A MODEL TO ANALYSE VMS EFFECTS ON FLOW IN MOTORWAY

LIVIO FLORIO
Dipartimento di Sistemi di Trasporto e Movimentazione, DSTM,
Politecnico di Milano, P.za L. Da Vinci, 32, 30123, ITALY

LORENZO MUSSONE
Dipartimento di Sistemi di Trasporto e Movimentazione, DSTM,
Politecnico di Milano, P.za L. Da Vinci, 32, 30123, ITALY
Tel +39 -2 - 23996612 fax +39 - 2 - 23996612 e-mail mussone@cdc8g5.cdc.polimi.it

Abstract

Influence of variable message signs (VMS) on flow characteristics, and on drive behaviour, is an important topic till now not completely investigated. A method already applied to analyse flow in motorway is extended to insert also the effects of VMS on traffic flow.
The models are worked out by means of using multi-layered feedforward neural networks. The first analyses the density flow relationship and the second the distribution of flow by lane. Input variables are density, total flow , speed, brightness, clearance, meteorological conditions, percentage of heavy vehicles and presence of messages on VMS. Output variables are flows by lane. Speed, flow and density values, necessary to ask the second models, are obtained from the first one.
Different scenarios are prepared to ask the model varying percentages of heavy vehicles, meteorological conditions, brightness, clearance and presence of VMS messages. Results show a great influence of VMS on capacity curve but not very significant on lane occupation.

INTRODUCTION

Investigation of the flow-density relationships by lane is an important topic both for control purposes and for planning. These relationships are usually treated as a unique one (summing up the effects of all lanes) but they can have very different shapes according to flow dynamic or vehicle composition. Up to now, how and how much these differences are conditioned by the presence of traffic disturbances, in particular those produced by a toll-gate, has not yet been investigated in depth.

A lot of models were developed to describe lane occupation of vehicles (Gipps, 1986; Alvarez et al., 1990; Schmidt et al., 1991; Wemple et al., 1991). They are based on hypotheses related to driver behaviour and on the used flow model such as, for example, the hydrodynamic or the car-following model. Real flow data are used only to fit model parameters but not to reveal particular driver behaviour.

Variable message signs (VMS) are especially suitable where fixed signs are not sufficient to control dynamic flow; they instead can intervene quickly to follow sudden change in flow or environmental conditions, or presence of incidents. Messages can be used to regulate speed or to change lane occupation. First experience in VMS started in the last 70s. One of the first signalling plant was in Aichelberg (Baden-Württemberg, Germany) and it was designed to tackle congestion in a quite short section (8km) and to improve just flow safety by avoiding too high dispersion of speed. Another experience is by Rijkswterstaat (Remeijin, 1984) who examined the aims and design of the motorway control and signalling system with VMS in detail. Benefits of this type of control is bettering of safety and improvement of capacity especially for bottlenecks caused by road-works or incidents (Klijnhout, 1984).

An other approach investigates of VMS on driver route choice (but not on flow characteristics) such as: (Bonsall and Merrall, 1996) tackled this problem by using a simulator to gather information on driver responses to a wide range of messages; (Hato *et al.*, 1996) investigated especially driver reactions to VMS information by analysing data of stated preferences surveys.

The aim of this paper is to carry out a model of lane occupation without making any a priori assumption about driver behaviour, influence on driving due to flow composition, environmental conditions and presence of VMS. Particularly the role of VMS is investigated.

In authors' previous works (Mussone, 1995; Florio and Mussone, 1996b) flow models for different sections and meteorological conditions are proposed by means of using neural networks and these models have shown the influence of meteorological data and flow composition to determine the capacity value for a single section. In the previous works the authors used neural network technique with the only assumption that a function (linear or not linear) relating variables exists and can describe that process. In the same way the relationships among flow distribution over lanes and flow, environmental data can be treated. Neural network approach guarantees approximability of any continuous function and therefore are well suited for the above-mentioned objective.

Feedforward neural networks (in the multi-layer perceptron version, MLP) were applied to work out the models used in this paper. MLP capabilities in the field of non-linear dynamic systems modelling are well known. These networks are well described in technical papers and a lot of theorems state that multi-layered feedforward neural networks by using neuron transfer functions of a sigmoidal type and linear input combinations, can approximate any continuous function. Input variables are density, total flow, speed, brightness, clearance, meteorological conditions, percentage of heavy vehicles by lane and presence of messages on VMS. Output variables for the first model are the three percentage values of flow by lane.

