Dynamic Predictive Control Algorithm for Hard Shoulder Running Highways

Sofia Samoili, EPFL
Prof. André-Gilles Dumont, EPFL
Prof. Nikolas Geroliminis, EPFL

Conference paper STRC 2015

STRC | **15th Swiss Transport Research Conference**
Monte Verità / Ascona, April 15-17, 2015

Dynamic Predictive Control Algorithm for Hard Shoulder Running Highways

Sofia Samoili
Laboratory of Traffic Facilities (LAVOC),
Urban Transport Systems Laboratory (LUTS)
École Polytechnique Fédérale de Lausanne (EPFL)
GC C1 383, Station 18, 1015, Lausanne
Phone: +41 21 693 06 02
Fax: +41 21 693 63 49
sofia.samoili@epfl.ch

Prof. André-Gilles Dumont
Laboratory of Traffic Facilities (LAVOC)
École Polytechnique Fédérale de Lausanne (EPFL)
GC C1 398, Station 18, 1015, Lausanne
Phone: +41 21 693 23 89
Fax: +41 21 693 63 49
andre-gilles.dumont@epfl.ch

Prof. Nikolas Geroliminis
Urban Transport Systems Laboratory (LUTS)
École Polytechnique Fédérale de Lausanne (EPFL)
GC C2 389, Station 18, 1015, Lausanne
Phone: +41 21 693 24 81
nikolas.geroliminis@epfl.ch

April 2015

Abstract

Recurrent congestion, traffic intra-class variability, adaptability of drivers' behaviour to ITS policies and networks' complexity in highways raise challenges in traffic management. Abrupt fluctuations ensued by transitional traffic states advocate the congestion phenomenon, could be anticipated and mitigated by a timely activation of designated control policies. In this scope, the study introduces a dynamic multi-level control algorithm that forecasts lane traffic distribution in succeeding states, examines the inter-dependence of lane vehicle allocation and congestion formation and provides the time frame and thresholds that could be enforced by a control scheme.

In the first level of the top-down approach, separate stochastic clustering procedures are invoked, in order to unveil unbiasedly the spatial patterns that define the prevailing traffic regimes, and the time span during which maximum capacity is attained. In the following level, based on the ensued clusters, multivariate modeling is integrated. Dynamic multivariate generalized regression models are employed to capture patterns of vehicle allocation in lanes, during free flow and congested regimes, and forecast impending traffic behaviour that could proactively trigger the efficient implementation of control policies before the actual expression of the need, thereby moderating delays and costs. The hypothesis regarding the underlying spatial inter-dependence between traffic allocation per lane and traffic states emergence is assessed, through the parameterization of stream dynamics with lane-related spatiotemporal variables. Furthermore, based on the rationale that congestion propagates from left to right in the onset of peak period, indicating a transition to another state, lane density distribution ratio (LDDR) and density of the left lane for congested conditions and of the right lane for uncongested, are addressed as promising determinant response variables. Both parameters are proven site-independent and are intermittently occurring during congestion and free flow conditions. Potential presence of multicollinearity between independent variables is negligible, on account of their normalization and the presence of the bounded response variables.

The algorithm is evaluated through implementation on a recent development in the Intelligent Transportation Systems (ITS) control policy, the hard shoulder running (HSR) system. The use of the hard shoulder as a temporarily regular lane, marginally increases the capacity of segments without additional costly infrastructure works. The case study, a two regular lanes highway, is located between Geneva and Lausanne (A1) and is equipped with a reactive HSR and variable speed limit (VSL) system. Statistical assessment indicates that the prediction models yield significant prediction accuracy.

Keywords

Dynamic spatial modeling – lane distribution – stochastic clustering – speed harmonization – managed lane – hard shoulder use

1. Introduction

Congestion mitigation is polarized between invasive approaches, such as infrastructure interventions, and traffic management through implementation of Intelligent Transportation Systems (ITS), which promote an ameliorated network performance with sustainable economic, spatial and temporal requirements (Brilon et al., 2008; Geistefeldt, 2012). Recent operational control systems, evoked more reliable monitoring and controlling methods to provide a robust prediction of traffic dynamics, with forecasting models that were deployed based on standard traffic spatiotemporal parameters, ensuring significant accuracy.

However, highway networks' complexity, traffic intra-class variability and adaptability of drivers' behaviour to current ITS policies, elicit abrupt fluctuations, delineated by transitional traffic regimes, that fortify congestion and challenge the management of control strategies. In this scope, a novel dynamic multilevel control algorithm with lane-scale parameterisation is proposed. The stochastic method that is introduced in this study, forecasts the succeeding traffic regimes, through the lane density distribution ratio (LDDR), based on patterns of lane stream dynamics, unveiling the inter-dependence of lane vehicle allocation and congestion formation.

As such, the developed control algorithm for highways, consists of three levels. The aim of the first level is the incorporation of an unbiased definition of traffic regimes and peak periods and is addressed with data mining through an artificial neural network (ANN) algorithm that identifies three predominant regimes, separates peak periods from off-peak periods and denoises the dataset.

In the following level, in order to predict impeding stream dynamics, separate multi-regime models are formed for each of the resulted homogeneous clusters. Two types of multivariate generalized regression models, static and dynamic, are deployed for congested or saturated regimes and for free flow regimes, in order to pursue simple and feasible, though integrable in real-time control policies models. Based on the rationale that congestion propagates from right to left in the onset of peak period, indicating a transition to another regime, left lane density distribution ratio (LDDR) and left lane density are addressed as promising determinant response variables for congested conditions. Density-related parameters are selected as they appear to be site-independent and because of density's monotonicity that could accord an unambiguous conclusion regarding the prevailing traffic regime.

The hypothesis regarding the link between lane distribution transitions and the onset of ascending or descending passages from saturated towards congested or free flow regimes, is evaluated at the lower level through a statistical assessment of the control algorithm,

implemented in a managed lane system site. The study concludes with the interpretation of the results and the perspectives of the current analysis.

1.1   State of Research

Traffic dynamics have been extensively analyzed through the main macroscopic parameters of flow, density and speed and traffic fundamental diagrams. Furthermore, as the geometrical attributes and the management strategies vary (number of lanes, ramps, existence of VLS, managed lanes etc.), lane-oriented behavior study could be suggested to delineate the trends (Daganzo, 2002; Chung and Cassidy, 2004; Duret et al., 2012). The impact of lane distribution at traffic flow near merging zones was observed during several traffic conditions, and patterns independent to the study area geometry and control policy were identified during free-flow regimes (Amin and Banks, 2005; Duret et al., 2012). A lane behavioral model that was established by Daganzo, classifies the drivers into aggressive and less aggressive and denotes that up to congested regimes fast drivers are reluctant to proceed to lane change (Daganzo, 2002). However, when the difference of speed becomes marginal, they are dispersing in an attempt to maximize their speed. Based in similar assumptions, a macroscopic theory of vehicle lane-changing in microscopic models was proposed to describe relaxation phenomena (Laval and Leclercq, 2008). Nevertheless, empirical researches acknowledged several regimes and transitional conditions whose traffic patterns could not be sufficiently reproduced by the fundamental diagram, as a high scatter was emerging (Helbing et al., 2009; Knoop et al., 2011; Duret et al., 2012).

