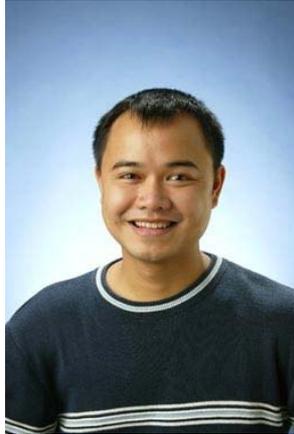


Minh Hai Pham

Ingénieur en Informatique
Assistant-Doctorant au Laboratoire des voies de circulation (LAVOC) à l'École Polytechnique Fédérale de Lausanne (EPFL) dans le domaine de la télématique

Vers un modèle proactif pour identifier des risques de trafic

La sécurité routière prend de plus en plus d'importance dans de nombreux pays. En mettant l'accent sur la prévention des accidents sur les autoroutes, nous proposons une méthodologie pour identifier les risques en utilisant des modèles de régression logistique. La caractéristique novatrice de cette méthodologie réside dans l'évaluation des risques de trafic par rapport à différentes régimes de trafic obtenues par des Cartes Auto-Organisatrices (Self-Organizing Maps - SOM). Nous agrégeons les données de trafic et les données météorologiques pour produire des états de trafic pour des intervalles de 5 minutes, que nous appelons des situations de trafic. En utilisant l'analyse en composantes principales (ACP), qui permet de réduire le nombre de dimensions et de supprimer le bruit aléatoire des données, nous transformons les situations de trafic pour le processus de regroupement par les SOM. Ce processus de regroupement produit des groupes de situations de trafic similaires que l'on appelle des régimes de trafic. De la base de données des accidents, nous déterminons les situations de trafic qui précèdent les collisions, appelées situations pré-accidentelles. Avec les résultats de regroupement obtenus, nous classons les situations pré-accidentelles dans des régimes de trafic. Pour chaque régime de trafic, nous développons un modèle de régression logistique pour identifier les situations à risque. Les résultats montrent que les modèles peuvent identifier correctement un pourcentage élevé de situations à risque, tandis que le taux de fausse détection est faible.



Vers un modèle proactif pour identifier des risques de trafic

Minh-Hai PHAM

Doctorant, EPFL-ENAC-LAVOC



EPFL
LAVOC

Motivation

- Accidents routières:
 - 1,2 million de personnes tuées par année dans le monde entier
 - 11e cause majeure de décès, 2,1% de tous les décès dans le monde
- Accidents autoroutières:
 - Sévères dus à la grande vitesse





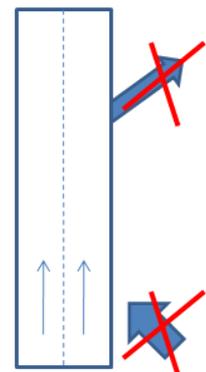
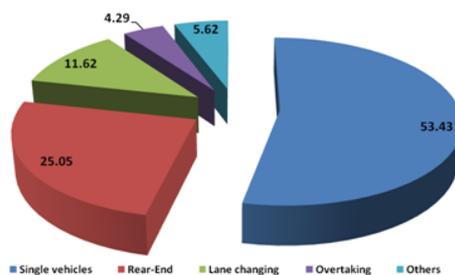
Contenu

- Motivation
- Champ d'étude
- Études précédentes
- Sites d'étude
- Méthodologie
- Résultats & Analyses
- Conclusions
- Perspectives