DATA COLLECTION AND EXTRACTION

The data used in the present study, were collected in the "Easy Driver" environment (FIAT, 1992), a traffic control system which was employed in Italy on the Padua-Mestre (Venice) motorway, a rectilinear section from the tollgate of Dolo to the end-motorway tollgate of Mestre, over a distance of about 11 km. Data collection has already been described in previous papers (Florio and Mussone, 1996a, 1996b, 1997) and here only some synthetic information are reported.

The subdivision of the flow into light vehicles and heavy goods vehicles was performed on the basis of ANAS Italian code, thus vehicles belonging to the first three categories (up to 5.5 meter long) were considered as light, whereas those belonging to the last three categories (longer than 5.5 meters) were considered as heavy. Information on weather conditions was not monitored on all the sections and the sampling period adopted within file data was different. Therefore a method to group together data is necessary considering both time and space. For the present study a section 3 km far from the tollgate of Mestre is considered. The spatial association concerns the files of data relative to brightness and meteorological conditions. Records available are almost 360,000.

Another topic is related to data distribution over the considered variables. In fact, the various flow conditions detected on a motorway are usually fairly different in their frequency. The samples of unstable flow or near to capacity, for example, are far less numerous that those relating to stable flow. Since the aim is to represent all the features of the flow, the extraction of the sample for the learning process should be preceded by a data classification, grouping data into categories.

For this purpose, a further variable was created, related to a first associating criterion which is based on the detection of the following information: presence of rain, snow/ice, percentage of heavy goods vehicles (the mean value among the three lanes), visibility, brightness and VMS presence. Considering these values as a part (one or two bits) of complex information (byte), the category which the datum belongs to, is then represented by a binary word, which is to be built up for this purpose. In the present case, 8 bits are enough to identify the various cases.

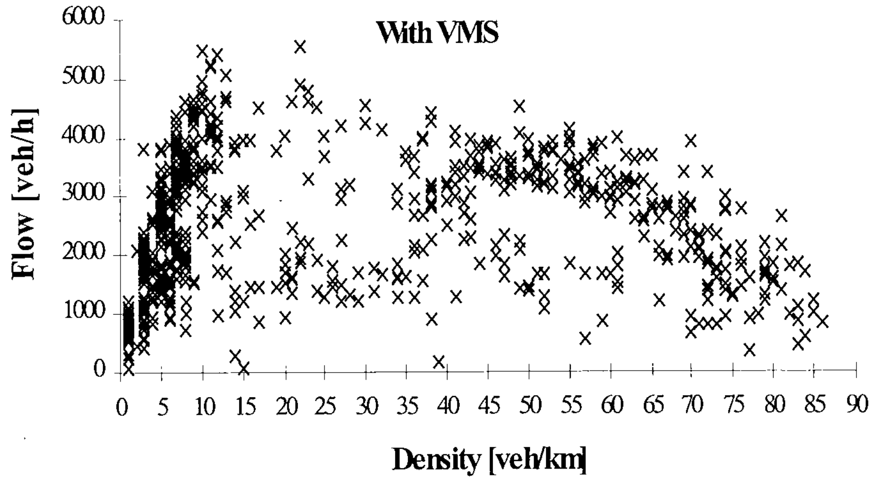
A second classification was performed according to density. A numeric variable represents a density category. Since all the characteristics of the process should be represented, the extraction of data is performed by class of data including a maximum number of elements, consistent with the mean frequency of classes. The number of extractions from each class was found to consist of 5 and the extracted data are quite homogenous around 1200.