Data-driven models (Zhang and Rice, 2003; Antoniou and Koutsopoulos, 2006) dominated the most recent researches, combined in several approaches with artificial neural network-based models (ANN) (Van Lint et al., 2005; Vlahogianni et al., 2008) and others, which were proved more accurate in capturing stream dynamics compared to the traditional statistical methods, since the optimal network parameters emerge, and the extreme traffic conditions, the rapid fluctuations and the transitions between states are better identified and predicted (Smith and Demetsky, 1997; Ishak and Alecsandru, 2003; Stathopoulos and Karlaftis, 2003; Zhang and Rice, 2003; Van Lint et al., 2005; Antoniou and Koutsopoulos, 2006; Vlahogianni et al., 2008). For real-time and near real-time forecasting implementations especially in highways, these are the key features that are required to be comprised for a representative modelling (Jiang and Adeli, 2004; Van Lint et al., 2005). In this framework, wavelets are also enlisted as methods with promising results (Karim and Adeli, 2002). However, based on the objectives of the current study, data-driven approaches are addressing better the subject in terms of traffic pattern recognition (Antoniou and Koutsopoulos, 2006). It is also to be noted that researches on traffic behaviour modeling on near-capacity or congested traffic conditions are better described when are introducing lane-oriented parameters (Daganzo, 2002; Chung and

Cassidy, 2004; Duret et al., 2012; Samoili et al., 2013) than adopting the kinematic-wave (KW) model (Lighthill and Whitham, 1955; Richards, 1956).

2. Methodology and Data

The methodological framework of the multi-level algorithm is presented as follows (Fig.1).
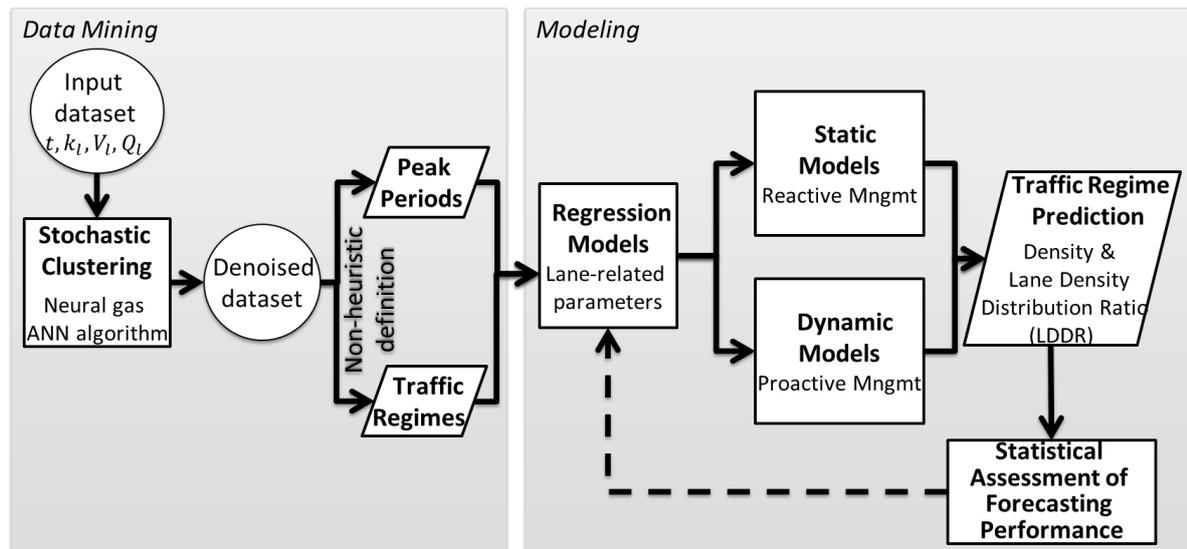
Figure 1. Methodological framework of stochastic lane distribution modeling and dynamic management of ITS policies

2.1 Case Study Setup and Data Description

The study site is a single managed lane per direction in a two-lane Swiss highway between Geneva and Lausanne (A1), separated from the general purpose lanes with pavement signalization. The facility was implemented in 2010, following a continuous increase in traffic demand and recurrent congestion during peak hours (annual average daily traffic of working days in 2008: 88,500 veh/day in both directions). The hard shoulder was broadened and the general purpose lanes were reduced, resulting in 3.50 m of width for every lane. For safety levels maintenance reasons, the system is additionally equipped with a partially automated variable speed limit (VSL) policy. During the operation of the emergency lane as temporary additional lane, a speed limit of 100km/h is enforced, as opposed to the 120 km/h during non-operation. Even though the system was designed to operate based on certain speed and density thresholds, currently it serves as decision support tool for the traffic control center operators, who open eventually the emergency lane.

The traffic dataset (courtesy of SMETRA) is derived from four radar sensors before and in the beginning of the system in both directions (L-59060, L-60590, J-64900, J-63145) (Fig. 2). These sensors measure the traffic volume per lane, speed, percentage of heavy vehicles and

indications of the opening and closure of the shoulder lane. The study period comprises four months in 2013 and 2014 (March, May of 2013 and 2014), of which holidays and days that accidents occurred are excluded. The data are aggregated per Δt=3min intervals, in order to neither cover nor overly pronounce any traffic dynamics characteristics.
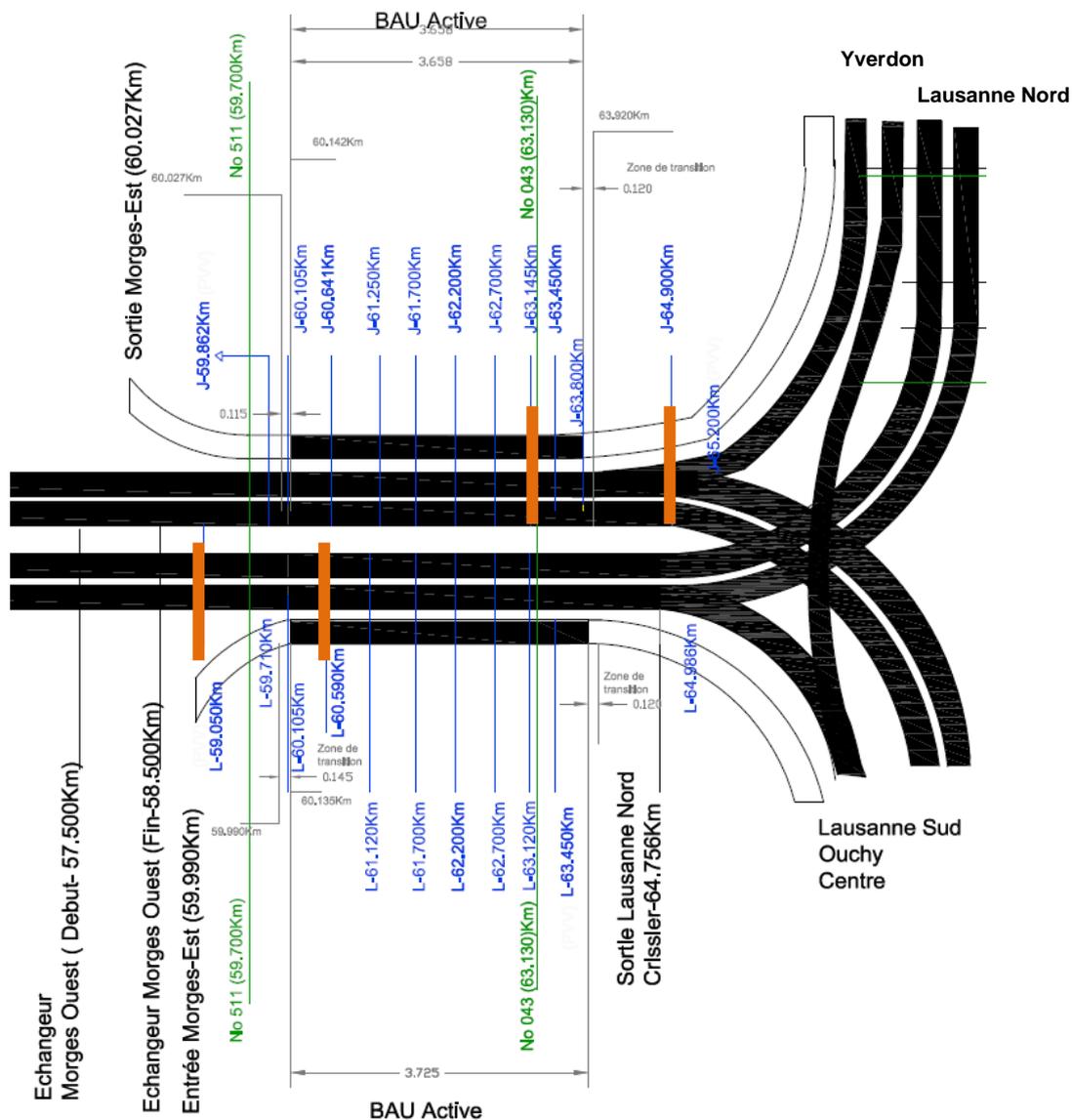
Figure 2 Case study site. The detectors that are used in the modeling section are depicted in orange (L-59060, L-60590, J-64900, J-63145).