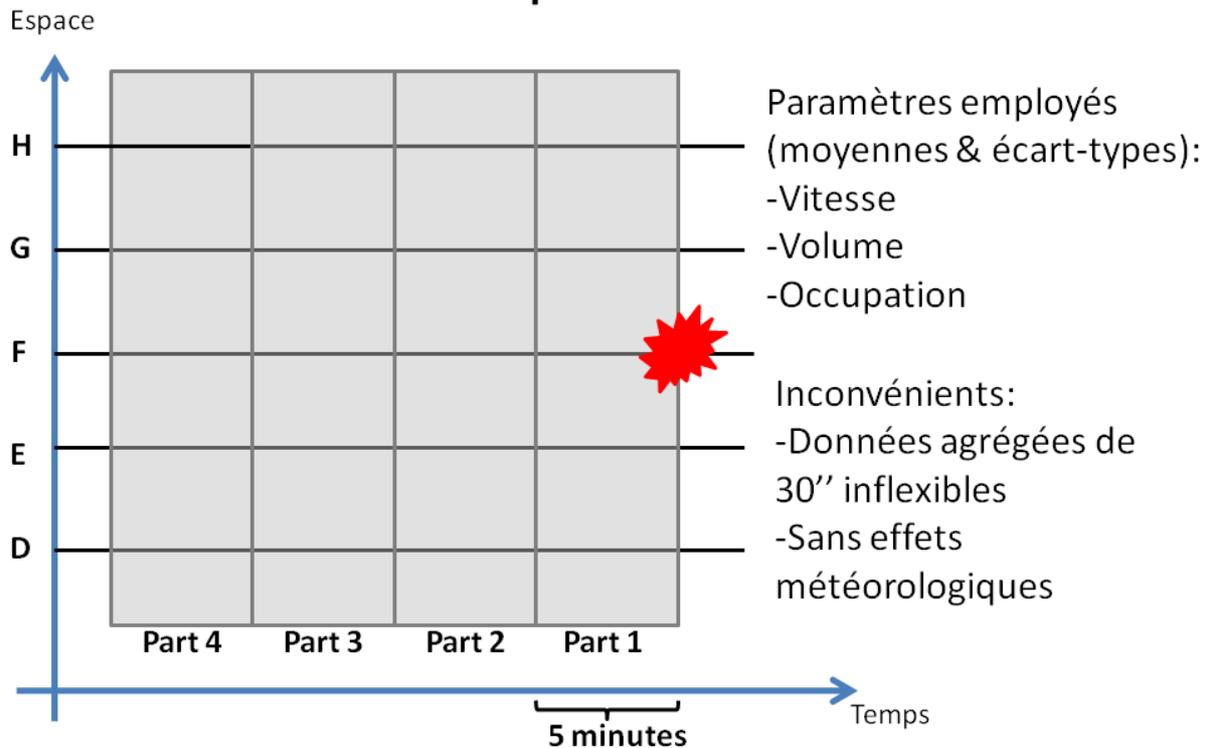
Champ d'étude

- Accidents induits par la circulation
 - Mauvaise coordination entre les conducteurs
 - Conducteurs non informés des conditions en aval
 - Mauvais comportements des groupes de conducteurs
- Sites d'étude potentiels
 - Sections d'autoroutes: planes, droites
 - Pas d'entrées/sorties
 - Pas d'échangeurs
- Types d'accidents
 - Accidents Non-DUI
 - Tamponnement (Frontal-Arrière)





