L'Institut wallon de l'évaluation,
de la prospective et de la statistique

Wallonie
Iweps

2020

MOBILITÉ

REGARDS STATISTIQUES

N° 4

Les accidents de la route en Wallonie de 2009 à 2018

De l'analyse locale des causes enregistrées aux pistes d'aide à la décision

COLOPHON

Auteur : **Baptiste Féraud** (IWEPS)

Édition : **Evelyne Istace** (IWEPS)

Éditeur responsable : **Sébastien Brunet**
(Administrateur général, IWEPS)

Création graphique : **Deligraph** http://deligraph.com

Reproduction autorisée, sauf à des fins commerciales, moyennant mention de la source.

Dépôt légal : **D/2020/10158/19**
ISBN : **2593-5542**

IWEPS

Institut wallon de l'évaluation, de la prospective et de la statistique

Route deLouvain-La-Neuve,2
5001 BELGRADE - NAMUR

Tel : 081 46 84 11

http://www.iweps.be

info@iweps.be

RESUME

Cette recherche s'intéresse à la compréhension au niveau communal des causes enregistrées des accidents de la route survenus en Wallonie entre 2009 et 2018. Lorsqu'un accident de la route est constaté, il est important pour éclairer les décideurs d'en comprendre les caractéristiques, les circonstances et les explications éventuelles. Comprendre ce qu'il s'est passé permet de mieux éviter que les mêmes scénarios ne se reproduisent. Pour cela, chaque accident doit être catalogué et référencé. Ce type de référencement va inévitablement provoquer un flux d'informations qui prendra le plus souvent la forme de variables catégorielles dans un jeu de données. Ces variables catégorielles ou facteurs (ex : type de route) impliquent chacune plusieurs niveaux ou modalités (ex : autoroute, route régionale, route communale).

Dans ce contexte, le premier objectif de cette publication est de fournir un état des lieux, une sorte de petit *vade-mecum*, afin de comprendre comment traiter statistiquement de telles variables catégorielles, de façon tout à fait générale. Partant de la notion d'indépendance entre variables catégorielles, des méthodes univariées (Analyse Factorielle des Correspondances simple) et multivariées (Analyse des Correspondances Multiples) sont montrées, décrites et interprétées. L'application directe de ces méthodes aux données d'accidents de la route (mortels ou non, au niveau des communes wallonnes, pour les accidents recensés entre 2009 et 2018) fournit des sorties graphiques diverses : représentations univariées, images instantanées dans une commune à la fois, représentations cartographiques globales des contributions des modalités des variables, etc… Tous ces outputs sont donc disponibles et fournissent tant des informations absolues que relatives.

Ils aboutissent *in fine* à une série de pistes de réflexion pouvant aider à une prise de décision au niveau local et/ou à l'activation de leviers d'action.

TABLE DES MATIÈRES

9 INTRODUCTION

11 DONNÉES DISPONIBLES

2.1. Tableaux de contingence indépendants sur des facteurs pouvant expliquer les accidents de la route 12

2.2. Données multivariées, profils élargis des accidents de la route 13

14 ANALYSES NUMÉRIQUES

3.1. Tableaux de contingence indépendants sur des facteurs pouvant expliquer les accidents de la route 15

3.2. Données multivariées, profils élargis des accidents de la route 17

3.2.1. Facteur « état de la chaussée » 17

3.2.2. Facteur « période de la semaine » 18

3.2.3. Facteur « luminosité » 20

3.2.4. Facteur « type de route » 22

3.2.5. Facteur « vitesse maximale autorisée » 23

3.2.6. Facteur « type de collision » 24

3.2.7. Autres conditions particulières de circulation 26

3.3. Tableaux de contingence indépendants sur des facteurs pouvant expliquer les accidents de la route 28

3.3.1. Représentations multivariées des causes d'accidents au niveau communal 28

3.3.2. Discussion : analyse par commune versus perte de généralisation 31

3.3.3. Représentations cartographiques générales des contributions des modalités des facteurs 31

3.3.4. Tableau récapitulatif des ACM commune par commune 37

3.3.5. Perspective suivante : *clustering* des communes 44

46 | PISTES D'AIDE À LA PRISE DE DÉCISION

48 | CONCLUSIONS

50 | RÉFÉRENCES

52 | ANNEXES

A.1. Cadre méthodologique élargi 53

A. Variables catégorielles et principe d'indépendance 53

B. Analyse factorielle des correspondances simples (AFC) 55

C. Analyse factorielle des Correspondances Multiples (ACM) 59

A.2. Exemple de sortie numérique détaillée pour une AFC (coordonnées, cosinus carrés et contributions des modalités et des communes – variable « état de la chaussée ») 62

A.3. Exemple de sortie numérique détaillée pour une ACM (commune de Liège – accidents mortels) 69

A.4. Exemple de sortie numérique détaillée pour une ACM (commune de Herbeumont – accidents mortels) 70