With regard to the learning process and to the cross-validation technique of the neural network, two sets of data, one for learning and one for validation, were created by means of an extraction from all the data. Data are then normalised according to their highest values; results are therefore normalised to those values. For the traffic flow, the maximum value of 6,000 veh./h for the three lanes was assumed; for speed the maximum value was 200 km/h and for density 150 veh./km per lane. In fig. 1 extracted data subdivided into two sets, with and without the presence of VMS, can be easy compared. The two data sets are very different for high value of density and they suggest also a different flow dynamic whether a VMS is present or not: values in fig. 1, a) with high density (greater than 35 veh./km) are very unusual in a flow-density relationship and probably depend on

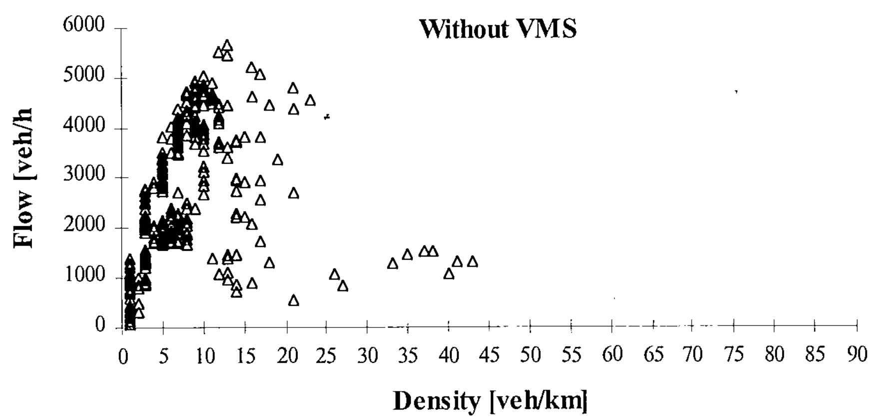
VMS effects; instead, those with low flow (less than 2000 veh/h) and high density (between 15 and 60 veh/km) are due to different percentages of heavy vehicles. There is not a great difference for low values of density (less than 15 veh/km) when flow is free.

THE NEURAL NETWORK MODELS

Models are worked out by means of using feedforward neural networks (multi-layer perceptrons, MLP), whose capabilities in the field of non-linear dynamic systems modelling are well known (Mussone, 1995).



a)



b)

Figure 1 - Collected data divided according presence a) or not b) of VMS.

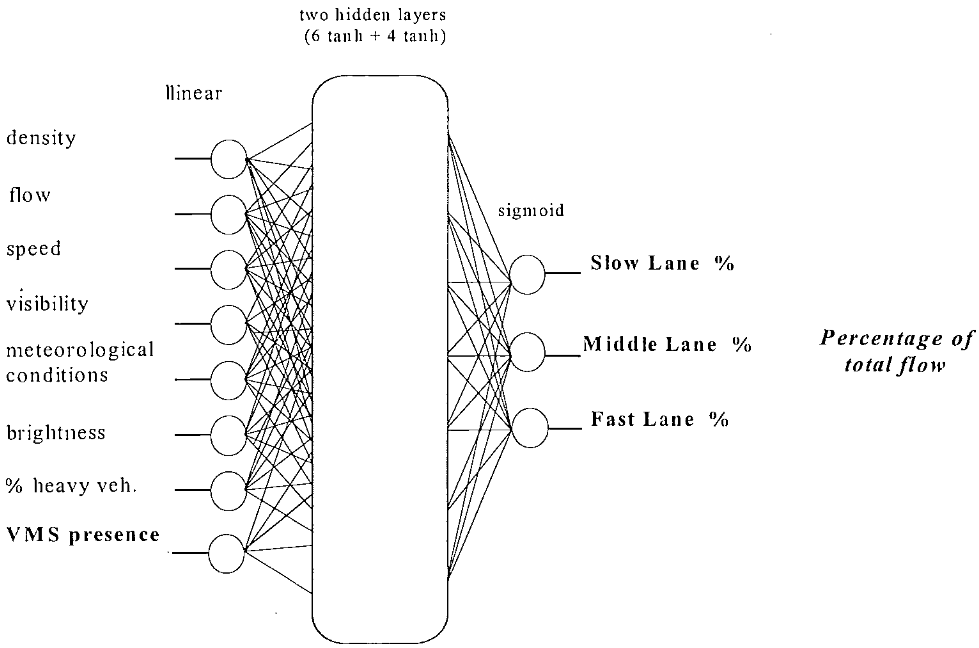


Figure 2 - The neural network scheme for the occupation model.