Études précédentes

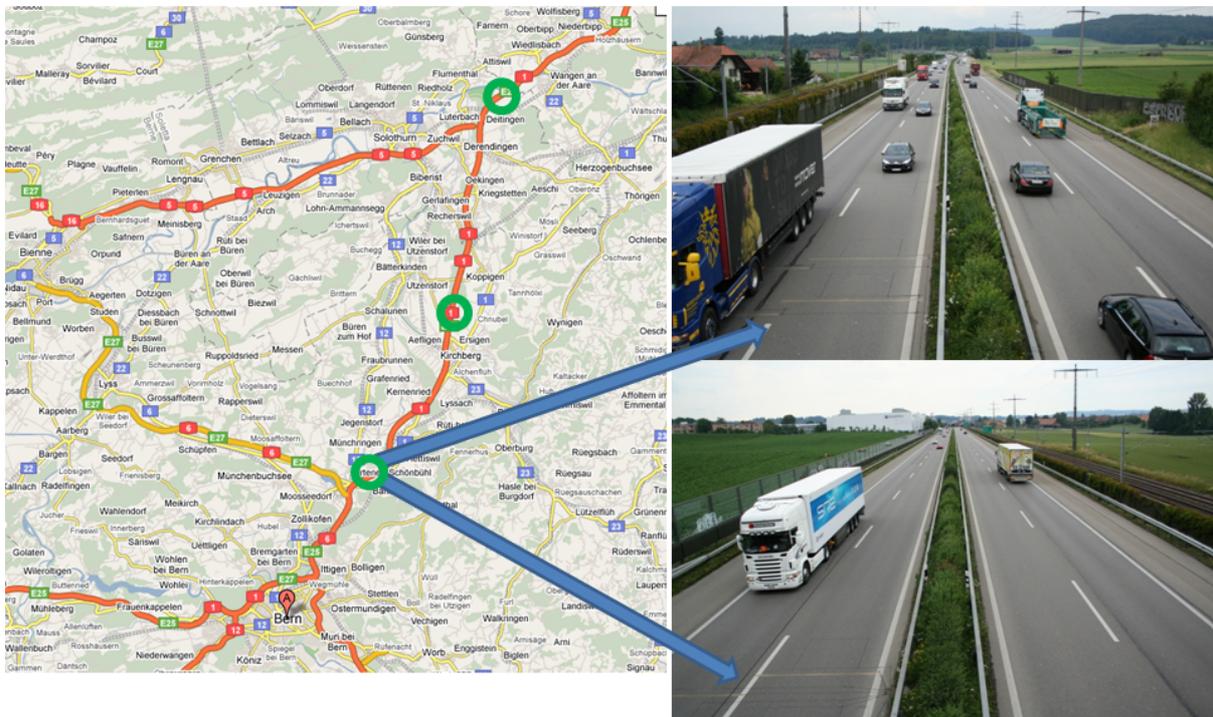


Conditions en Suisse

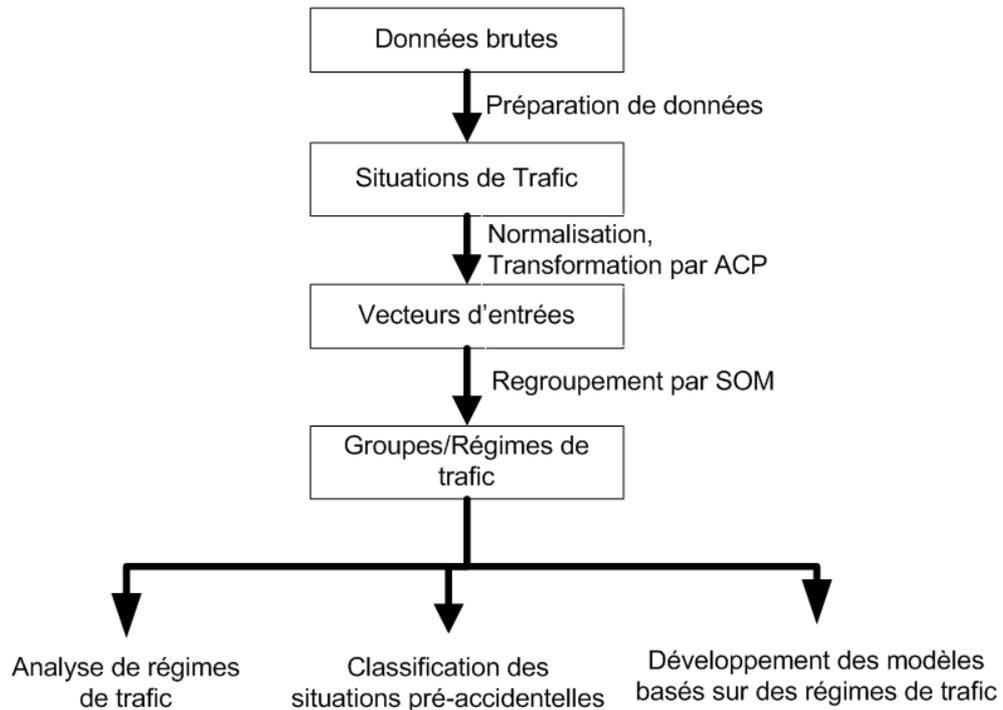
- En janvier 2009, plus de 300 détecteurs à double boucles installés sur les autoroutes
- Disponibilité des données individuelles
- Stations météorologiques à proximité des détecteurs (MétéoSuisse/Boschung)
- **Mais, distance élevée entre les détecteurs**



Sites d'étude



Méthodologie





Situations de trafic

Paramètre	Variable	Validité	Note
Moment du jour	TimeDay	2 voies	Index de la période de 5 minutes pour une journée, de 1 à 288
Jour de la semaine	WDay	2 voies	lundi (1) à dimanche(7)
Nombre de véhicules	Flow	Chaque voie	Pour chaque période de 5 minutes
Vitesse Moyenne	ASpeed	Chaque voie	en km/h
TIV moyen	AHeadway	Chaque voie	en seconde
Occupation	Occupancy	Chaque voie	en %
Variation de vitesse	VSPEED	Chaque voie	
Variation de TIV	VHeadway	Chaque voie	
Pourcentage de poids lourds	%HeavyVeh	Chaque voie	
Diff. de vitesse entre les deux voies	Speed#	2 voies	Différence des vitesses moyennes entre la voie droite et la voie gauche (km/h)
Précipitation	Precipitation	2 voies	
Diff. de vitesse par rapport à la situation précédente	SpeedChange	Chaque voie	en km/h
Diff. de débit par rapport à la situation précédente	FlowChange	Chaque voie	

