ESTIMATION OF THE PROBABILITY OF MERGING AT AN ENTRANCE RAMP OF A SUBURBAN FREEWAY

J. Mint Moustapha

Encadrant: D. Daucher
Lepsis, IFSTTAR

Directeur de thèse: B. Jourdain
Cermics, ENPC

Réunion Géri Animatic
Lyon, June 7th 2013

Plan

Site Angevin de Référence et d'Observations pour le Trafic

The site is located on the A87 in the east of Angers (western France)

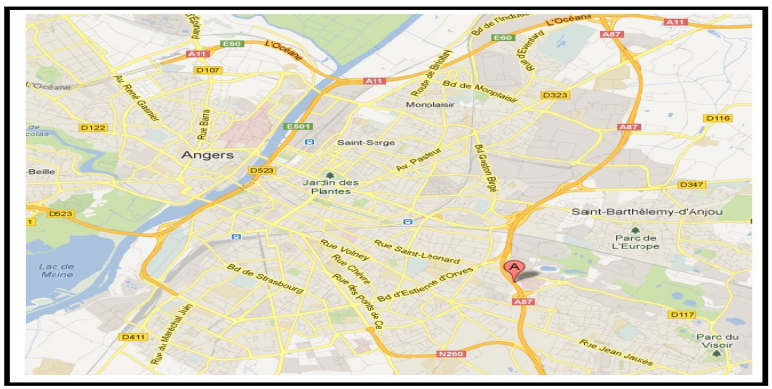


FIGURE : Position of the site

- The average flow can reach 22000 vehicles per day
- The speed is limited at 90km/h and the merging cars represent 18% of traffic
- 15% of the heavy vehicles

The aim of the installation

TO MEASURE IN REAL TIME THE DRIVERS BEHAVIOUR AND THE TRAFFIC ROAD STREAM IN ORDER TO ALLOW RESEARCH AGENCIES AND PROFESSIONALS TO ANALYZE AND DETECT THE DANGEROUS SITUATIONS.

Equipments

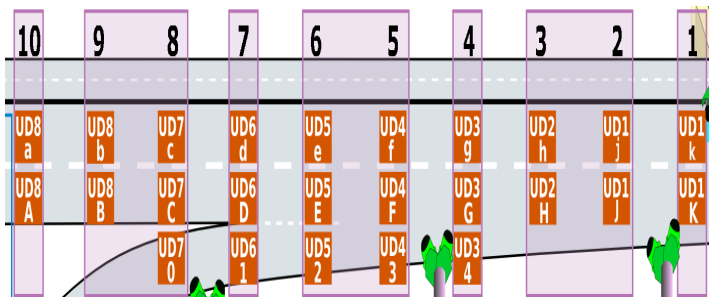


FIGURE : SAROT site

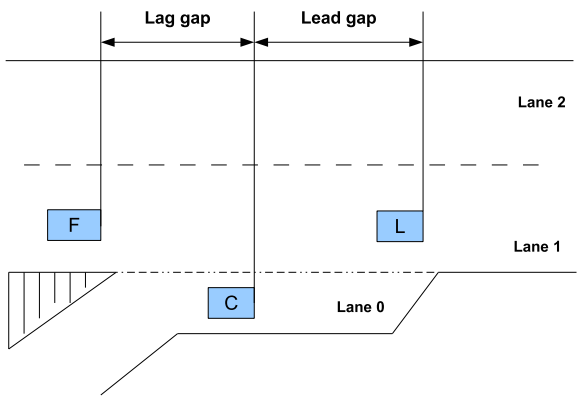
The measure instruments consist of :

- 10 sensors equidistributed and spaced of 50 meters, numbered from 10 to 1 in the direction of traffic flow.
- Each sensor line includes a couple of electromagnetic loops on the freeway lanes and 1 loop on the on-ramp.
- The merging vehicles are detected at sensor line 8 and the merging lane ends at point 3.

Definitions of the variables

Data : timestamp, speed, lane number, length of car, index reliability.