These networks are well described in technical papers and a lot of theorems state that multi-layered feedforward neural networks (with at least one hidden layer), by using neuron transfer functions of a sigmoidal type and linear input combinations, can approximate any function which belongs to L^2 space with a small margin of error (Cybenko, 1989; Hornik et al., 1990; Hornik, 1991; Girosi et al, 1991; Leshno et al, 1993). These models are referred to without inferring anything from their physical characteristics, using the so called "black-box" approach. It should be said that the attempt to deduce anything from the values of the connections appears a long and difficult operation which gives few useful results. The number of hidden neurons, and the number of layers needed to obtain the desired approximation is still being studied.

The paradigm used in learning networks was backpropagation (BP) (Rumelhart et al., 1986): it is a heuristic solution to the training problem. Many authors, such as (Masters, 1993) underline the difficulties of training, and in particular the problems of overfitting or overtraining which adversely affect the performance of a neural network.
The building of a feedforward neural network model, with backpropagation learning, requires the determination of the proper number of hidden neurons and layers in an attempt to minimise the error on both learning and test data. The proper number of neurons refers to the minimisation of the output error i.e. to the best performance of the network.

There are many techniques to an optimum exploitation of data: cross-validation, Jacknife and Bootstrap (Efron, 1982), (Weiss et al., 1991). Cross-validation is the most known technique and it essentially consists in dividing data into two disjoint sets, one for training and the other for validation. Different topological networks are evaluated on the basis of these two sets. The Jacknife technique (Efron, 1982) divides data into N disjoint sets; each of these sets contains one element for validation and N-1 for training. Then N+1 identical neural networks are created; only one network, called master, is trained on the basis of all data; the other networks, called slaves, are trained and tested on the basis of the N disjoint sets. Test error evaluated on slaves networks is an estimation of generalization capability of the master network. In Bootstrap technique extraction of the two sets is randomly and repeated more times. Then, for validation cross-validation or Jacknife technique can be used. It must be said that some authors suggest the use of a third set to evaluate performance because the test set is already used to stop training and therefore in the training phase. But this is not so crucial in evaluation of performances.

The optimum configurations of the network (in the sense of performance) are necessarily related to the phenomenon of overfitting. Recent studies suggest that overfitting occurs essentially because of two main reasons: firstly, the network is not properly sized compared to the available data; secondly, data are not sufficiently representative of the function to be implemented, thus the two sets of data, the test and train sets, are remarkably different from each other.

To work out the models the authors used the cross-validation technique which appears to fit better to the dimension of data set. To overcome the above-mentioned learning problems, a lot of network configurations with one or two hidden layers, with a different number of hidden neurons and different transfer function (sigmoid, tangent hyperbolic and sine function) were built up. The first layer, the input one, has always neurons with linear transfer function.

After several trials the optimal models are carried out. The number of learning iterations varies according to the model. In Tab. 1 these features and other performance (on the test set) of the optimal models are reported. The first column refers to the motorway section considered, the following three refers to the characteristic of the hidden layers (number of hidden layers and neurons, and neuron transfer function); the number of iterations is the number of learning iterations for the model; correlation is the correlation between desired and real data calculated over the three lanes; RMSE is the root mean square error calculated according to eqn (1):

$$\mathrm{RMSE} = \sqrt{\left(\frac{\sum (x-y)^2}{n}\right)} \tag{1}$$

where y,x, n are the real data, the desired data, the dimension of the test set respectively; the summation is over the n data. The best model (fig. 2) has two hidden layers: the first has 6 neurons and the second 4, in both layers neurons have a tangent hyperbolic transfer function.

Table 1: Features and test errors of lane occupation model.

Hidden Layers	Hidden neurons	Transfer function	Iterations No.	RMSE
1	6	tanh	60.000	0.10205
2	6 + 4	tanh + tanh	90.000	0.09647

Table 2: Features and test errors of flow-speed-density model.