2.2   Clustering Analysis

Independent stochastic clustering procedures are invoked for the unbiased definition of a) the three prevailing traffic regimes that capture adequately stream dynamics in three homogeneous groups (free flow, saturated, congested) through spatial patterns, b) time span, during which maximum capacity is attained per sections of a network, hereinafter referred to

as the peak period, through temporal patterns. Additionally, any detected outliers are removed, composing a denoised and concise input space.

A neural network algorithm, namely the "neural-gas" algorithm is employed for each clustering (Martinetz et al., 1993). The algorithm is a fuzzy extension of the k-means clustering, setting the neighbourhood of each data vector, v, based on proximity ranking with weight vectors, also referred to as cluster centers, $c_j \in \mathcal{R}^D$ with $j = 1, ..., N$ the number of observations, and $\mathcal{R}^D$ the topological manifold of real coordinate D-space. The studied data vectors are: (a) a six-dimensional vector of normalized densities per lane and normalized speeds per lane, $v_{k_{n_l}v_{n_l}}$, and a six-dimensional vector of normalized densities per lane and lane density distribution ratios (LDDR) per lane, $v_{k_{n_l}r_{k_l}}$, for the identification of traffic regimes, and (b) a two-dimensional vector of time and densities per direction, $v_{tk_{dir}}$, for the separation of peak from off-peak periods. Hence, three separate clustering procedures are invoked for each of the three vectors, $v_{k_{n_l}v_{n_l}}$, $v_{k_{n_l}r_{k_l}}$, $v_{tk_{dir}}$. For every reference to v hereinafter, each of these vectors is implied, depending on the case.

The convergence with the "neural-gas" to low distortion errors $d(v, c_{i(v)})$ is proven faster than the standard k-means, as each v is represented by the most suitable reference vector $c_{i(v)}$ of the submanifold sb, or else the input dataset of empirical observations for four months selected over the entire years 2013 and 2014, for which the $d(v, c_{i(v)})$ (e.g. the squared error $\left\| v - c_{i(v)} \right\|^2$) is minimal. This divides sb into Voronoi polygons subregions, out of which each v is described by the corresponding reference c_i. The average distortion error lies on the error surface E that is given by eq.1:

$$E = \int d^D \, vP(v)\left(v - c_{i(v)}\right)^2 \tag{1}$$

where the data point distribution $P(v)$ is a stochastic sequence of sample data points $v(t = 1), v(t = 2), ...$ that determines the adaptation steps for the cluster centers. Since E has several local minima instead of the distance $\left\| v - c_{i(v)} \right\|$, a "soft-max" adaptation rule is applied that sets the cluster centers c_i based on a neighborhood-ranking $(c_{i_0}, c_{i_1}, ..., c_{i_{N-1}})$ for the given data vector v, with c_{i_0} being closest to v, then c_{i_1} being second closest to v and so forth, and an adaptation step for adjusting the c_i given by eq.2:

$$\Delta c_i = \epsilon \cdot e^{-\frac{\delta_i(v,c)}{\lambda}} \cdot (v - c_i) \tag{2}$$

that is in fact a deterministic annealing procedure that starting from a high temperature and decaying gradually to 0 with each adaptation step Δc_i increasing δ_i with a characteristic decay constant λ, results to a slow emergence of the local minima of E, avoiding thus that the set of

c is being captured in suboptimal states. For a step size $\epsilon \in [0,1]$ and $\lambda \to 0$, the Δc_i (eq. 2) becomes equivalent to the k-means adaptation rule, but for $\lambda \neq 0$ all cluster centers c_i of the neighborhood ranking are updated and not only the c_{i_0} that sets the center for the cluster in question. Consequently, the dynamics of c_i follows the stochastic gradient descent on the cost function that from eq. 1 becomes eq. 3, and is the average distortion error that has to be minimized:

$$E_{ng}(c,\lambda) = \frac{1}{2C(\lambda)} \sum_{i=1}^{N} \int d^D \ vP(v) h_\lambda(\delta_i(v,c))(v - c_i)^2 \qquad (3)$$

where $\varphi(\lambda) = \sum_{i=1}^{N} h_\lambda(k_i) = \sum_{\delta=0}^{N-1} h_\lambda(\delta)$ is the normalization factor that only depends on λ. This sets a fuzzy clustering approach where v is assigned to a cluster i with a fuzzy membership function $p_i(v) = \frac{h_\lambda(\delta_i(v,c)}{\varphi(\lambda)}$ and ranges from 0, namely not belonging to cluster, to 1, belonging to cluster. Consequently, through the annealing procedure, each data point is assigned to a cluster based on the fuzzy membership function, denoted by the minimization of the average distortion error of every Voronoi polygon subregion, and hence assembling an error surface of several local minima.

Therefore, neural-gas approach outperforms standard clustering techniques as the k-means, the maximum-entropy and Kohonen's feature map algorithm on a number of separated data clusters, since they might tend to converge to local minima for non-smooth data, as in the case of rapidly fluctuating traffic flow, while the algorithm reduces convergence time to reach smaller distortion errors than the aforementioned procedures. Moreover, it requires an order of magnitude fewer weights to achieve the same prediction error.

2.3   Regression Analysis

Following the issued formation of homogeneous clusters and the compression and denoising of data space by the elimination of outliers, dynamic multivariate generalized regression models are employed in order to depict patterns of vehicle allocation in lanes during free flow and congested regimes and forecast impending traffic behaviour that could proactively trigger the efficient implementation of control policies. The aim is to predict and hence alleviate part of the causes of the mechanisms that trigger the hysteresis phenomenon, namely spatiotemporal distributions, thereby moderating delays and costs.

The hypothesis regarding the underlying spatial inter-dependence between traffic allocation per lane and traffic states emergence is assessed, through the parameterization of stream dynamics with lane-related spatiotemporal variables. Based on the rationale that congestion propagates from left to right in the onset of peak period, indicating a transition to another regime, lane density distribution ratio (LDDR) and density of the right or left lane depending

on the regime, are addressed as promising determinant response variables. They are selected representative indicators, for the monotonic properties of density that accords an unambiguous conclusion regarding definition of congested or free flow conditions. Potential presence of multicollinearity within the independent variables is inquired, even though their normalisation and the account of bounded response variables, renders it negligible.

Pursuing initially a simple approach that would ensure feasibility of implementation, static models of type $Y(t) = \sum(\alpha + \beta_i X_i(t))$ with α, β constant coefficients $\in \mathbb{R}\backslash\{0\}$ are formed with limited number of variables. Nevertheless, in order to address realistically the potential application of the models to a real-time control policy, dynamic models of type $Y(t + 1) - Y(t) = \sum(\alpha + \beta_i X_i(t))$ are separately developed for each cluster of the uncongested and congested regimes. Their dynamic character lays on the fact that they are providing predictions for an interval, *t+1*, one time step subsequent to the current time *t*.