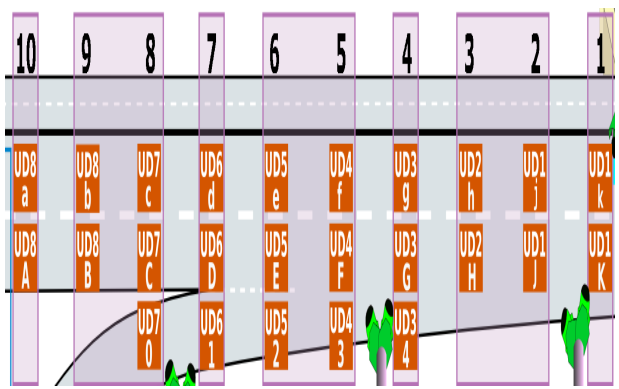
- Lead and lag gap time (T_L and T_F)
- Lead and lag relative speed (Dv_L and Dv_F)
- $T_L := t_L - t_C$
- $T_F := t_C - t_F$
- $Dv_L := V_C - V_L$
- $Dv_F := V_F - V_C$



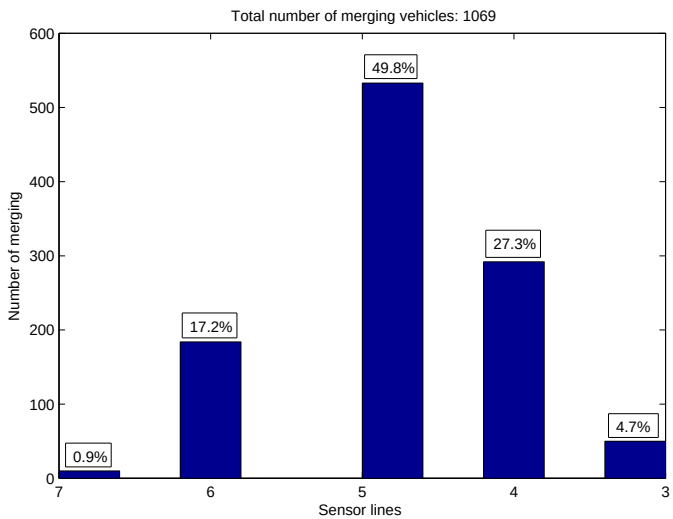
Merging framework for the merging vehicle C between its partners L and F

Merging Locations

Merging maneuvers may bring about an inconvenience either to the partner follower or to the merging vehicle if its partner does not give way and overtakes it.
An analysis of merging locations shows :



SAROT site

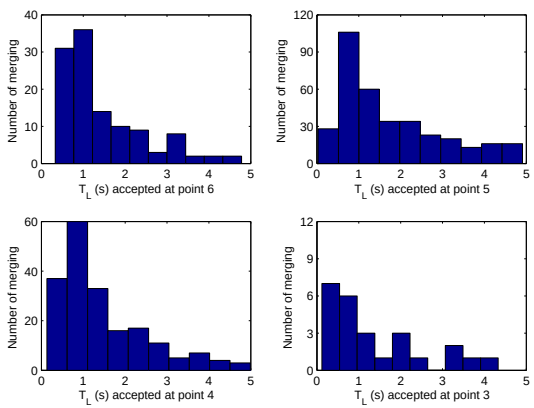


Histogram of merge location

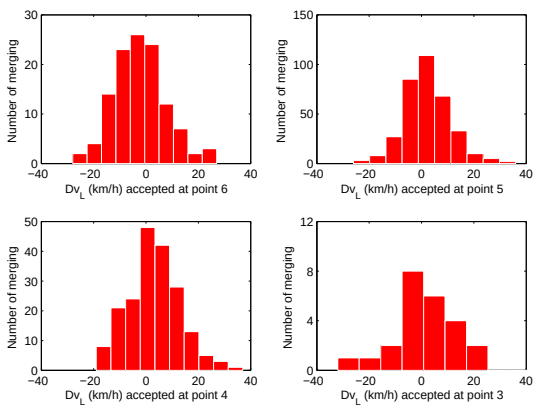
Median Analysis

Analysis of the median values shows :