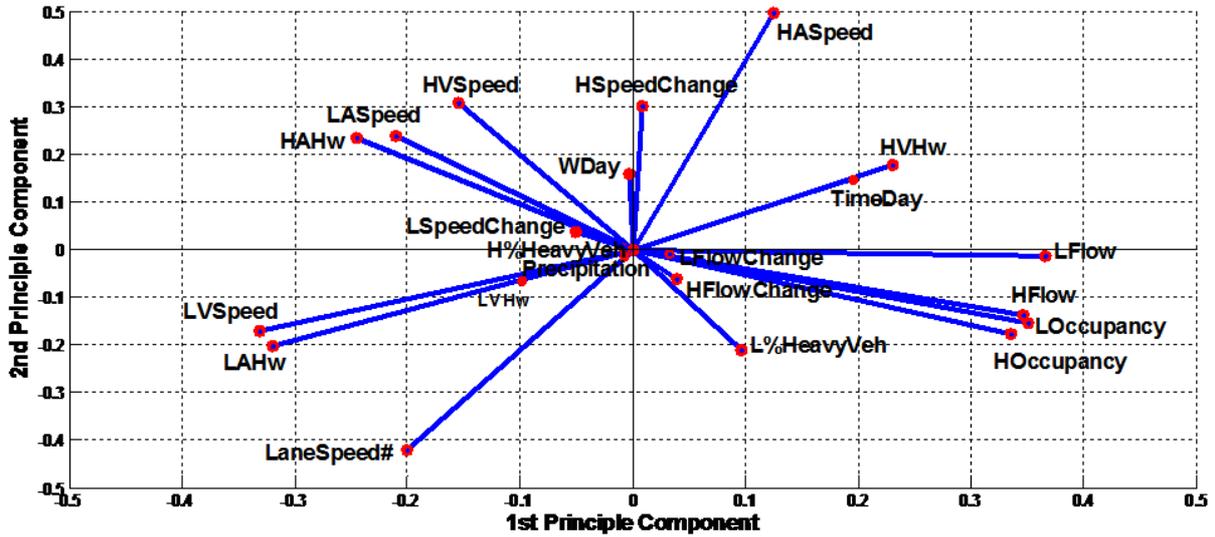


Vecteurs d'entrées

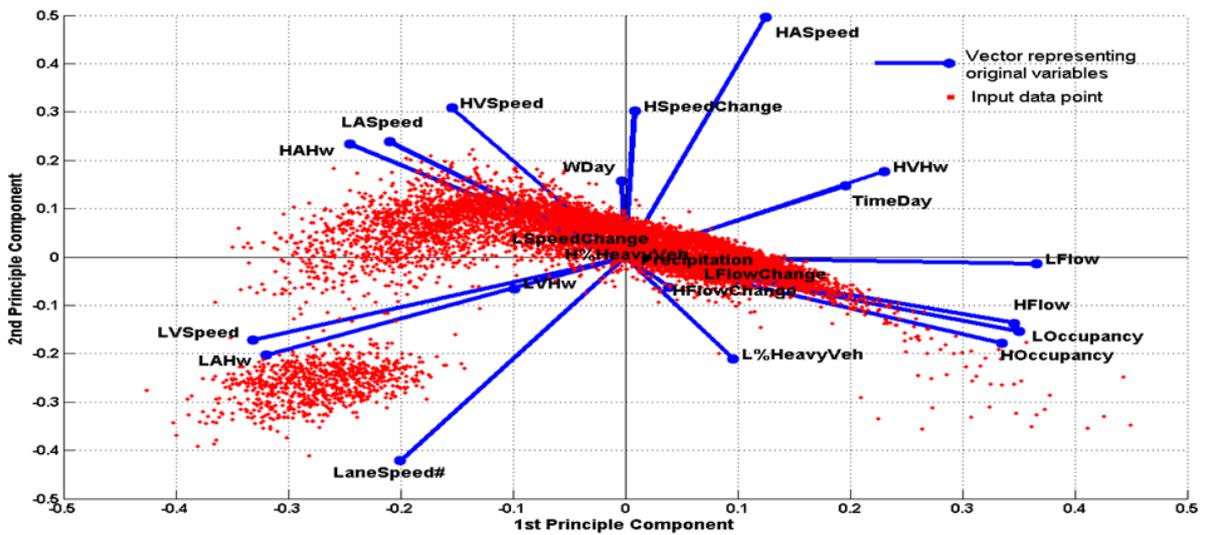
- Normalisation
 - Mise à la même échelle des champs de données
- Transformation par ACP
 - Réduction de la dimension
 - Elimination des bruits aléatoires
 - Visualisation des données grâce aux composantes principales