The forecasting parameters are subject to unity-based normalization, in order to be scale invariant for comparability reasons. The restriction of the range of values in the dataset of observations x_k between two arbitrary points, thus $[a, b]$ or in this case $[0,1]$, is achieved with the following equation (eq. 4):

$$x_{k_n} = a + \frac{(x_k - x_{\min})(b-a)}{x_{\max} - x_{\min}} = \frac{(x_k - x_{\min})}{x_{\max} - x_{\min}} \tag{4}$$

2.4 Statistical Assessment

The attainment of prediction accuracy of the developed regression models is assessed through the residual standard error $\hat{\sigma}$ that is an estimate of the standard deviation σ of the sample of the given dataset, which sets the magnitude of difference of each observation of the sample from the sample mean. As the least squares regression approach is used, the residuals $\hat{\varepsilon}_i$, namely the estimated errors that occur from the differences between the sample of empirical observations Y_i and the predicted $\hat{Y}_i$ (eq. 5), are described by a Gaussian (normal) distribution with properties of variables with mean 0 and standard deviation σ.

$$\hat{\varepsilon}_i = Y_i - \hat{Y}_i = Y_i - (\hat{a} + \hat{\beta} X_i) \tag{1}$$

Thereby, the residual standard error $\hat{\sigma}$ is computed as follows (eq. 6):

$$\hat{\sigma} = \sqrt{\frac{1}{n-p-1} \sum(\hat{\varepsilon}_i)^2} \tag{2}$$

where $n - p$ are the statistical degrees of freedom. The assumption of normality of the residuals has been confirmed by appropriate Q-Q plots.

To evaluate the goodness of fit of the regression, a version of the coefficient of determination R^2, the adjusted R-squared is employed. The $R^2 adj.$ penalizes every dependent variable that is added to model without contribution to the explanation of the regression, with 1 signalizing the best fit to the model.

3. Analysis

3.1 Data Mining for Traffic Patterns Identification & Optimal Activation Range

In the first level of the introduced control algorithm, peak periods (morning, evening) and three prevailing traffic regimes (free flow, saturated, congested) are derived from a non-heuristic method, the clustering approach with the "neural-gas" algorithm. Furthermore, outliers are removed, resulting to a denoised input space. The traffic dynamics of each regime are emerged and are presented in Figures 3 and 4. Each cluster comprises indications about stream patterns that contribute to the targeted delineation of traffic dynamics, since separate models are developed per regime that can be applied to activate policies according to the traffic behavioural patterns of each cluster.

The morning and evening peak periods serve to reduce the research area and target the optimal thresholds of a traffic responsive activation. Both periods are defined by a two-dimensional input vector of time and densities per direction that was processed by the "neural-gas". Since density is monotonic, an association to another parameter so as to determine a traffic regime, is not required as it is the case of the dual character of flow. In Fig. 3, time is plotted against density per direction aggregated per time intervals of Δt=3-min. The ensued peak periods of morning range between 6:45 a.m. and 9:30 a.m (6:45 a.m.-9:15 a.m. for the representative day 18.03.2014, in Fig. 3a), and of the respective evening between 15:00 and 20:00 p.m. (15:21 p.m.-19:18 p.m., in Fig. 3b). The peak periods are denoted by the sign "*", and the off-peak periods before the onset or after the offset of the peak period by the sign "Δ", or by the sign "○" between two peak periods. The operation of the system, as it is signalised by the current threshold, is illustrated by a continuous line, whereas the operation as it was effectuated by the operation center for the presented day, with a dashed lane. The green line depicts the activation and the red line the deactivation of the managed lane system.

The aforementioned signs correspond to the three homogeneous clusters of free flow, saturated and congested regimes (cluster 1: "Δ", 2: "○", 3: "*" respectively), which are defined based on the six-dimensional vector of the per lane density and the lane density distribution ratio (LDDR) (Fig. 4). The lane density distribution, namely ratio of the

density of a lane divided by the total density of the direction, can be described as follows through the relationship that relates the ratios of each lane (eq. 7):

$$\sum_{l=1}^{l=3}(\beta_l \cdot r_{k_l} + \alpha) = \sum_{l=1}^{l=3}\left(\beta_l \cdot \frac{k_l}{k_{dir}} + \alpha\right) = 1 \tag{3}$$

for $\quad k_{dir} = \sum_l k_l \in \mathbb{R} \setminus \{0\}$

where l is denoting the lanes ($l=1$ is the left lane), r_{k_l} are the ratios of lane density distribution (LDDR), hereinafter expressed in percentage, over the total density per direction k_{dir}, computed by the per lane density k_l, and constrained by 0 (as it is appointed as the divisor of the k_l,), and β_l, α the coefficients that are estimated by locally weighted regression (Cleveland and Devlin, 1988).

The relationships of per lane density to LDDR aim to provide an insight into traffic behaviour and traffic distribution dynamics in lanes, which can unveil the transitions between the regimes and highlight the thresholds for a timely activation of an ITS. In these transitions that occur between free flow and congested regimes, represented by the saturated or otherwise known as synchronised regimes, lay the range of thresholds to be inquired for a traffic responsive system operation, since they precede or succeed the periods of rapid fluctuations of traffic demand in highways, which require the system's activation or deactivation. Cluster 2, is called synchronised regime as it represents the transitions, or else fluctuations see passages, that comply with the free-flow or congested regimes upstream or downstream, towards a regime different than the current. The saturated regime may only result from a) an ascending passage from the free flow regime (1→2), b) a descending passage from the congested regime (3→2), or from c) an interrupted passage by the congested regime (2→3→2). The congested regime may only occur either by a) an ascending passage from the saturated regime (1→2→3), or by b) an interrupted passage by a saturated regime (3→2→3). In the contrary, the free flow regime can only occur from a descending passage from the saturated regime (2→1), hence the passages 1↛2↛1 and 3↛1 cannot happen.

More explicitly, if a series of observations is attributed to the saturated cluster (cluster 2), and there is at least one observation $k(t_i)$ attributed to cluster 3 (2→3), then the transitions period and thus the saturated regime is over, and the following observations, $k(t_i + t)$,will be classified to cluster 3, since the distortion error has been already used to compare the $k(t_i)$ to the minima of every vector of the sectors of the managed lane section both upstream and downstream, resulting to permitting the classification, until the time t_x that an observation $k(t_x)$ maximises the error. While a saturated cluster succeeds a free flow cluster (1→2), only a congested cluster (3) may follow (1→2→3). Therefore, when this condition is satisfied (1→2→3), the respective values of density of this interval, which begins with the first observation attributed to cluster 2 and ends with the first observation of 3, are proposed to be

studied as candidate thresholds for the activation of the hard shoulder. Likewise, the candidate range for the inquiry of the optimal evening activation, can be located among the last observations of the saturated regime and the first of the congested, when the condition of passage 3→2→3 is satisfied. Regarding the system's deactivation, when a congested regime cluster is preceded and followed by a saturated regime (2→3→2), the respective density values of the first observations classified to the second appearance of cluster 2, form the candidate set of thresholds to examine. This clustering approach has the advantage that an observation is classified to a cluster, only when the result of the comparison of each of the distorting errors of the observations of the surface between them, is minimal. This explains the fact that there is no recurrence when after the emergence of a cluster. It is noted that this surface is created by the vectors of the *t-1* observations of all the detectors of the study area, and the error is recalculated and updated for each added observation.