- Median of the T_L accepted vary from 0.89s at sensor line 3 to 1.3s recorded at point 5
- The median of the T_F accepted are slightly higher : vary from 1.5s at point 5 to 2.1s at point 3
- The T_L refused are around 1.2s and the T_F refused are around of 1.5s at all measuring points
- Gaps (lead & lag) accepted less than 0.2s were recorded



Lead gaps T_L accepted at sensor lines 6,5,4 and 3



Relative Speeds Dv_L accepted at sensor lines 6,5,4 and 3

Kind of Modelling

THE STRUCTURE OF INDIVIDUAL DATA INVOLVE A MICROSCOPIC LEVEL FOR THE MODELLING

- $X_1 := V_C$
- $X_2 := T_L$
- $X_3 := Dv_L$
- $X_4 := T_F$
- $X_5 := Dv_F$
- $M_g := (X_1, X_2, X_3, X_4, X_5)$

LOGIT Model

Response variable

$$\begin{cases} Y = 1 & \textit{merge occurs} \\ Y = 0 & \textit{no merging} \end{cases}$$

The a posteriori merging probability at each sensor :

$$\eta(x) = \mathbb{P}(Y = 1 | X = x) = \frac{\exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x^j\right)}{1 + \exp\left(\beta_0 + \sum_{j=1}^{p} \beta_j x^j\right)}$$

- $p+1$ coefficient estimators of β_j are obtained from the likelihood maximization method
- The decision rule is to assign a "merge" when $\hat{\eta}_i > \frac{1}{2}$, otherwise we assign "no merge" for the vehicle i

The used data

- Data collected between 7a.m. and 10a.m.
- 3049 merging vehicles with both a leader and follower partners at the sensor line before their merging
- The data sample is separated into two parts : a learning basis and another set to test the model
- No merging before point 7 and no merging at point 3

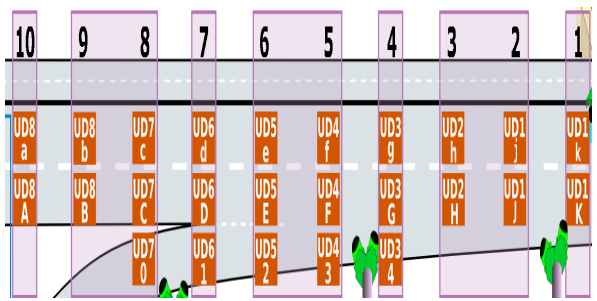
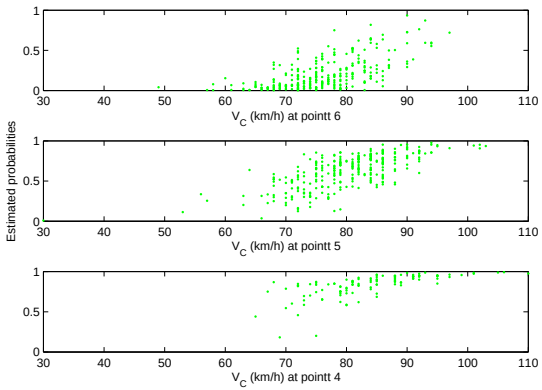
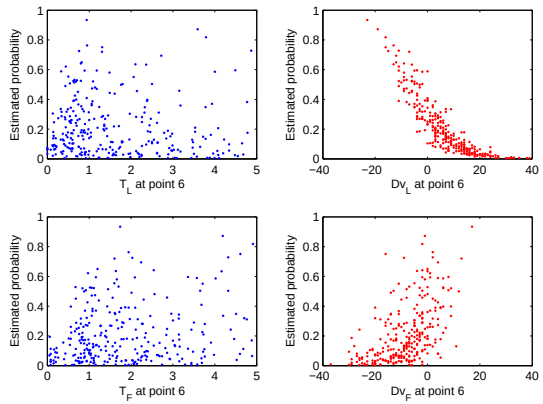