Transformation par ACP



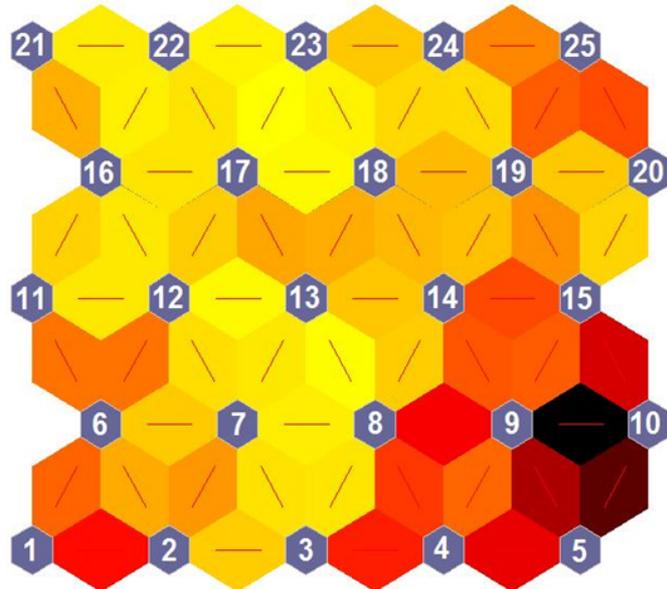
Transformation par ACP





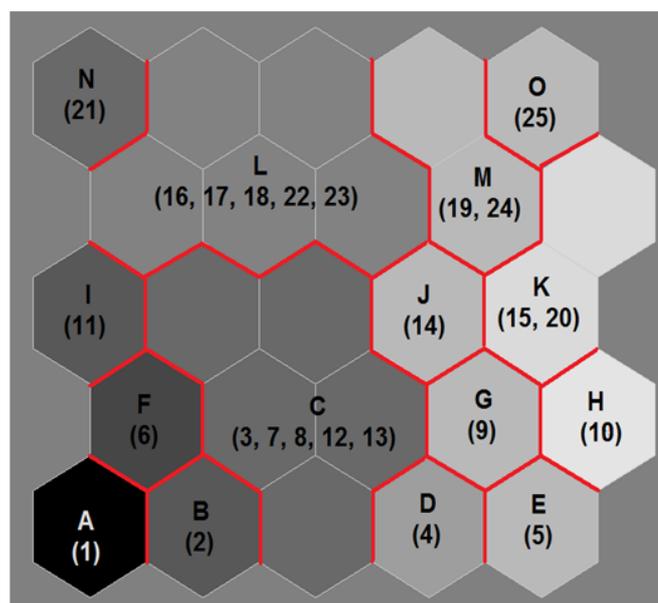
Regroupement par SOM

- 230'000 entrées
- 6 dimensions
- Cartes Auto-organisatrice de 5x5



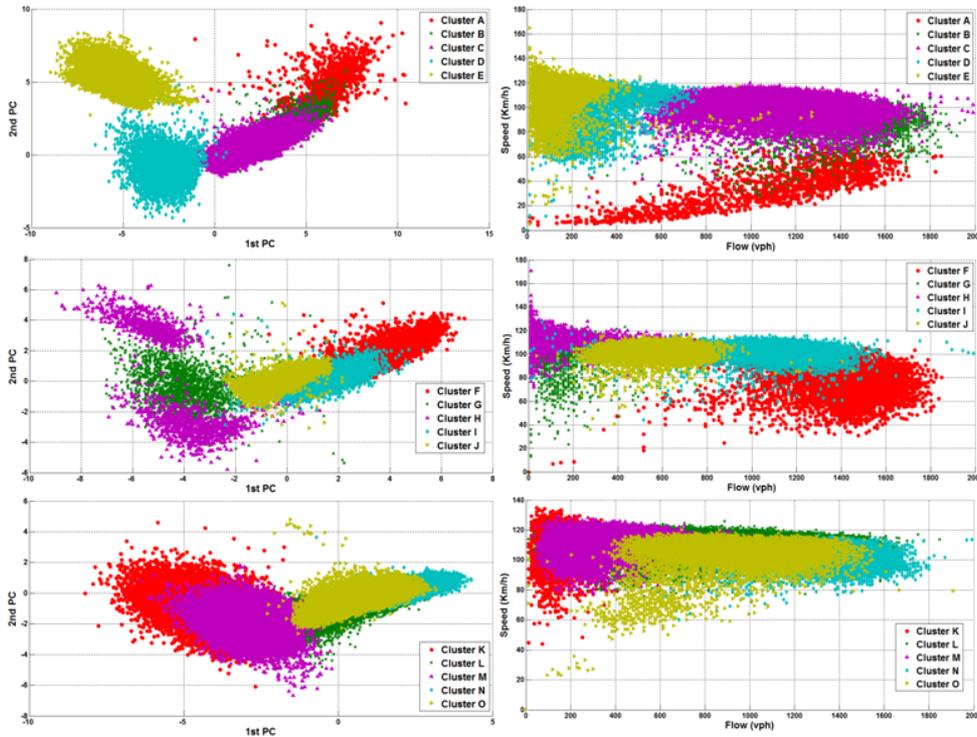
Post-regroupement

- Fusion des groupes similaires



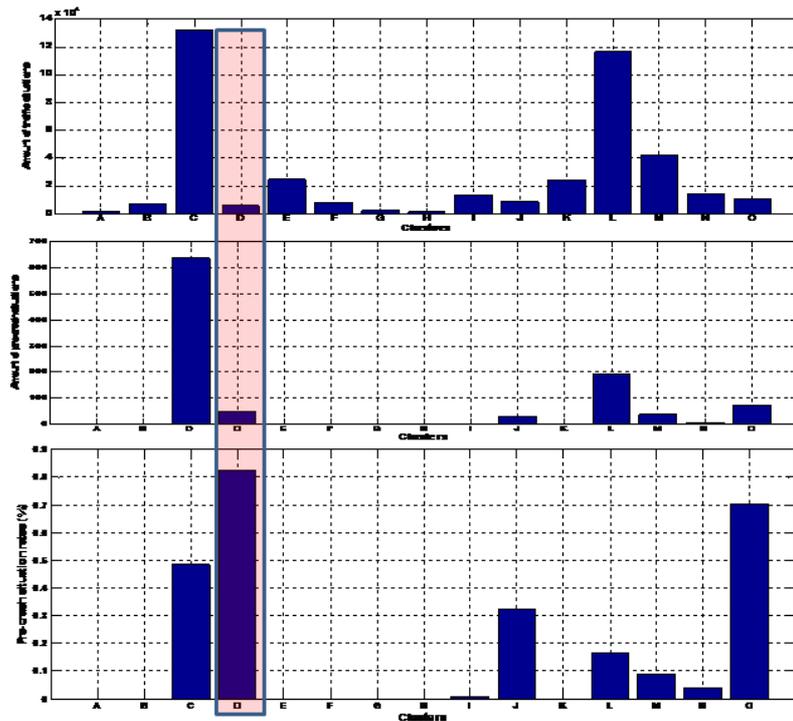


Régimes de trafic



Situations pré-accidentelles

- Situations pré-accidentelles classées dans des régimes de trafic





Modèles de régression logistique (1)

- Pour chaque régime de trafic, construction d'un modèle de régression logistique:

$$TS_i = (X_1, X_2, \dots, X_n)$$

$$Y(TS) = \frac{1}{1 + e^{\alpha TS + \beta}}$$

- Y=0 si TS est une situation normale
- Y=1 si TS est une situation pré-accidentelle



Modèles de régression logistique (2)

- Données pour calibration: 80%
- Données pour validation: 20%
- Au niveau statistique, seulement les régimes C, D, J, L, et O sont représentatifs