Based on the LDDR and density per lane relationships, the hypothesis regarding the lane vehicle allocation and the saturated regime is confirmed. In a range of normalised density of left lane k_{n_L} [0.03,0.08] veh/km/lane that depicts the synchronised cluster (2), the maximum value, $k_{n_L} \cong 0.08$ veh/km/lane, indicates the onset of congestion. In addition, a underutilisation of the left lane is observed, as the emergency lane system is activated when the left lane is occupied only by a $r_{k_L} \cong 40\%$ for $k_{n_L} \cong 0.08$ (Fig. 4a), when the $k_{n_R} \cong 0.10$ and heads towards the onset of congestion (Fig. 4b), uncovering the existence of margins for amelioration of the operation. The uneven distribution is further verified also in free flow regimes that the right lane is favored over the left, as for $k_{n_R} \leq 0.06$ veh/km/lane the right lane's LDDR, r_{k_R}, ranges from 65% to 95% (Fig. 4b), hence it is preferred over the left lane that presents a r_{k_L} [15%,35%] (Fig. 4a). Only during congested regimes, $k_{n_L} > 0.08$ veh/km/h, an equilibrium between left and right LDDR is noted.
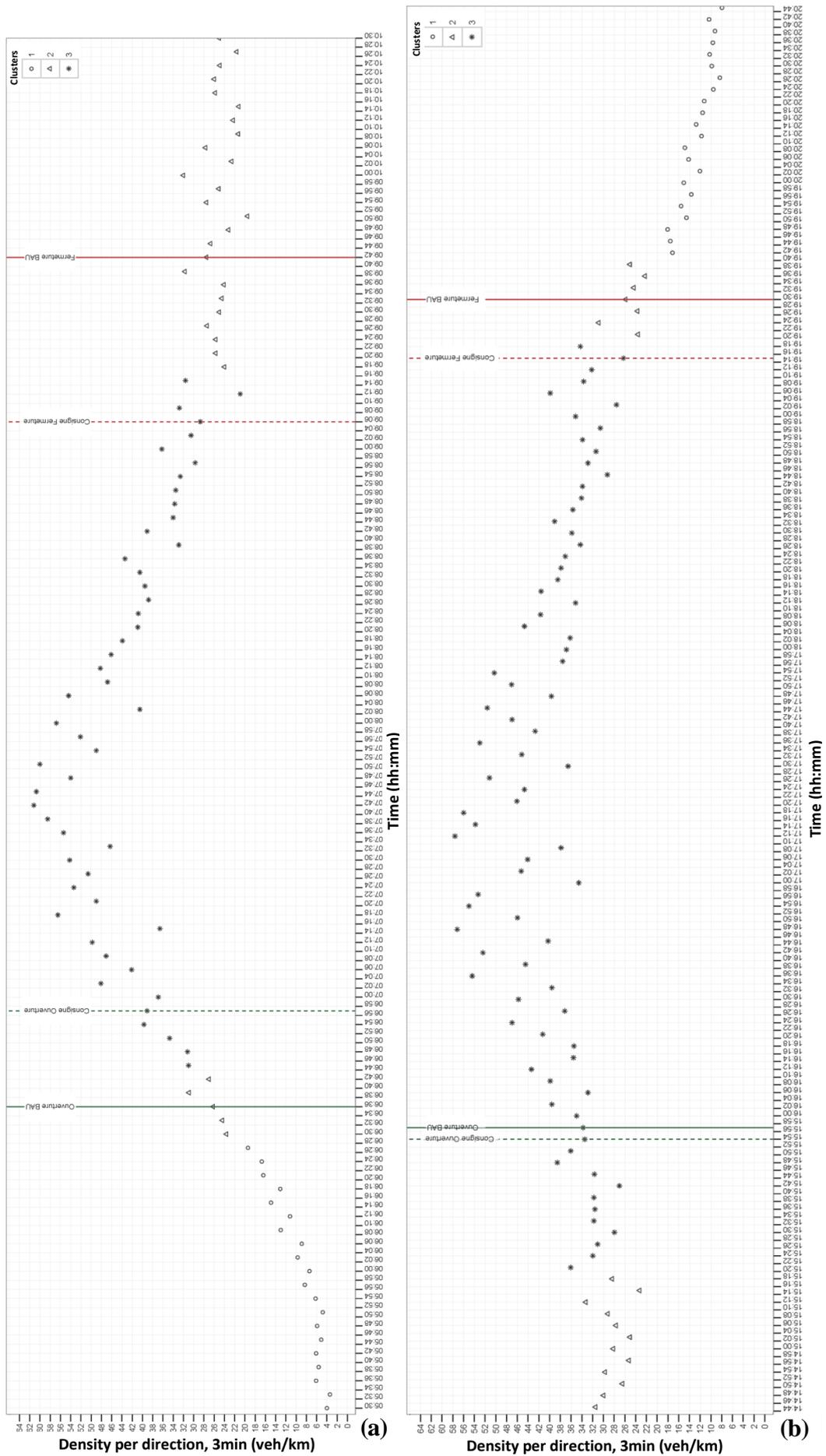
Figure 3 Clusters of peak periods of (a) morning and (b) evening (L-60590, 18.03.2014)
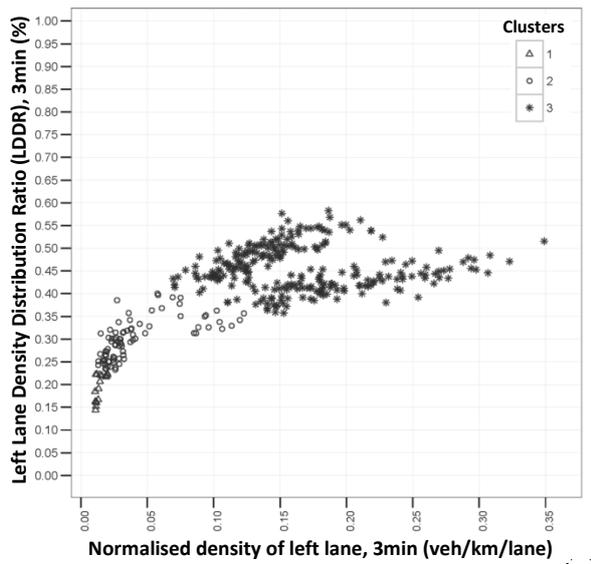
(a)
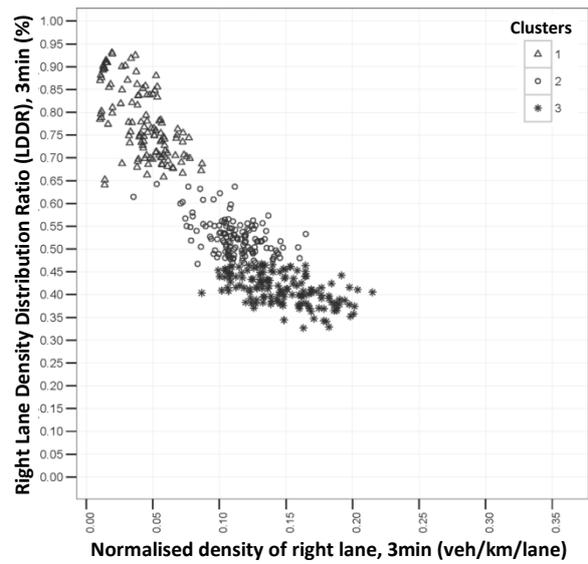
(b)
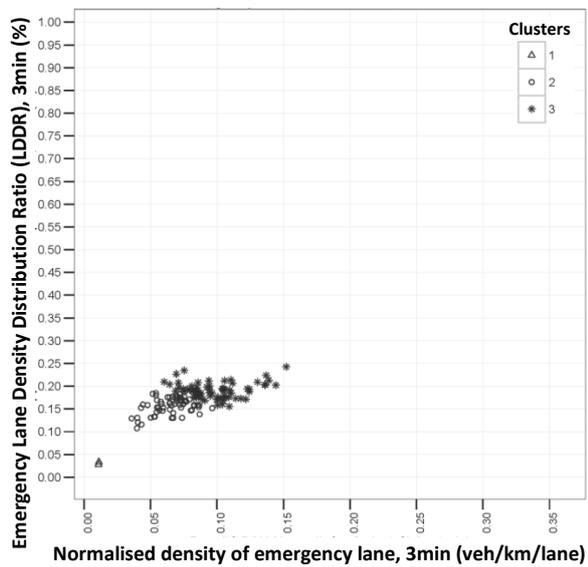
(c)

Figure 4 Clusters of traffic regimes per lane (a) left, (b) right, (c) emergency lane (L-60590, 18.03.2014)

3.2   Lane Distribution Modeling for Managed Lanes Management

In the level in question, the aim is to predict and hence alleviate part of the causes of the mechanisms that cause extended congested phenomena, through impending spatiotemporal distributions that could be integrated to existing control algorithms of ITS schemes, moderating thereby delays and costs. The hypothesis regarding the underlying spatial inter-dependence between traffic allocation per lane and traffic states emergence is assessed, through the parameterization of stream dynamics with lane-related spatiotemporal variables. Separate models are formed for each of the resulted homogeneous clusters, enhancing thus their statistical power and contributing to the acquired significant accuracy for the modeling of both congested and uncongested conditions.