FIGURE : SAROT site

Graphs of a posteriori Probabilities



Estimated probabilities depending on V_C at lines 6, 5 and 4



Estimated probabilities depending on T_L, Dv_L, T_F, Dv_F at line 6

- On the left, the highest speeds have a higher probability to merge at lines 5 and 4 : drivers accelerate to merge
- On the right, vehicles C moving slower than L have a higher predicted probability to merge
- Model discriminates smaller dependence on the variables with follower

Results of statistic model

- M_g yields a success rate of 82% at line 6, 72% and 90.3% at sensor lines 5 and 4
- M_g predicts 93% of the good no-merging and 31.5% of the good merging at point 6
- At point 5, it predicts 86.4% of the good merging and 49.6% of the good no-merging
- Highest success rate for the model at point 4 : 99% of prediction of the good merging but only 23% of the good no-merging

Model Selection

- Some variables do not provide information about the estimation of the a posteriori probability and complicate the model.
- To perform, we use statistical tests and model selection criteria to test the significance of parameters β_j of X_j
- Testing the $2^p - 1$ models derived from the possible combinations of X_j.

Selection Criteria : AIC, BIC

The log-likelihood function of the model

$$\ell(\Theta; y) = \sum_{i=1}^{n} y_i \ln(\eta_i) + (1 - y_i) \ln(1 - \eta_i)$$

- The Akaike information criterion : $AIC = -2 \times \ell + 2p$
- The Schwartz information criterion : $BIC = -2 \times \ell + p \times \ln n$.
  The selected model is the one that has the smallest information AIC, BIC

Hypothesis Testing : Wald test

- Allows testing the nullity of q ($1 \leq q \leq p$) parameters of M_g :

$$\mathcal{H}_0 : \beta_{j_1} = \beta_{j_2} = ... = \beta_{j_q} = 0, \ \forall 1 \leq j_1 \leq ... \leq j_q \leq p$$

against,

$$\mathcal{H}_1 : \overline{\mathcal{H}_0}$$

Wald statistic : $\Lambda_q = \hat{\Theta}'_q \hat{\Sigma}_q^{-1} \hat{\Theta}_q$, reject area :$\Lambda_q \geq \chi^2_{1-\alpha}(q)$

1. Start to share the set of submodels of M_g into four groups of same number of parameters
2. Apply the Wald test to all models of each cluster.
3. The model validated by the test is identified
4. When several models are accepted, the model choosen is the optimal model selected for the AIC criterion
5. We obtain an optimal model on each group
6. **The final selected model is the one with the Wald statistic Λ_q the most discarded from its reject area**.

Hypothesis Testing : LRT test

- The LRT selects the best model among nested models.
  1. Set a benchmark model M_g to compare to its submodels M_s with p' ($\leq p$) parameters. The hypothesis to be tested is :
     $\mathcal{H}_0$: M_s *is more suitable than* M_g
     against, $\mathcal{H}_1$: $\overline{\mathcal{H}_0}$
  2. For each identified submodel, we calculate the associated deviance $D_s = -2 \times \ell$
  3. LRT statistic : $\Delta D = D_s - D_g$. Under $\mathcal{H}_0$, $\Delta D \hookrightarrow \chi^2(p - p')$.
  4. Reject area : $\Delta D > \chi^2_{1-\alpha}(p - p')$
  5. If $\mathcal{H}_0$ is favoured for several models, the model which has the smallest AIC value is kept.