Hidden Layers	Hidden neurons	Transfer function	Iterations No.	RMSE
1	6	tanh	80.000	0.083858
1	7	tanh	80.000	0.085826
2	4+3	tanh + sigmoid	10.000	0.127746
2	4+3	tanh + tanh	50.000	0.124887

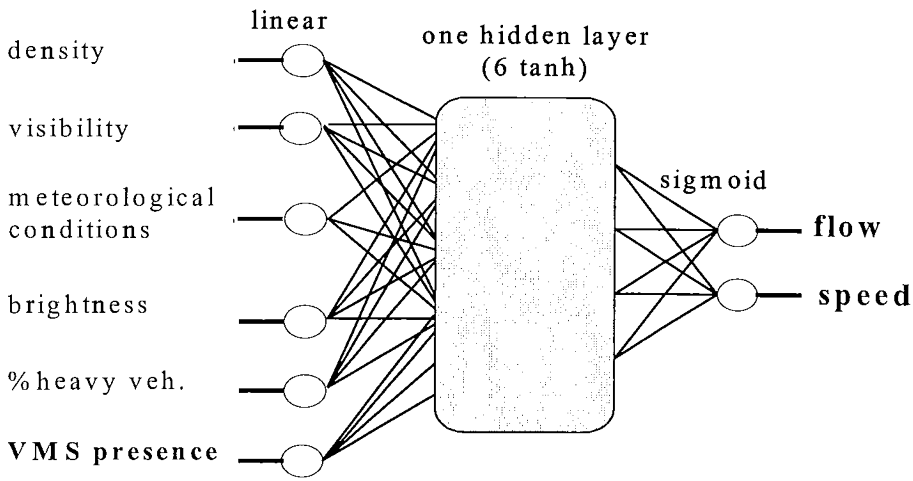


Figure 3 - The model used to calculate the fundamental law relating flow, speed and density variables.

It must be noted that the sum of the three percentages is almost always correct and quite close to 1 with a average error of 0.028 and a standard deviation of 0.030; maximum error of 0.12 is reached only in very a few cases. This result is particularly interesting because in the learning data this information was not explicit. Nothing it is possible to say about the use of tangent hyperbolic or sigmoid transfer function in the hidden layer: probably it depends on data distribution but until now it cannot be demonstrated.

Flow, speed, and density values, necessary to ask the models, are obtained from a model of density-flow and speed relationships (fig 3) worked out as explained in previous papers (Mussone, 1995; Florio and Mussone, 1995) and which relates these variables to the environmental variables (visibility, meteorological conditions, brightness) and the percentage of heavy vehicles, one for each section considered.

RESULTS

The model analysis reflects the neural network black-box approach so that a large number of input cases are prepared to ask the network and to know its characteristics and the following figures are only a small number of those obtainable by the model.

The cases consider different values of density from 1 to 100 veh./km (higher values are not very meaningful in this context because they represent only congestion where flow dynamic is rather unpredictable leading to stop-and-go situations).

In fig. 4 the output of the second model is shown when VMS are present and not. Speed is not shown because doesn't vary much with input variables. Three extreme cases are considered: standard conditions, brightness reduced to 0.1 of its maximum (practically night-time) and a heavy vehicles percentage on flow of 40%. The other variables affect the curve, in the section considered for the models, to a less significant degree.

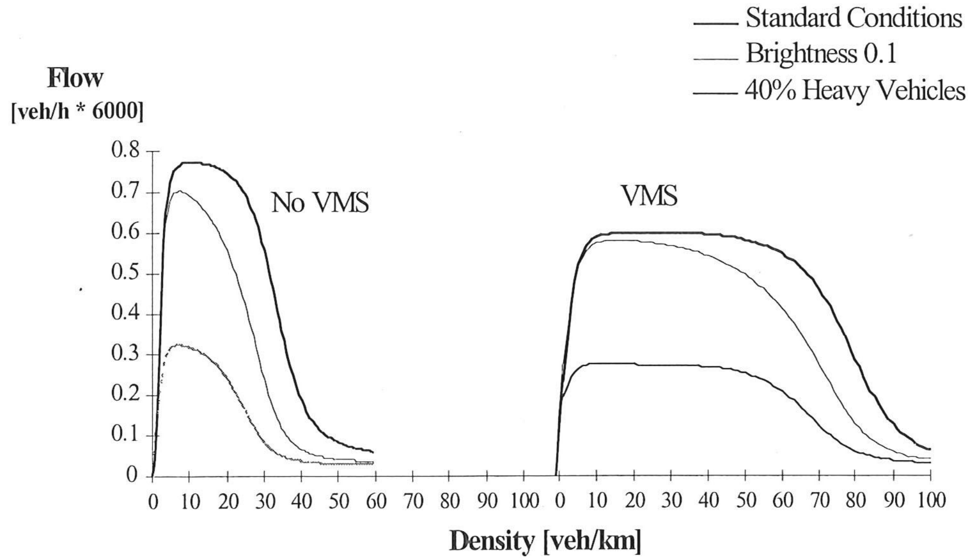


Figure 4 - Flow-density relationships of the first model.