The data mining of the preceding level of the algorithm, revealed that the left lane related LDDR and density could serve as indicative predictors during congested conditions, as congestion propagates in this site from right to left, and thus any alteration could suggest transition to/from free flow or saturated regimes. In the same scope, the LDDR of the right lane and the density could be evoked as representative response variables during uncongested regimes. Density was selected as an appropriate response variable as it is site-independent and its monotonic property accords an unambiguous conclusion regarding the potential association of lane density distribution and traffic regimes.

Two types of multivariate generalized regression models, static and dynamic, are deployed for congested or saturated regimes and for free flow regimes, in order to pursue simple and feasible models for reactive management schemes, and respectively integrable models in real-time proactive control policies, as a result of their rapid computation time, that will eventually improve the policies performance. Although multicollinearity within independent variables is negligible, since they are normalised and on account of the bounded response variables, it is inquired based on the significance of the estimated regression coefficients.

To assess the impact of the implementation of the developed models in terms of network performance, an exploratory analysis is effectuated at the aforementioned reactive managed lane and variable speed limit (VSL) system of a segment of a Swiss highway. In order to provide a timely operation of the Hard Shoulder Running (HSR) control policy system, forthcoming stream dynamics are monitored in two sections per direction; upstream (L-59060, J-64900) and in the HSR system (L-60590, J-63145), ensuring sufficient time interval for its activation before the propagation of any upstream triggering conditions to downstream. In the following subsections, the proposed models are given for one direction, denoted as "up" for the upstream section L-59060, and "dwn" for the downstream section L-60590.

3.2.1 Reactive Management – Static Models

Initially, simple static models were established with limited number of variables, introducing though LDDR as response variable. Nevertheless, statistical assessment demonstrated that precision and principally fit could be improved. This is remedied by raising the parameterization complexity with adding parameters that describe better the forthcoming motion, granting though a significant improvement in fit to data. An adaptation of a certain threshold of a control algorithm in line with the developed static models could be proved efficient.

Two separate models for congested and uncongested regimes are formed, based on the previously stated rationale that left lane is associated to congested conditions and right lane to free flow regimes. For the congested conditions model, given the congested and saturated clusters, the left lane LDDR upstream of the system, $r_{k_L}^{up}$, is set as response variable, based on the rationale that traffic conditions can be described by observing only the regime in the acceleration (left) lane that merge to the leftmost lane to prevent the anticipated incoming traffic to the system downstream. Thereby, a lower LDDR of left lane implies that the rightmost lanes are in a regime of similar or higher magnitude of density. The explanatory variables, refer to the current time t, and are the normalized density of the left lane downstream, $k_{n_L}^{dwn}(t)$, the lane flow distribution ratio (LFDR) of the right lane downstream, $r_{k_R}^{dwn}(t)$, the derivative of the normalised left lane density upstream, $\frac{dk_{n_L}}{dt}^{up}(t)$, and the derivative of the normalised left lane speed upstream, $\frac{dV_{n_L}}{dt}^{up}(t)$. The function that describes the model is the following (eq.7)

Multivariate static regression model for congested conditions:

$$Y_{r_{k_L}^{up}}(t) = \beta_{k_{n_L}^{dwn}(t)} \cdot X_{k_{n_L}^{dwn}}(t) + \beta_{r_{Q_R}^{dwn}(t)} \cdot X_{r_{Q_R}^{dwn}}(t) + \beta_{\frac{dk_{n_L}}{dt}^{up}(t)} \cdot X_{\frac{dk_{n_L}}{dt}^{up}}(t) + \beta_{\frac{dV_{n_L}}{dt}^{up}(t)} \cdot X_{\frac{dV_{n_L}}{dt}^{up}}(t)$$

$$(7)$$

The statistical results presented in Table 1, are based on n=4293 observations and indicate that the predicted left lane LDDR (LLDDR) upstream is inversely related to the right lane LFDR downstream, the derivative of left lane density upstream and the derivative of the left lane speed upstream. This implies that greater LLDDR upstream indicates the onset of congestion, and entails slow attainment of lower densities and speeds in the left lane, since it is highly occupied. It also denotes lower right lane LFDR downstream, which is expected as the right lane is not preferred during congestion or when an ascending passage towards congestion occurs. Consequently, the hypothesis regarding the underlying spatial inter-dependence between traffic allocation per lane and traffic states emergence is justified.

Table 1 Static multivariate models for congested and uncongested conditions, as resulted by clustering.

	Static		
	Congested		**Uncongested**
Variables*	**β**	**Variables***	**β**
$Y_{r_{k_L}^{up}}(t)\sim$	(n=4293)	$Y_{r_{k_R}^{up}}(t)\sim$	(n=3765)
Intercept	0.81	Intercept	0.43
$\beta_{k_{n_L}^{dwn}(t)}$	0.30	$\beta_{r_{k_R}^{dwn}(t)}$	0.19
$\beta_{r_{Q_R}^{dwn}(t)}$	- 0.77	$\beta_{k_{n_R}^{up}(t)}$	2.02
$\beta_{\frac{dk_{n_L}}{dt}^{up}(t)}$	- 44.00	$\beta_{k_{n_R}^{dwn}(t)}$	- 0.37
$\beta_{\frac{dV_{n_L}}{dt}^{up}(t)}$	- 11.41	$\beta_{k_{n_L}^{up}(t)}$	- 1.96
		$\beta_{k_{n_L}^{dwn}(t)}$	0.11
		$\beta_{\Delta V_{n_{RL}}^{up}(t)}$	0.05
		$\beta_{\Delta V_{n_{RL}}^{dwn}(t)}$	- 0.07
Residual S.E.	0.05	**Residual S.E.**	0.01
R^2 adjusted	0.84	**R^2 adjusted**	0.92

* All variables are statistically significant at the 99.9% confidence level, based on t-test.

The developed static model for uncongested conditions (eq. 8), consists of the explanatory variables, which all refer to the current time t, of right lane LDDR downstream, $r_{k_R}^{dwn}(t)$, the right lane normalised density upstream, $k_{n_R}^{up}(t)$, the right lane normalised density downstream, $k_{n_R}^{dwn}(t)$, the left lane normalised density upstream, $k_{n_L}^{up}(t)$, the left lane normalised density downstream, $k_{n_L}^{dwn}(t)$, the normalised difference of right and left lane speed upstream, $\Delta V_{n_{RL}}^{up}(t)$, and the normalised difference of right and left lane speed downstream, $\Delta V_{n_{RL}}^{dwn}(t)$. The right lane LDDR upstream, $r_{k_R}^{up}(t)$, is assigned as response variable, based on the justified hypothesis that the right lane is preferred during any other regime but the congested.

Multivariate static regression model for uncongested conditions:

$$Y_{r_{k_R}^{up}}(t) = \beta_{r_{k_R}^{dwn}(t)} \cdot X_{r_{k_R}^{dwn}}(t) + \beta_{k_{n_R}^{up}(t)} \cdot X_{k_R^{up}}(t) + \beta_{k_{n_R}^{dwn}(t)} \cdot X_{k_R^{dwn}}(t) + \beta_{k_{n_L}^{up}(t)} \cdot X_{k_L^{up}}(t) +$$
$$\beta_{k_{n_L}^{dwn}(t)} \cdot X_{k_L^{dwn}}(t) + \beta_{\Delta V_{n_{RL}}^{up}(t)} \cdot X_{\Delta V_{n_{RL}}^{up}}(t) + \beta_{\Delta V_{n_{RL}}^{dwn}(t)} \cdot X_{\Delta V_{n_{RL}}^{dwn}}(t)$$

$$(8)$$

Indeed, the LDDR of the right lane upstream is inversely related to the right lane normalised density downstream, the left lane normalised density upstream and the normalised difference of right and left lane speed downstream. This confirms that an increase to the right lane

_____

LDDR downstream results to a decrease of the left lane density and hence depicts a transition towards uncongested conditions, which triggers the decrease of the right lane density upstream and the attainment of similar speeds between the right and left lane, since low densities concede circulation in maximum allowed speeds and distribution to any lane with acceptable time or distance gap (Table 1).