A.5. Cartes supplémentaires basées sur les accidents mortels 71

Liste des graphiques, cartes et tableaux

Graphique 1 : Dimensions de l'AFC et variance expliquée (variable « état de la chaussée ») 17

Graphique 2 : Carte des profils ligne et colonne dans le plan formé par les deux premiers axes principaux de l'AFC (variable « état de la chaussée ») 18

Graphique 3 : Dimensions de l'AFC et variance expliquée (variable « période de la semaine ») 19

Graphique 4 : Carte des profils ligne et colonne dans le plan formé par les deux premiers axes principaux de l'AFC (variable « période de la semaine ») 19

Graphique 5 : Dimensions de l'AFC et variance expliquée (variable « luminosité ») 20

Graphique 6 : Carte des profils ligne et colonne dans le plan formé par les deux premiers axes principaux de l'AFC (variable « luminosité ») 21

Graphique 7 : Carte des profils ligne et colonne dans le plan formé par les deux premiers axes principaux de l'AFC (variable « type de route ») 22

Graphique 8 : Dimensions de l'AFC et variance expliquée (variable « vitesse maximale autorisée ») 23

Graphique 9 : Carte des profils ligne et colonne dans le plan formé par les deux premiers axes principaux de l'AFC (variable « vitesse maximale autorisée ») 24

Graphique 10 : Dimensions de l'AFC et variance expliquée (variable « type de collision ») 25

Graphique 11 : Carte des profils ligne et colonne dans le plan formé par les deux premiers axes principaux de l'AFC (variable « type de collision ») 25

Graphique 12 : Carte des profils ligne et colonne dans le plan formé par les deux premiers axes principaux de l'AFC (sur base des conditions particulières de circulation) 27

Graphique 13 : Projections des modalités des facteurs, commune de Liège. En haut, pour les accidents corporels mortels ; en bas, pour les accidents corporels avec blessés 29

Graphique 14 : Projections des modalités des facteurs, commune de Herbeumont. En haut, pour les accidents corporels mortels ; en bas, pour les accidents corporels avec blessés 30

Carte 1 : Représentation des contributions ACM cumulées – Accidents avec blessés – Modalité « autoroute » 33

Carte 2 : Représentation des contributions ACM cumulées – Accidents mortels – Modalité « autoroute » 33

Carte 3 : Représentation des contributions ACM cumulées – Accidents avec blessés – Modalité « agglomération » 34

Carte 4 : Représentation des contributions ACM cumulées – Accidents avec blessés – Modalité « hors agglomération » 34

Carte 5 : Représentation des contributions ACM cumulées – Accidents avec blessés – Modalité « aube, crépuscule » 35

Carte 6 : Représentation des contributions ACM cumulées – Accidents avec blessés – Modalité « nuit sans éclairage public » 36

Carte 7 : Représentation des contributions ACM cumulées – Accidents avec blessés – Modalité « avec un piéton » 36

Carte 8 : Représentation des contributions ACM cumulées – Accidents avec blessés – Modalité « contre un obstacle situé sur la chaussée » 37

Tableau 1 : Récapitulatif des principales contributions cumulées commune par commune 38

Graphique 15 : Exemple de *clustering* basé sur les quatre premières composantes d'une ACP et projection sur les deux premiers axes 45

Tableau A1 : Effectifs observés 54

Tableau A2 : Exemple simplifié d'un passage de variables catégorielles initiales à un tableau disjonctif complet 60

Carte A1 : Représentation des contributions ACM cumulées – Accidents mortels – Modalité « en agglomération » 71

Carte A2 : Représentation des contributions ACM cumulées – Accidents mortels – Modalité « hors agglomération » 71

Carte A3 : Représentation des contributions ACM cumulées – Accidents mortels – Modalité « aube, crépuscule » 72

Carte A4 : Représentation des contributions ACM cumulées – Accidents mortels – Modalité « nuit, sans éclairage public » 72

Carte A5 : Représentation des contributions ACM cumulées – Accidents mortels – Modalité « avec un piéton » 73

Carte A6 : Représentation des contributions ACM cumulées – Accidents mortels – Modalité « contre un obstacle situé sur la chaussée » 73

REMERCIEMENTS

Pour cette première publication à l'IWEPS, je tiens à remercier Frédéric Vesentini de m'avoir en premier lieu orienté vers cette thématique de la sécurité routière et des causes supposées d'accidentalité sur les chaussées wallonnes. Je le remercie aussi pour ses conseils et sa relecture.

Je remercie Yvan Casteels et Nathalie Focant (AWSR) pour les échanges d'idées et le précieux partage de données.

Je remercie Julien Charlier pour son expertise et sa grande aide lorsqu'il a été question d'obtenir de la meilleure des façons les objets cartographiques de ce travail.

Je remercie Julien Juprelle, pour sa relecture et ses idées qui annoncent sans doute des futurs travaux prometteurs autour de la thématique de la mobilité. Toujours pour la relecture, je remercie également Dominique Fasbender et Sébastien Brunet.

