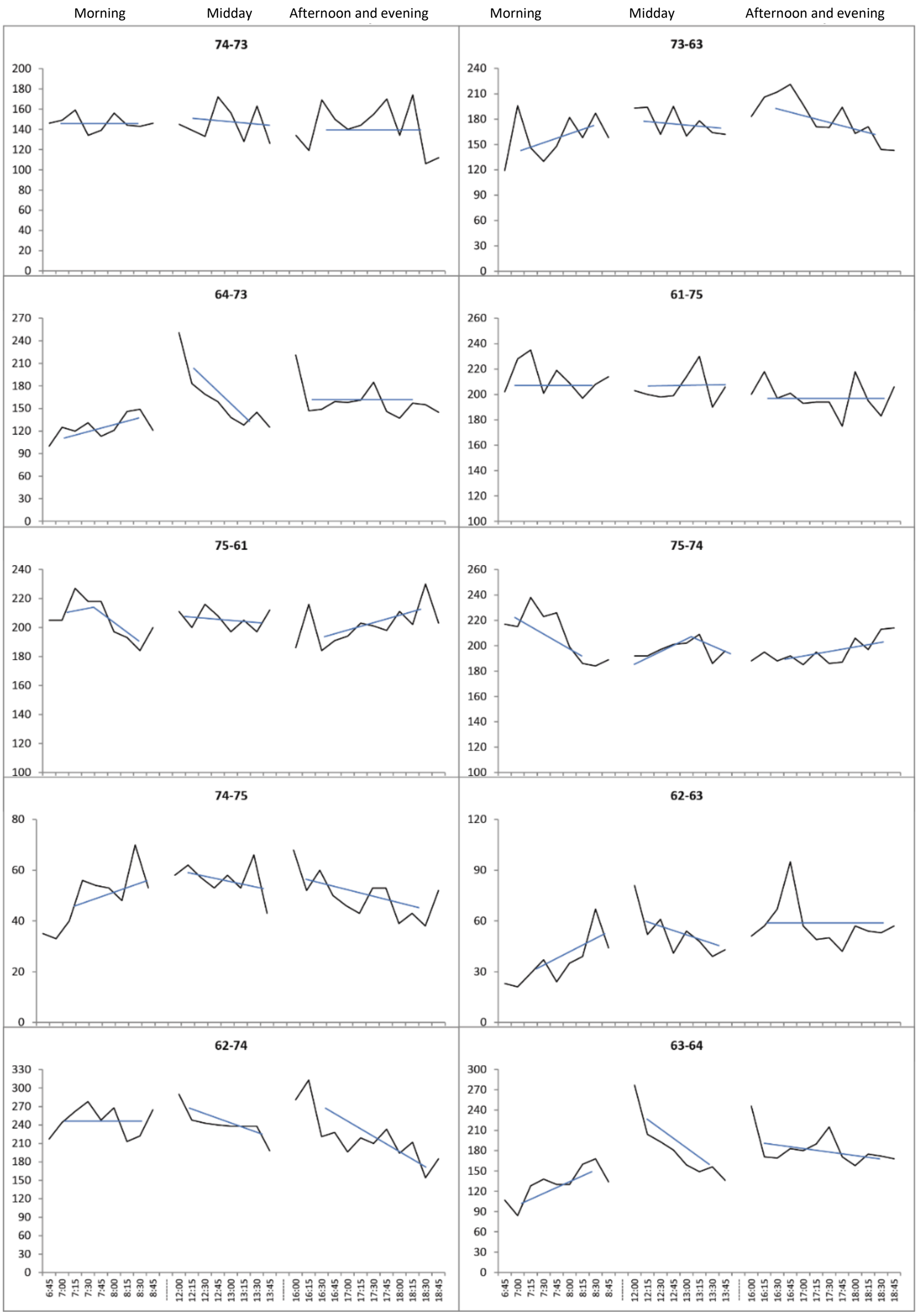


Figure 6. Time series graph of example network arcs

3.2.2. Dynamic Linear Models (DLM): The mathematical modeling of time series processes is based on dynamic models; the dynamic term refers to changes in those processes due to the passage of time as a fundamental driving force. Classic linear models are usually written as an equation:

$$Y = F'\theta + v \tag{5}$$

where Y is the response vector, F' this is the matrix of regressors or independent variables, θ this is a vector of unknown parameters, and v, it is the errors vector.