Based on the statistical assessment, both static approaches model adequately traffic behaviour. The static model for uncongested conditions yielded borderline higher accuracy (1% vs. 5%) and notably better fitting than the model for congested (92% vs. 84%), as uncongested patterns are more repetitive and thus more predictable for modeling.

3.2.2   Proactive Management – Dynamic Models

For the operation of real-time control systems, dynamic models are formed. The dynamic character of the following models for congested and uncongested conditions lays on the fact that they are providing predictions for an interval, t+1, one time step subsequent to the current time t. The prediction horizon is set on t=3-minutes, commensurating with the aggregation interval, which is considered as adequate to detect essential alterations, without overleaping or accentuating rapid fluctuations that may signalize the impendence of a successive regime. For scalability purposes the variables are normalized.

A dynamic model for congested conditions (eq. 9) is intended to describe stream dynamics during this regime, through a left lane related variable for the reasons described to the corresponding static model. The model predicts the response variable of left lane LDDR upstream one time step subsequent $r_{k_L}^{up}(t+1)$ to the current, $r_{k_L}^{up}(t)$, with an accuracy of 5% (Table 2), when as explanatory variables are set the current values of normalized density of the left lane downstream, $k_{n_L}^{dwn}(t)$, the right lane LFDR downstream, $r_{Q_R}^{dwn}(t)$, the derivative of the normalised left lane speed upstream, $\frac{dV_{n_L}}{dt}^{up}(t)$, and the discrete normalised number of lanes that are used as general purpose lanes downstream (0 for 2 lanes, 1 for 3 lanes), $nb_{lanes}^{dwn}(t)$, for the normalised variable of time that the managed lane system is open downstream, $T_n^{dwn}(t)$.

Multivariate dynamic regression model for congested conditions:

$$Y_{r_{k_L}^{up}}(t+1) - Y_{r_{k_L}^{up}}(t) = \beta_{k_{n_L}^{dwn}(t)} \cdot X_{k_{n_L}^{dwn}}(t) + \beta_{r_{Q_R}^{dwn}(t)} \cdot X_{r_{Q_R}^{dwn}}(t) + \beta_{\frac{dV_{n_L}}{dt}^{up}(t)} \cdot X_{\frac{dk_{n_L}}{dt}^{up}}(t) +$$
$$+ \beta_{nb_{lanes}^{dwn}(t)} \cdot X_{nb_{lanes}^{dwn}}(t) + \beta_{T_n^{dwn}(t)} \cdot X_{T_n^{dwn}}(t)$$

<div align="right">(9)</div>

Based on n=4293 observations, the LLDDR upstream one time step following to the current is inversely related to the right lane LFDR (RLFDR) downstream and the derivative of the normalised left lane speed upstream, which suggests that a decrease to the latter indicates a subsequent 3-min transition towards denser conditions, thus the decrease of the left lane LDDR upstream is justified as the congestion propagates up to the section in question. The model successfully provided adequate fitting to the data with an adjusted R^2 of 81% (Table 2).

Table 2 Dynamic multivariate models for congested and uncongested conditions, as resulted by clustering.

Dynamic			
Congested		**Uncongested**	
Variables*	**β**	**Variables***	**β**
$Y_{r_{k_L}^{up}}(t+1) - Y_{r_{k_L}^{up}}(t) \sim$	$(n=4293)$	$Y_{k_{n_R}^{up}}(t+1) - Y_{k_{n_R}^{up}}(t) \sim$	$(n=3765)$
Intercept	0.43	Intercept	0.02
$\beta_{k_{n_L}^{dwn}(t)}$	0.30	$\beta_{k_{n_L}^{dwn}(t)}$	0.12
$\beta_{r_{Q_R}^{dwn}(t)}$	- 0.46	$\beta_{k_{n_R}^{dwn}(t)}$	0.31
$\beta_{\frac{dV_{n_L}}{dt}^{up}(t)}$	-19.40	$\beta_{\frac{dV_{n_L}}{dt}^{dwn}(t)}$	- 0.18
$\beta_{nb_{lanes}^{dwn}(t)}$	0.01	$\beta_{Q_{n_L}^{up}(t)}$	- 0.04
$\beta_{T_n^{dwn}(t)}$	0.23	$\beta_{Q_{n_R}^{up}(t)}$	- 0.06
		$\beta_{\frac{dQ_{n_R}}{dt}^{up}(t)}$	0.13
		$\beta_{nb_{lanes}^{dwn}(t)}$	0.004
		$\beta_{T_n^{dwn}(t)}$	0.01
Residual S.E.	0.05	**Residual S.E.**	0.02
R^2 adjusted	0.81	**R^2 adjusted**	0.79

* All variables are statistically significant at the 99.9% confidence level, based on t-test.

Idem, the developed dynamic model for uncongested conditions (eq.10), predicted with a significant low residual standard error (2%), the normalised left lane density upstream one time step subsequent, $k_{n_L}^{up}(t+1)$, to the current, $k_{n_L}^{up}(t)$. The explanatory variables comprise the left lane normalised density downstream, $k_{n_L}^{dwn}(t)$, the right lane normalised density downstream, $k_{n_R}^{dwn}(t)$, the derivative of the normalised left lane speed downstream, $\frac{dV_{n_L}}{dt}^{dwn}(t)$, the normalised left and right lane flow upstream, $Q_{n_L}^{up}(t)$ and $Q_{n_R}^{up}(t)$ respectively, the derivative of the normalised right lane flow upstream, $\frac{dQ_{n_R}}{dt}^{up}(t)$, and the discrete normalised number of lanes that are used as general purpose lanes downstream (0 for 2 lanes, 1 for 3 lanes), $nb_{lanes}^{dwn}(t)$, for the normalised variable of time that the managed lane system is open downstream, $T_n^{dwn}(t)$.

Multivariate dynamic regression model for uncongested conditions:

$$Y_{k_{n_L}^{up}}(t+1) - Y_{k_{n_L}^{up}}(t) = \beta_{k_{n_L}^{dwn}(t)} \cdot X_{k_{n_L}^{dwn}}(t) + \beta_{k_{n_R}^{dwn}(t)} \cdot X_{k_{n_R}^{dwn}}(t) + \beta_{\frac{dV_{n_L}}{dt}^{dwn}(t)} \cdot X_{\frac{dk_{n_L}}{dt}^{dwn}}(t) +$$

$$+ \beta_{Q_{n_L}^{up}(t)} \cdot X_{Q_{n_L}^{up}}(t) + \beta_{Q_{n_R}^{up}(t)} \cdot X_{Q_{n_R}^{up}}(t) + \beta_{\frac{dQ_{n_R}}{dt}^{up}(t)} \cdot X_{\frac{dQ_{n_R}}{dt}^{up}}(t) + \beta_{nb_{lanes}^{dwn}(t)} \cdot X_{nb_{lanes}^{dwn}}(t) +$$

$$\beta_{T_n^{dwn}(t)} \cdot X_{T_n^{dwn}}(t)$$

$$(10)$$

The normalised left lane density upstream one time step subsequent to the current is inversely related to the normalised left and right lane flow upstream and to the derivative of the normalised left lane speed downstream, implying that an increase to the latter variables induces a subsequent 3-min transition towards less dense conditions that signify less attractiveness to the left lane, thus an increase to the right lane density.

4. Conclusions

This research is aiming to provide a stochastic approach that adapts to a designated ITS proactive or reactive system that may be implemented, providing amelioration of designated policies' performance and hence of the impending stream dynamics. As such, a multilevel control algorithm is proposed that forecasts succeeding traffic regimes through lane distribution parameterisation, based on hypothesis of the existence of an inter-dependence of a patternised lane vehicle allocation and traffic regimes.