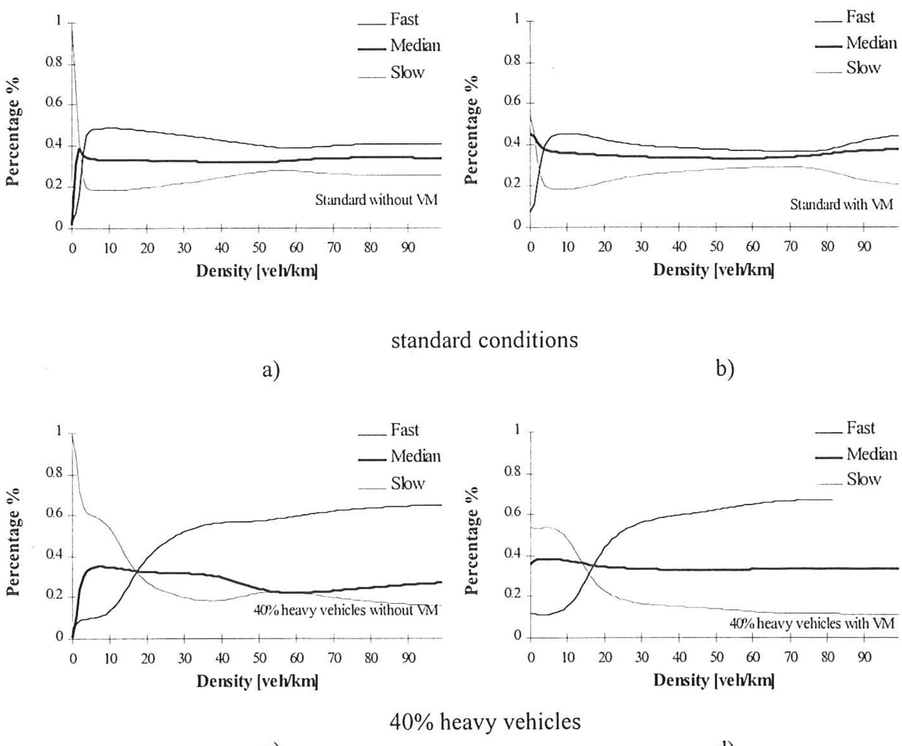


Figure 5 - Comparisons of occupation by lane with (b) and d)) and without VMS (a) and c)) for standard conditions and 40% of heavy vehicles.

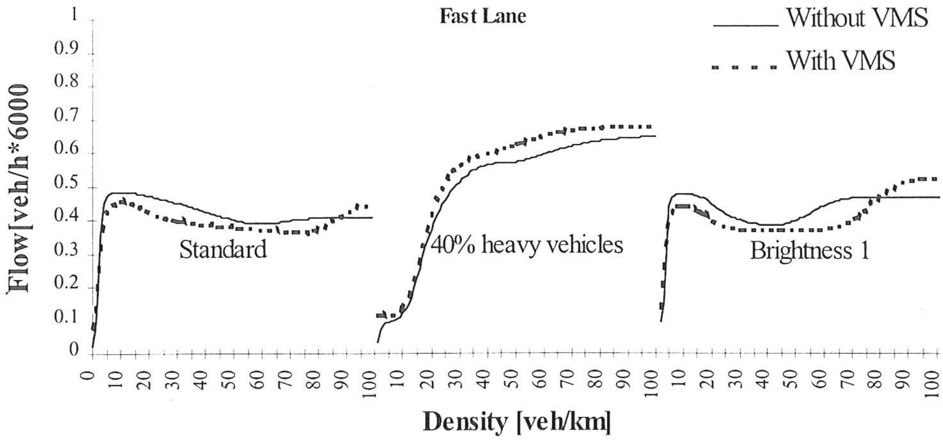


Figure 6 - Fast lane without and with VMS standard, 40% and brightness equal to 1.