If you want to give a dynamic or evolutionary approach to the previous system, you must can be vary the elements of the model over time t and allow the set of parameters θ to evolve, in turn, with time.

Evolution takes the form of a Markov first-order process. In this way, a general normal dynamic linear model (DLM) could be written according to:

$$Y_t = F_t\theta_t + v_t\,, \qquad v_t \sim N_m(0, V_t) \tag{6}$$

$$\theta_t = G_t\theta_{t-1} + w_t\,, \qquad w_t \sim N_p(0, W_t) \tag{7}$$

$$\theta_0 \sim N_p(m_0, C_0) \tag{8}$$

which will be characterized by: $\{F_t, G_t\,, W_t\,, V_t\,\}$ (9)

where, for each instant of time t, F_t is a known vector of dimension n x 1, G_t is a known matrix $n\ x\ n$, V_t is a known variance and W_t a matrix of known variances $n\ x\ n$.

The specific model to use is a linear growth DLM, which is a state space model with: (10)

$$Y_t = \mu_t + v_t\ , \qquad v_t \sim N(0, V),$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + w_{t,1}\ , \qquad w_{t,1} \sim N\left(0, \sigma_\mu^2\right) \tag{11}$$

$$\beta_t = \beta_{t-1} + w_{t,2}\ , \qquad w_{t,2} \sim N\left(0, \sigma_\beta^2\right). \tag{12}$$

When β is equal or very concentrated in 0, the linear polynomial model is of order 1 (as in the arc 74-73, in the morning). Otherwise, it is of order 2 (as in arc 75-61, afternoon schedule). In the DLMs they allow modeling typical features in temporary series as trends (polynomials), regression and seasonal effects, see Petris et al. [18].

Application of the DLM to the example network

In Figure 6 you can see the trend lines marked with blue, which suggest the level of the polynomial be first or second order. This helps to determine one of the main parameters to obtain values by applying the dynamic linear model (DLM) of prediction.

The data of the arch 75-61 represented in Figure 6 is taken as an example, since according to its trend lines in the morning schedule it is of second order, while in the hours of noon, and afternoon and night are of the first order. The values for W and V used in the DLM have been determined by maximum likelihood; the initial value of m_0 is equal to the mean of the data vector of the corresponding time band, and the C_0 initial, the variance thereof. The DLM mentioned has been applied in R, and the following models and graphs have been obtained:

Morning schedule (6:45 - 7:00 - 7:15 - 7:30 - 7:45 - 8:00 - 8:15 - 8:30 - 8:45)