Comparison of models selected by different criteria

Sensor	AIC	BIC	LRT	Wald test	Success Rate
6	Dv_L	Dv_L	$T_L + Dv_L + T_F$	$Dv_L + T_F + Dv_F$	$T_L + Dv_L + Dv_F$
					82.6%
5	$V_C + Dv_L$	Dv_L	$V_C + Dv_L$	$V_C + Dv_L$	$V_C + T_L + Dv_L$
					71.6%
4	V_C	V_C	V_C	$V_C + Dv_L$	$V_C + T_L + Dv_L + T_F + Dv_F$
					90.3%

- At sensor 6, the drivers mainly consider the relative speed Dv_L to make their "merging" or "no-merging" decision
- At sensor 5, the drivers seem to consider also their own speed V_C together with Dv_L to take a decision
- All criteria and tests agree on the importance of the variable V_C at point 4

Conclusion

The statistical modelling argued that to take its decision (merge or not merge), the merging driver mainly refers to V_C and Dv_L, and this, at the beginning and middle of lane 0.
But at the end of the merging lane, only its own speed adjustment matters because he/she wants to merge even forcing to avoid stopping at the end of merging lane.

Considering only the partner leader in *logit* model

The same procedure was applied to data of merging vehicles with only leader. The general model is : $M_g = (V_C, T_L, DV_L)$

- 80.8% of succes at line 6 against 82% previously
- 74% of succes at line 5 against 72%
- 91.5% of succes at line 4 against 90%

The most influential variables still are the speed of the merging vehicle and the relative speed of the merging vehicle with respect to its partner leader

THE CHOICE OF THE MOST INFLUENTIAL VARIABLES IS DECISIVE IN THE *logit* MODELLING AND WILL SERVE IN ORDER TO CONSTRUCT A BEHAVIOURAL MODEL.

Gap acceptance model

- The concept of the gap acceptance models is to select for each variable a given gap and to choose to merge if this critical gap is exceeded.
- The tested gap acceptance model are composed from combinations of conditions on the following variables.

Building of the Model

- The individual variables : V_C, T_L, T_F, Dv_L, Dv_F.
- Ratio of the speed of lead vehicle and the speed of merging vehicle $\frac{V_L}{V_C}$.
- Ratio of the speed of lag vehicle and the speed of merging vehicle $\frac{V_F}{V_C}$.
- Gap distance with the leader vehicle $T_L \times V_C$.
- Gap distance with the follower vehicle $T_F \times V_C$.
- Some estimators of the time to collision such as : $\frac{T_L \times V_C}{V_C - V_L}$, $\frac{T_L \times V_C}{V_L - V_C}$, $\frac{T_F \times V_C}{V_F - V_C}$, and $\frac{T_F \times V_C}{V_C - V_F}$.

Choice of thresholds

The output of X_j from the *logit* model will be used to choose the gaps :
The realizations belonging the hyperplane

$$\mathbb{H} = \{x_j \in \mathbb{R}^p : \hat{\eta}(x_j) = \frac{1}{2}\}$$

are used as input for critical gaps.

Results

- When the vehicles have a leader and a follower partners, the selected models are less successful at points 5 and 4 where the success rate does not exceed respectively 70% and 88%.
- When the vehicles just have a partner leader, the selected models are :
  - At sensor line 5 :

    $$V_C \geq \delta_v \text{ and } V_C \times T_L \geq \delta_x$$

    with a success rate equal to 74%. Where $\delta_v = 75km/h$ and $\delta_x = 9.6m$.
  - At measuring point 4, several models give the same success rate equal to 91.5% with always the shared condition on $V_C \geq \delta_v$.
    One of these models is : $\frac{V_L}{V_C} \geq \delta$ and $V_C \times T_L \geq \delta_x$ with thresholds $\delta_v = 58km/h$, $\delta = 0.74$ and $\delta_x = 7.5m$.
  - At point 6 : $V_L \geq V_C - \delta_v$ yields 84% of success rate and the threshold $\delta_v = 13km/h$.

CONCLUSION

- Over 95% of merging were finished before the end of merging lane : resizing the lane merging.
- Some drivers accept a short time gap $< 2s$.
- Merging drivers consider mostly the information with their partner leader.
- Both modelling agree on the significance of Dv_L in the merging decision at the beginning of merging lane and V_C at the ending of lane.
- The methodology would be validated and the results would be improved if we had a finer mesh of data, which would allow a better follow up of vehicles.