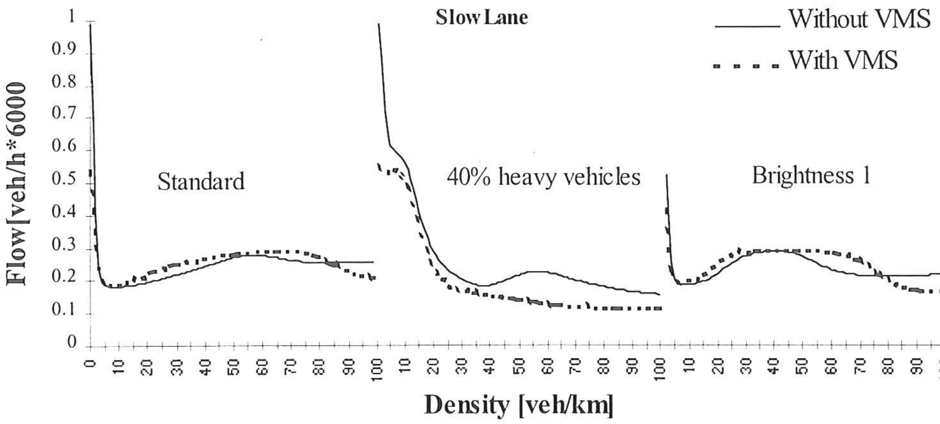


Figure 7 - Slow lane without and with VMS standard, 40% and brightness equal to 1.

The influence of heavy vehicles percentage and brightness is clearly shown by the figure but it is more interesting to note the different extension of the capacity zone: when VMS is present the capacity is reduced but it is constant for a double or triple interval of density. It decreases only when density is about 60 veh/km which is a value when dynamic usually cannot be other than unstable. This effect is more evident when not standard conditions are examined: for 40% of heavy vehicles case capacity doesn't vary much so as the effect of VMS is more relevant. This result is very important because maintaining the same value of capacity for a wider range of density means that flow becomes more stable and some problems related to unstable flow (such as incidents or catastrophic dynamics) can be avoided.

In fig. 5 the output of the first model is shown when considering standard conditions and a 40% of heavy vehicles with and without VMS. The middle lane has the most stable values especially when VMS is present and it is not affected by heavy vehicles percentage. The other two lanes have a very different shape: the slow lane starts with higher values (near 1 without VMS and 0.55 with VMS) then, increasing density, decreases to low values (between 0.1 and 0.2); the fast lane instead starts with lower values and then increases to reach the highest ones (close to 0.5). The point where curves intersect (which could represent the point where free flow finishes) changes with the scenario examined but it doesn't change significantly with VMS presence.

Another effect of VMS presence consists in limiting differences of percentage occupation between lanes at low density: it is particularly evident for the slow and median lane, the values of which go from 1 to 0.55 and from 0.02 to 0.4 respectively. In fig. 6 and 7 VMS comparisons is shown for the fast and the slow lane (the median lane changes in a still lesser extent). It is possible to observe that VMS doesn't affect much occupation except for very low values of density. In these figures the brightness scenario is included and it is rather similar to the standard conditions scenario.

FINAL REMARKS

Results highlight the influence on lane occupation of the proposed input parameters. Density, brightness and percentage of heavy vehicles lead to considerable modifications in the curves, and the consequent effects are clearly distinguishable. In respect of previous flow-density models, without the VMS presence as input variable, results are statistically more significant assessing that the process is really affected by VMS.

It is possible to assert the following remarks:

- the presence of a VMS message affects significantly the shape of the relationship flow-density;
- the fast lane, in standard conditions, has the highest flow but when VMS messages are present it is quite similar to the median lane flow especially for density higher than 20 veh/km;
- increasing the heavy vehicles percentages, flow over lanes changes drastically (it must be remembered that in the Italian motorways heavy vehicles must run on the slow lane and they can use the median lane only to overtake), there is a point of equal percentage for about 18 veh/km;
- decreasing brightness affects flow-density curve but less distribution by lane;
- decreasing visibility the median lane flow (and less the fast lane) increases;
- rain affects curves but not significantly the percentage of distribution over lanes.

The last result is surprising and it may be explained by the fact that this parameter reduces capacity and decreases speed but does not affect the driver choice of the lane. VMS affect flow but less lane occupation though generally VMS presence reduces differences between the percentages of lane occupation and stabilises their values in a more limited range. Results are obviously valid for the section examined, generalisation to other sections is possible but not sure because the other sections could be affected by different flow dynamics due to the presence of curves, merging points and so on.