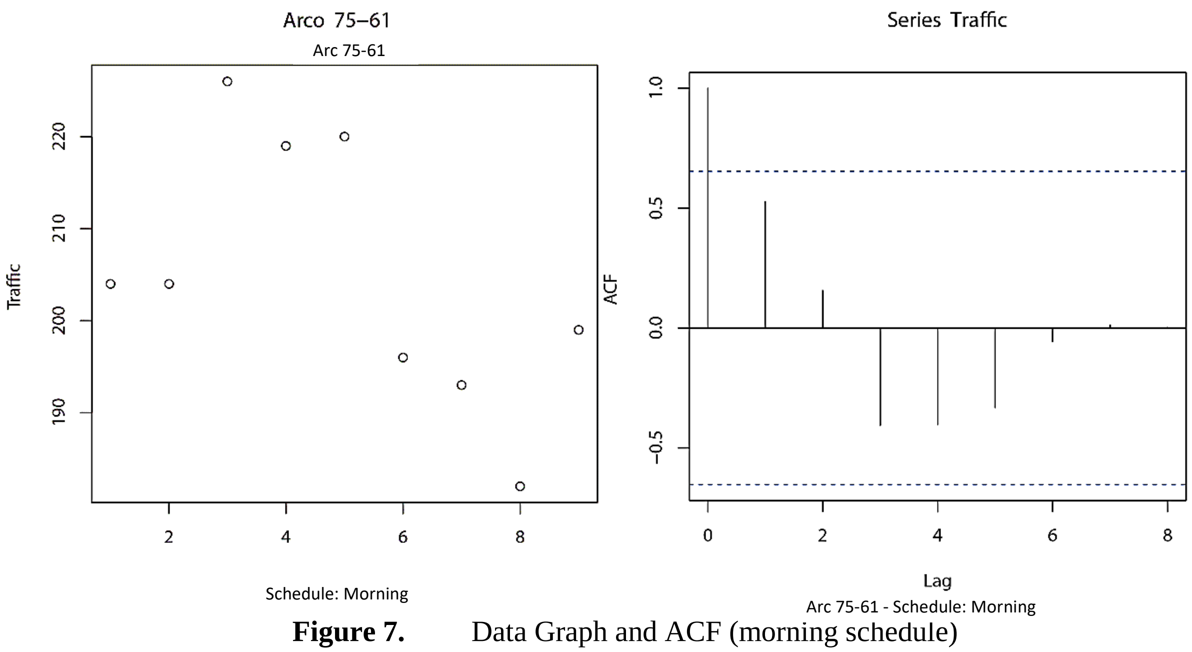


Figure 7. Data Graph and ACF (morning schedule)

IOP Conf. Series: Materials Science and Engineering **844** (2020) 012008 doi:10.1088/1757-899X/844/1/012008

In Figure 7 you can see, in the first table, the scatter plot of the morning data. The second table shows the autocorrelation function suggested by the trend line of order 2 as indicated in Figure 6 for the arc 75-61 in the morning schedule.

Figure 8 shows the predicted values marked in gray that follow a trend similar to the original values in black, with a good fit.

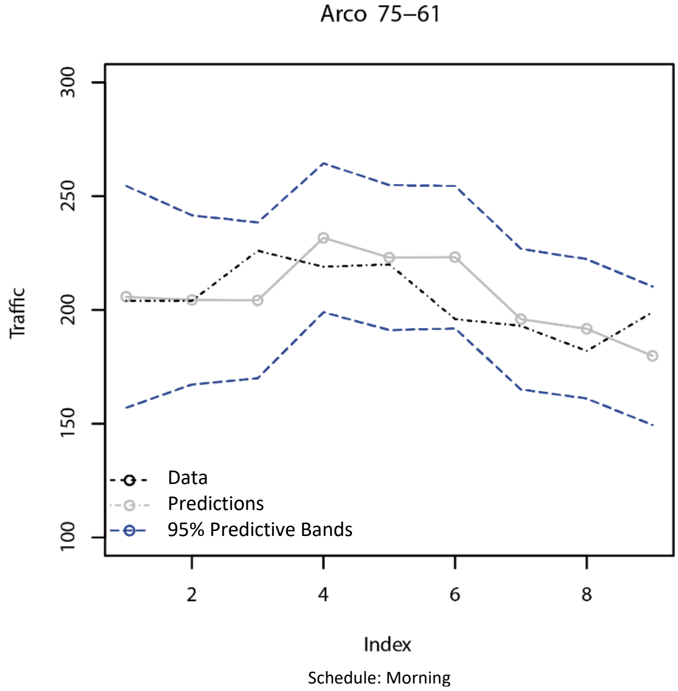


Figure 8. Graph of original and calculated data when applying the DLM (morning schedule)

In Figures 7 and 8, as can be seen, the generic DLM used is adjusted satisfactorily to predict the following quantities of vehicles that may circulate in the following periods of time. Similarly, under the same conditions, they apply to the rest of the arcs and thus predict 30, 45, 60, ..., n minutes.

3.2.3. Saturation Prediction: One of the most relevant uses of the previous DLM is a prediction of saturation in the corresponding network paths. It is defined as the capacity of a road at the maximum hourly intensity of people or vehicles when crossing a uniform stretch of a lane or roadway for a defined period of time under the prevailing road conditions, traffic and regulatory systems [21].