The first level of the algorithm comprises data mining, in order to incorporate an unbiased definition of traffic regimes and peak periods into the subsequent level of modeling, which generates a concise input dataset and thus promotes accuracy prediction. With the "neural-gas" NN algorithm and independent clustering procedures, three homogeneous groups of the predominant traffic regimes and two of the time span of peak/off-peak periods are derived. The clustering analysis on the relationships between density per direction and lane density distribution ratios revealed an uneven vehicles' distribution, which favours the right lane even when saturation is reached, while left lane remains underutilised up to the onset of congestion conditions. The reason lies potentially on the destination of the drivers'. These lane density patterns that are intermittently occurring during respective traffic conditions, induce the conjecture that congestion moves from right to left in the onset of peak periods and from left to right in the offset, and that traffic inter-lane propagation delineates a potential transition to another state. This justifies the initial hypothesis that certain lane distribution denotes the offset of a current state and the ascending or descending passage to another.

Based on the findings of the first level of the algorithm, the sequence between congested and uncongested conditions can be anticipated through lane-related parameterisation of the stream. Therefore, the second level serves to alleviate part of the causes of the mechanisms that cause extended congested phenomena, through prediction of impending spatiotemporal distributions that could be integrated to existing control algorithms of ITS policies,

moderating delays and costs. In this scope, two types of multivariate generalised regression models are developed; static models, pursuing a simple approach that would ensure feasibility of implementation to reactive control management systems, and dynamic models with one time step (Δt=3-minutes) ahead prediction, for integration into real-time proactive systems. Furthermore, for each type a model per regime is formed, as it is ensued by the clustering, which results to a static and a dynamic model for congested regimes, and a static and a dynamic model for uncongested regimes. Following the aforementioned analysis and the rationale that in the onset of congested conditions the left lanes receive the inflow, implying that the right lanes are already saturated, left lane density distribution ratio (LDDR) is introduced as determinant response variable for the congested regimes model, and of the right lane for the uncongested.

To assay the impact of the implementation of the algorithm to an existing system in terms of network performance, an exploratory analysis is effectuated at a managed lane and variable speed limit (VSL) system in a segment of a Swiss highway. Statistical assessment of the models indicates that both static and dynamic prediction models yield significant prediction accuracy, with the models for uncongested conditions maintaining marginally stronger prediction estimation, as their traffic patterns are more repetitive.

The perspectives of the study are lying on revalidating the conjecture, by assessing stationarity of the models in case of networks equipped or deprived of ITS control policies (e.g. managed lanes system, ramp metering etc.), of incident occurrence or of non-homogeneous flow. A possible extension could be the modeling of transition phases in the same location, as a system of nonlinear equations.

5. References

Amin, M. R. and J. H. Banks (2005). Variation in Freeway Lane Use Patterns with Volume, Time of Day, and Location. *Transportation Research Record: Journal of the Transportation Research Board*, Transportation Research Board of the National Academies, Washington, D.C., **1934**, 132–139.

Antoniou, C. and H. Koutsopoulos (2006). Estimation of Traffic Dynamics Models with Machine Learning Methods. *Transportation Research Record: Journal of the Transportation Research Board*, Transportation Research Board of the National Academies, Washington, D.C., **1965**, 103 – 111.

Brilon, W., J. Geistefeldt, and H. Zurlinden (2008). Implementing the Concept of Reliability for Highway Capacity Analysis. *Transportation Research Record: Journal of the Transportation Research Board*, Transportation Research Board, National Research Council, Washington D.C., **2027**, 1-8.

Chung, K. and M. J. Cassidy (2004). Test of theory of driver behavior on homogeneous freeways. *Transportation Research Record: Journal of the Transportation Research Board*, Transportation Research Board of the National Academies, Washington, D.C., **1883**, 14–20.

Cleveland, W. S., and S. J. Devlin (1988). Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association*, **83**, 596–610.

Daganzo, C.F. (2002). A behavioral theory of multi-lane traffic flow. Part I: long homogeneous freeway sections. *Transportation Research Part B: Methodological*, **36** (2), 131–158.

Duret, A., S. Ahn, and C. Buisson (2012). Lane Flow Distribution on a three-lane freeway: General features and the effects of traffic controls. *Transportation Research Part C: Emerging Technologies*, **24**, 157-167.

Geistefeldt, J (2012). Operational Experience with Temporary Hard Shoulder Running in Germany. Proceedings of the 91st Annual Meeting of the Transportation Research Board, Washington, D.C., January 2012.

Helbing, D., Treiber, M., Kesting, A. and M., Schönhof (2009). Theoretical vs. empirical classification and prediction of congested traffic states. *European Physical Journal B*, **69** (4), 583–598.

Ishak, S. and C. Alecsandru (2003). Optimizing Traffic Prediction Performance of Neural Networks under Various Topological, Input, And Traffic Condition Settings. Proceedings of the 82th Annual Meeting of the Transportation Research Board, Washington, D.C., January 2003.

Jiang X. and H. Adeli (2004). Wavelet Packet-Autocorrelation Function Method for Traffic Flow Pattern Analysis. *Computer-Aided Civil and Infrastructure Engineering*, **19**, 324–337.

22

Karim, A. and H. Adeli (2002). Comparison of the Fuzzy—Wavelet RBFNN Freeway Incident Detection Model with the California Algorithm. *Journal of Transportation Engineering*, **128** (1), 21–30.

Knoop V.L., A. Duret, C. Buisson and B. Van Arem (2010). Lane distribution of traffic near merging zones influence of variable speed limits. Proceedings of the *13th International IEEE Conference on Intelligent Transportation Systems* (ITSC), Madeira, Portugal 2010, 485-490.

Laval , J. and L. Leclercq (2008). Microscopic Modeling of The Relaxation Phenomenon Using a Macroscopic Lane-Changing Model. *Transportation Research Part B: Methodological,* **42**, 511-522.

Lighthill, M.J. and G. B. Whitham (1955). On kinematic waves. II. A theory of traffic flow on long crowded roads. Proceedings of the *Royal Society of London: Series A, Mathematical and Physical Sciences*, **229** (1178), 317–345.

Martinetz T., S. Berkovich and K. Schulten (1993). Neural-Gas Network for Vector Quantization and its Application to Time-Series Prediction. *IEEE Transactions on Neural Networks*, **4** (4), 558-569.

Richards, P.I. (1956). Shock waves on the highway, *Operations Research*, **4**, 42–51.

Samoili S., D. Efthymiou C. Antoniou and A.-G. Dumont (2013). Lane Flow Distribution Investigation of Hard Shoulder Running Freeways. *Transportation Research Record: Journal of the Transportation Research Board*, Transportation Research Board of the National Academies, Washington D.C., **2396**, 133-142.

Smith, B. L. and M. J. Demetsky (1997). Traffic Flow Forecasting: Comparison of Modelling Approaches. *Journal of Transportation Engineering*, **123** (4), 261–266.

Stathopoulos, A. and M. Karlaftis (2003). A Multivariate State-Space Approach for Urban Traffic Flow Modelling and Prediction. *Transportation Research Part C: Emerging technologies*, **11** (2), 121-135.

Van Lint J.W.C., Hoogendoorn S.P., H.J. van Zuylen (2005). Accurate Freeway Travel Time Prediction with State-Space Neural Networks Under Missing Data. *Transportation Research Part C: Emerging technologies*, **13** (5-6), 347-369.

Vlahogianni, E., N. Geroliminis and A. Skabardonis (2008). Empirical and Analytical Investigation of Traffic Flow Regimes and Transitions in Signalized Arterials. *Journal of Transportation Engineering*, **134** (12), 512-522.

Zhang, X. and J.A. Rice (2003). Short-term Travel Time Prediction. *Transportation Research Part C: Emerging technologies,* **11** (3-4), 187-210.