AKNOWLEDGEMENTS

Thanks are due to "Società delle Autostrade Venezia e Padova" for providing data. This paper is partially funded by MURST.

REFERENCES

Alvarez A., Brey J.J., Casado J.M. (1990) A simulation model for traffic flow with passing. **Transpn. Res.**, Vol. 24B, **3**, 193-202.

Bonsall P.W., Merrall A.C. (1996) Analysing and modelling the influence of roadsidevariable message displays on drivers' routes choice. In **Proceedings of the 7th WCTR**, eds. D. Hensher, J. King, and T.Oum, Vol. 1, pp.11-25, Elsevier Science.

Cybenko G. (1989) Approximation by superpositions of sigmoidal functions. **Mathematics control, signals and systems**, **2**, 303-314.

Efron B. (1982) **The Jacknife, the Bootstrap and other resampling plans**. Society for sindustrial and applied mathematics, 1982.

Florio L., Mussone L. (1995a) Neural network models for classification and forecasting of highway traffic flow stability. In: **Transportation Systems: Theory and Application of advanced technology**, *eds* Liu B., Blosseville J.M., pp. 773-784, A postprint volume from the 7th IFAC Symposium, Tianajin, August 24-26, 1994, Pergamon.

Florio L., Mussone L. (1996a) Neural network models for classification and forecasting of highway traffic flow stability. **Control Engin. Practice**, Vol. 4, No.**2**, 153-164.

Florio L., Mussone L. (1996b) Freeway section capacity for different meteorological conditions using neural networks. In **Proceedings of the 7th WCTR**, eds. D. Hensher, J. King, and T.Oum, Vol. 2, pp.3-18, Elsevier Science.

Florio L., Mussone L. (1997) Lane occupation models in motorway sections before a toll-gate station., Chania, Greece, 16-18 June, 1997, **Proceedings 8th IFAC**, pp. 1431-1477.

Gipps P.G. (1986) A model for the structure of lane-changing decisions. **Transpn. Res.**, Vol. 20B, **5**, 403-414.

Girosi F., Poggio T. (1991) Networks for learning. In: **Neural networks: concepts, applications and implementation** (Antognetti P., Milutinović V., Ed.s), 110-154, Prentice Hall, Inc., 1991.

Hato E., Taniguchi M., Sugie Y. (1996) Influence of traffic information on drivers' route choice. In **Proceedings of the 7th WCTR**, eds. D. Hensher, J. King, and T.Oum, Vol. 1, pp.27-40, Elsevier Science.

Hornik K., Stinchombe M., White H. (1990) Universal approximation of an unknown mapping and its derivative using multilayer feedforward networks. **Neural Networks 3**, 551-560.

Hornik K. (1991) Approximation capabilities of multilayer feedforward networks. **Neural Networks 4**, 251-257.

Klijnhout J.J. (1984) Motorway control and siganlling: the test of time. **Traffic Engng. Control**, **25**, 193-197.

Leshno M., Lin V. Ya., Pinkus A., Schocken S. (1993) Multilayer feedforward networks with a non polynomial activation function can approximate any function. **Neural Networks**, **6**, 861-867.

Masters, T. (1993) **Practical Neural Network recipes in C++**. Academic Press, Inc.,1993.

Mussone L. (1995) Neural Networks to evaluate motorway flow relationships in different meteorological and flow conditions. In: **Artificial Intelligence Applications in Engineering**,eds Chianese A., Sansone L., Teti R., Zollo G., pp. 145-154, Liguori Editore, Napoli 1995.

Remeijn H. (1984) **The Dutch Motorway Control and Signalling System**. Rijkswaterstaat, Traffic Engineering Division.

Schmidt G., Stappert K.H.(1991) Some aspects of speed-flow relations on German motorway. In: **Highway Capacity and Level of service**, eds Brannholte U., pp. 329-336, Karlsruhe, July 1991, Balkema.

Wemple E.A., Morris A.M., May A.D. (1991) Freeway capacity and flow relationships. In: **Highway Capacity and Level of service**, ed Brannholte U., pp. 329-336, Karlsruhe, July 1991, Balkema.