To make the calculation of maximum flow in each arc, considering that the capacity of a lane is the maximum number of vehicles that can pass through it, assuming a uniform speed.
Generally, it is expressed in vehicles per hour. Then the capacity in vehicles per hour would be:

$$C = \frac{v}{s}1000, \tag{13}$$

where C: maximum capacity of a lane, v: speed in km/h, s: minimum average separation in meters between the front or rear parts of two successive vehicles to determine the speed. The value of s is a function of the length of the vehicles, the reaction time of the drivers and the braking distance, can be expressed as a function of v using a formula of the type:

$$s = a + bv + cv^2 \tag{14}$$

The independent term (a) corresponds to the length of the vehicles, the term (bv) to conductors reaction time and (cv^2) to the braking distance. In a study linked in England by Smeed [22], he found the following formula for the most frequent separation based on speed:

$$s = 8 + 0.2v + 0.003v^2. \tag{15}$$

Replacing s in the formula of capacity C, we would have:

$$C = \frac{v}{8 + 0.2v + 0.003v^2} * 1000 \tag{16}$$

When the vehicle running speed is below 25km/h it becomes intolerable for the driver, so we have considered the use of the minimum tolerable to drive. Therefore, $v = 25Km/h$ and $C = 1681\ vehicles/hour$. Thus, we would obtain that for every 15 min for arcs with a lane, the maximum flow is 420, and 70% of its capacity would be 294, this information being necessary since when exceeding this percentage saturation problems begin to exist. The saturation probabilities are calculated by:

$$\Pr(Z > 0{,}7f), \tag{17}$$

where Z is the predictive distribution of the corresponding DLM.

After determining the probability values for each arc, it is observed in Figure 9, in the first table it indicates that in the arc 62-63 there is a very low probability of saturation, since the maximum value reached is 0.5 in the first two quarters of an hour in the morning and in the rest of the vehicle flow measurement periods, the probability is close to 0.

Regarding the arc 62-74, the opposite can be observed, in the first two periods of time the probability of saturation is higher than 50%, while in the rest of the periods it indicates that it is possibly always saturated, with a probability close to 1. This would suggest the need to interact in such traffic, through public service, limitations of traffic and other similar measures.

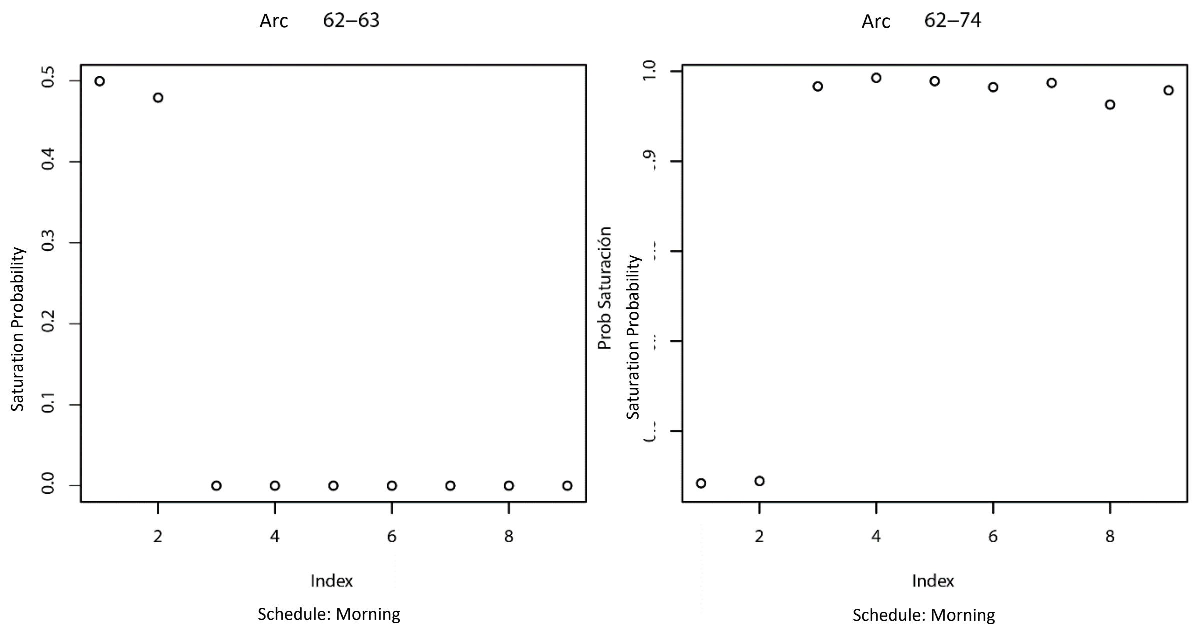
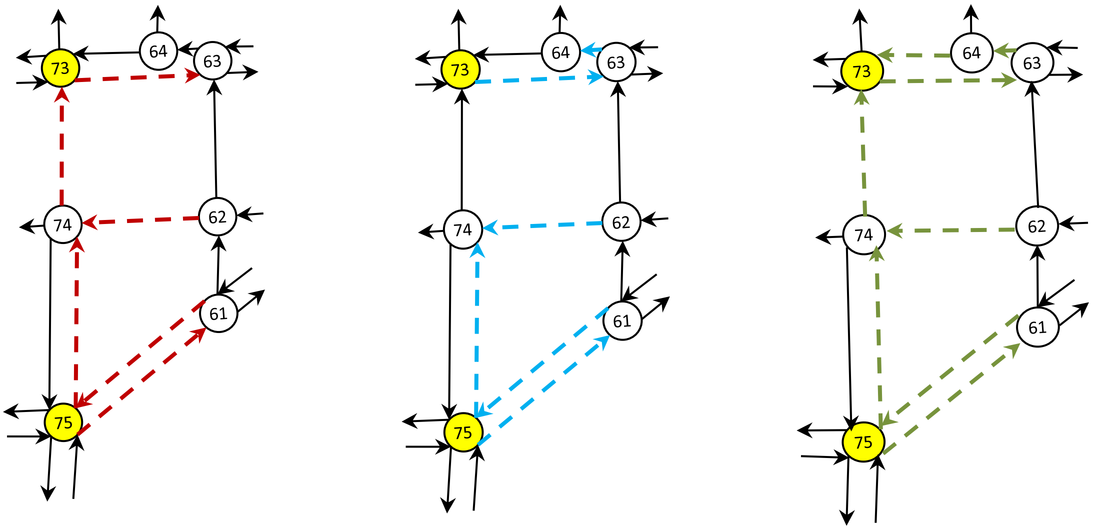