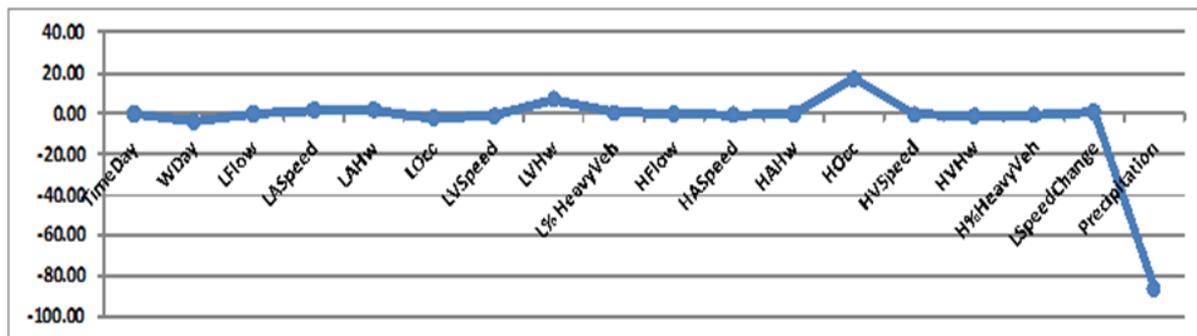


Résultats pour régime D (1)

Source	Sum of Squares	Degree of Freedom	MS	F Test	Probability	Coefficients
Constant	0.38	1	0.38	89.00	<0.001	-235.95
TimeDay	0.10	1	0.10	24.00	1.20E-06	-0.03
WDay	0.37	1	0.37	89.00	<0.001	-3.87
LFlow	0.77	1	0.77	180.00	<0.001	0.05
LASpeed	0.97	1	0.97	230.00	<0.001	1.78
LAHw	3.57	1	3.60	850.00	<0.001	1.81
LOccupancy	0.21	1	0.21	49.00	3.00E-12	-2.01
LVSpeed	1.95	1	1.90	460.00	<0.001	-1.02
LVHw	9.54	1	9.50	2300.00	<0.001	6.92
L% HeavyVeh	0.21	1	0.21	51.00	1.20E-12	0.33
HFlow	0.05	1	0.05	12.00	0.00047	-0.04
HASpeed	0.16	1	0.16	37.00	9.90E-10	-0.63
HAHw	0.03	1	0.03	6.90	0.0087	-0.10
HOccupancy	0.04	1	0.04	10.00	0.0013	17.03
HVSpeed	0.19	1	0.19	45.00	2.40E-11	-0.35
HVHw	0.07	1	0.07	16.00	7.30E-05	-1.26
H%HeavyVeh	0.29	1	0.29	69.00	1.10E-16	-0.51
LSpeedChange	0.14	1	0.14	34.00	5.80E-09	0.63
Precipitation	3.36	1	3.40	800.00	<0.001	-86.00
Residual	23.61	5603	0.004			
Measure	46	5622				



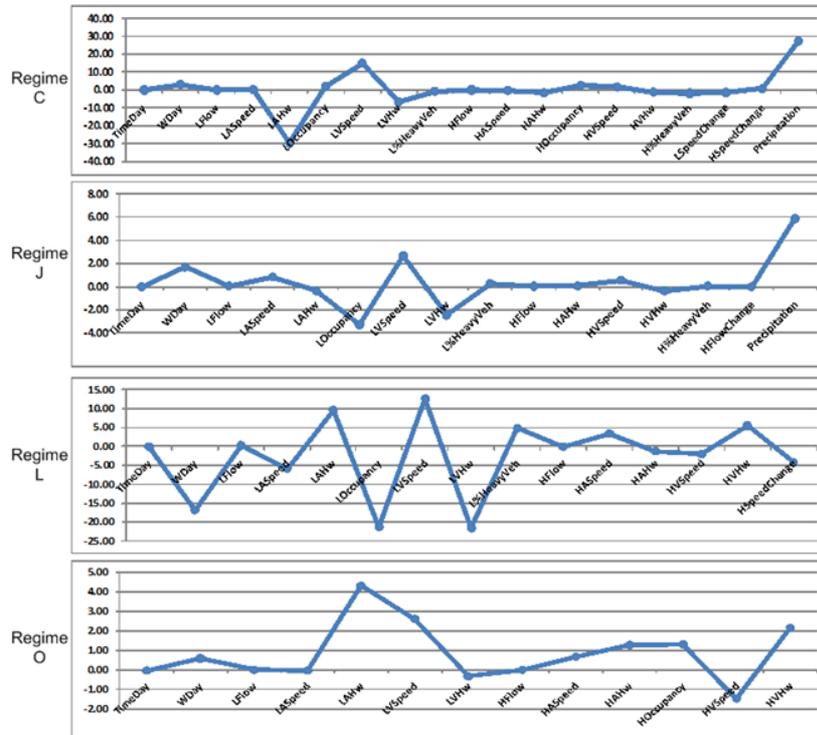
Résultats pour régime D (2)