Figure 9. Saturation probability graph

4. Results

Finally, in Figure 10 you can see the results obtained after applying the methods described above in the example network. With the help of the PCA, the conflict nodes marked with yellow colour were determined, the same ones that require more attention within the network, therefore, improving the vehicular flow in these nodes would probably cause an improvement in the flow throughout the network. The arcs with greater probability of saturation determined with the help of the DLM, are marked with dashed lines in red, blue and green colors for each schedule respectively. This information can help to take the necessary corrective measures to avoid the probable problems of congestion.



a) Schedule: morning b) Schedule: midday c) Schedule: afternoon and evening

Figure 10. Saturation probability graph

Applying the indicated methodology to a larger road network in the city, the results are shown in Figure 11. were obtained, where the flows with the highest probability of vehicular congestion can be clearly seen, becoming priority places where measures must be applied to avoid saturation.

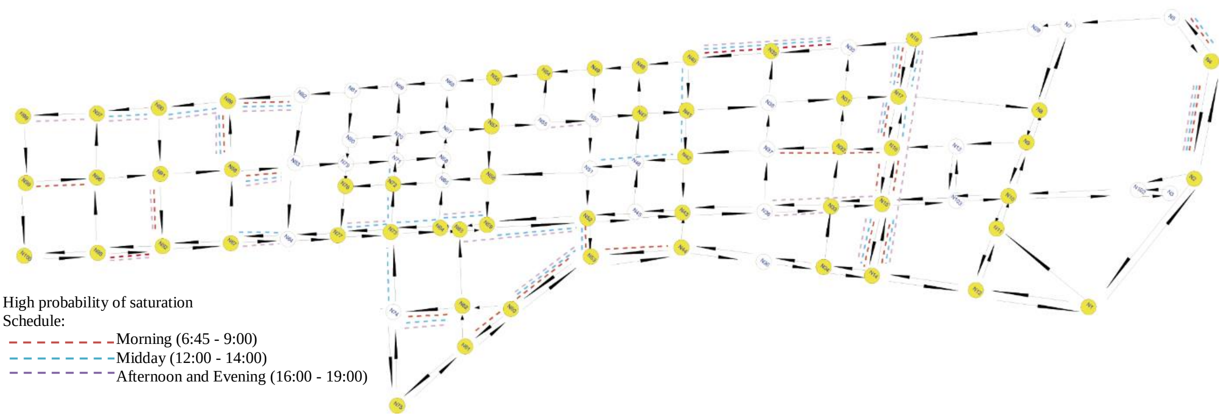


Figure 11. Hot spots and Arcs with high saturation probability

5. Conclusion

After applying the PCA to the study network, several of the so-called hot spots or with greater incidence in the vehicular flow are determined. Therefore, the nodes in which more care must be taken when making decisions regarding vehicle traffic planning. Another alternative offered by these points is determine the possible key locations where vehicle counting devices can be placed to obtain a traffic data history in the city and thus achieve traffic predictions with a lower margin of error by applying the DLMs.
The saturation probability values obtained helps to determine the possibility of congestion problems occurring in the different periods of time, with which the necessary corrective measures can be taken to avoid them. For example, to locate traffic lights, to eliminate parking lanes, to change the direction of tsome arcs in certain hours or permanently, etc., and thus allows the flow of vehicles is optimal and thus avoid congestion in strategic locations.

6. Future work

Based on the prediction model provided by the MDLs at the traffic measurement points, the concept of dynamic O-D matrices is intended to predict the main routes in the city. The concept of static O-D matrix can be seen, p. ex. in Vardi [23]. Having thus to fit to a model of simulation to experiment different measures to improve the vehicular traffic.

Acknowledgments
The work of Marcelo Pilamunga and Milton Aldás has been possible made by the project "1572-CU-P-2017 Models of management and optimization of traffic".

References

[1] Nel-lo O 1998 *Los confines de la ciudad sin confines. Estructura urbana y límites administrativos en la ciudad difusa* de *La ciudad dispersa. Suburbanización y nuevas periferias* ed Monclús J. (Barcelona: Centro de Cultura Contemporánea de Barcelona) p 35

[2] Taha H A 2004 *Investigación de Operaciones*, (México: Pearson Educación)

[3] Fernandez R and Valenzuela E 2004 Gestión Ambiental de Tránsito: Cómo la Ingeniería de Transporte *Eure* **29** 97

[4] Castillo E, Pedregal P, García R and Alguacil N 2002 *Formulación y Resolución de Modelos de Programación Matemática en Ingeniería y Ciencia* (México DF: Alfaomega Grupo Editor)

[5] Tebaldi C and West M 1998 *J American Stats Assoc* **93** 557

[6] Akcelik R 1981 *Analysis, Traffic Signals: Capacity and Timing* (Vermont: ARRB Group Limited)

[7] IHT 1997 *Transport in the Urban Environment* (London: The Institution of Highway and Transportation)

[8] Williams B M and Hoel L A 1999 *Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results* Ph.D. Dissertation, Dept. Civil Eng., University of Virginia, Charlottesville, 1999.

[9] Head K L 1995 *Transp Research Rec* **1510** p 46

[10] Lee S and Fambro D B 1999 *Transp Research Rec* **1678** 179

[11] Okutani I and Stephanedes Y J, 1984 *Transp Research Part B: Method* **18** p 1.

[12] Banks J 1998 *Handbook of Simulation: Principles, Methodology, Advances, Applications, and Practice* (Atlanta, Georgia: EMP Books)

[13] Barceló J 2010 *Fundamentals of Traffic Simulation* (California: Springer)

[14] Willumsen L 1978 *Estimation of an O-D Matrix from Traffic Counts – A Review* (University of Leeds,Institute of Transport Studies)

[15] Van derVoort C, Dougherty M and Watson S 1996 *Transp Research Part C: Emerging Tech* **4** p 307

[16] Yu G QHu J M Zhang C S, Zhuang L K and Song J Y 2003 *Intelligent Vehicles Symposium, 2003. Proceedings IEEE,* p 208

[17] West M and Harrison J 2005 *Bayesian Forecasting and Dynamic Models* (New York: Springer)

[18] Petris G, Petrone S and P. Campagnoli P 2009 *Dynamic Linear Models with R* (New York: Springer)

[19] French S and Ríos Insua D 2000 *Statistical decision theory* (Oxford: Oxford University Press Inc)

[20] Hair J, Anderson R and Tatham R 1987 *Multivariate data analysis* (New York: Macmillan Pub. Co)

[21] National Academy of Sciencies 2000 *Highway Capacity Manual (HCM2000), Transp Research Board*

[22] Smeed R and Bennett G 1949 Research on Road Safety and Traffic Flow *Inst Civil Engineers - Road Paper* **29**

[23] Vardi Y 1996 *J American Stats Assoc* **91** p 365