Régime de trafic	Données pour calibration (erreurs)		Données pour validation (erreurs)	
	Normal	Pré-acc.	Normal	Pré-acc.
D	4461 (0)	37 (0)	1115 (0)	9 (1)



Facteurs importants



Résumé des résultats

Régime de trafic	Données pour calibration (erreurs)		Données pour validation (erreurs)	
	Normal	Pré-acc.	Normal	Pré-acc.
C	106291 (0)	510 (0)	26573 (1)	128 (1)
D	4461 (0)	37 (0)	1115 (0)	9 (1)
J	6961 (0)	22 (0)	1740 (0)	6 (0)
L	92352 (0)	154 (0)	32312 (0)	38 (2)
O	8239 (0)	58 (0)	2060 (0)	14 (0)



Conclusions (1)

- Processus de développement comprend deux étapes:
 - Définition des régimes de trafic
 - Construction des modèles de régression logistique
- Les facteurs importants:
 - Occupation
 - TIV moyen
 - Écart-type de vitesse
 - Précipitation



Conclusions (2)

- Bonne performance des modèles de régression logistique:
 - Taux de détection élevé (reconnaissance du risque)
 - Taux de fausse détection faible (normale->pré-accidentelle)

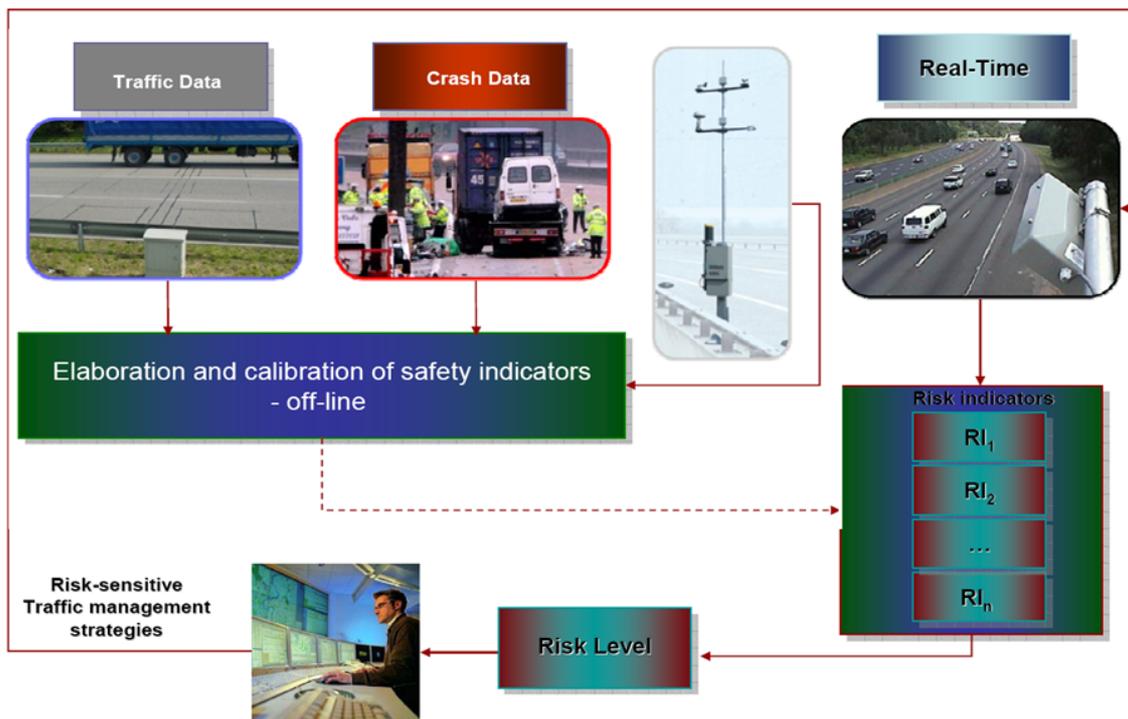


Perspectives

- Plus d'informations météorologiques lors de la définition des situations de trafic
- Test avec autres sites en Suisse
 - Test de robustesse du modèle
- Test avec intervalle d'agrégation plus courte
 - Plus vite le risque est identifié, plus tôt la décision peut être prise
- Considération de l'évolution des régimes avant les accidents



Applicabilité



Merci de votre attention