Enfin, pour de précieuses orientations techniques et pratiques, et pour la mise en page, un merci aussi à Évelyne Istace et Aurélie Hendrickx.

CHAPITRE 1

INTRODUCTION

La thématique de la sécurité routière et des accidents de la route demeure un phénomène de grande ampleur pour notre société par la diversité de ses impacts sociaux, économiques, voire même environnementaux. Ces accidents touchent notamment à la question de la voiture dans notre vie quotidienne et à l'importance toujours accrue de la mobilité des personnes (voir Masuy *et al.*, 2018, La Rocca et Masuy, 2019, Brunet *et al.*, 2020).

Lorsqu'un accident de la route est constaté, il est toujours important d'en comprendre les caractéristiques, les circonstances et les explications éventuelles. Comprendre ce qu'il s'est passé est une étape indispensable à la mise en place d'une politique d'évitement ou de réduction du risque, et contribue ainsi à mieux éviter que les mêmes scénarios ne se reproduisent. Pour cela, chaque accident doit être catalogué et référencé. Ce type de référencement va inévitablement provoquer un flux d'informations[1] qui prendra le plus souvent la forme de variables catégorielles dans un jeu de données. Ces variables catégorielles ou facteurs (ex. : type de route) impliquent chacune plusieurs niveaux ou modalités (ex. : autoroute, route régionale, route communale).

Partant de ces constats, et après avoir détaillé les bases de données à notre disposition (cf. chapitre 2, sources : Statbel et AWSR), cette publication va tenter d'atteindre deux objectifs. Tout d'abord, une partie de ce travail, très générale et méthodologique, est dédiée à la compréhension des outils et méthodes statistiques propres à l'analyse de variables catégorielles. En effet, nombre de modélisations ne sont pas adaptées à de telles variables et un rappel théorique peut s'avérer nécessaire. Plusieurs niveaux d'étude seront ainsi parcourus, allant d'un simple test d'indépendance entre deux variables à des analyses multivariées du type Analyse factorielle des Correspondances Multiples (ACM), en passant par des analyses univariées du type Analyse Factorielle des Correspondances Simples (AFC). Les résultats numériques et les sorties graphiques issus de ces modèles seront montrés et interprétés. Cette partie davantage théorique est reportée en annexe afin de permettre au lecteur de s'y référer à son aise, sans pour autant alourdir le corps du texte principal.

Ces analyses vont être directement appliquées au contexte des accidents de la route survenus dans les communes wallonnes entre 2009 et 2018 (cf. chapitre 3). Les accidents mortels et ceux ayant entraîné des blessures corporelles sont pris en compte, le plus souvent séparément. Le deuxième objectif sera alors de faire parler les données, de comprendre comment fonctionne chacun des facteurs structurels d'occurrence des accidents de la circulation (cf. section 3.2) et de proposer des analyses multivariées pouvant se résumer en des cartes géographiques facilement interprétables (cf. section 3.3).

Pour chaque analyse, pour chaque facteur, des communes seront mises en avant, avec des caractéristiques propres. Ces interprétations et ces différents constats pourront par la suite servir à la mise en œuvre de leviers d'action éventuels pour, une nouvelle fois, éviter que certains risques ne se répètent. L'exercice reste cependant difficile dans la mesure où des situations ou des réalités diverses (partie urbaine et rurale au sein d'une même commune, etc.) peuvent agir simultanément sur les facteurs qui en découlent.

[1] Les variables retenues n'apparaissent jamais de manière spontanée dans un jeu de données. Elles sont le fruit d'un processus de sélection qui mêle administratif, technique, politique, limites technologiques…

CHAPITRE 2

DONNÉES DISPONIBLES

Les modèles d'Analyse Factorielle des Correspondances simples (AFC) et d'Analyse factorielle des Correspondances Multiples (ACM) détaillés dans la section suivante et en annexe 1 seront par la suite directement appliqués sur des données réelles liées aux **causes supposées des accidents de la route dans les communes wallonnes entre 2009 et 2018**. Tout en gardant en commun la dimension territoriale (**niveau communal**, Wallonie) et la dimension temporelle (**entre 2009 et 2018**), deux jeux de données différents sont utilisés pour des finalités elles-mêmes différentes.

2.1. TABLEAUX DE CONTINGENCE INDÉPENDANTS SUR DES FACTEURS POUVANT EXPLIQUER LES ACCIDENTS DE LA ROUTE

Ce premier jeu de données a été extrait et fourni par l'Agence wallonne pour la sécurité routière (AWSR)[2]. On y retrouve les 262 communes wallonnes qui jouent naturellement le rôle des individus (en lignes). En colonne, les niveaux de différentes variables catégorielles sont disponibles. Ces informations sont généralement récoltées par la police sur le lieu de l'accident (à partir du procès-verbal rédigé par la police lors de tout accident sur voies publiques ayant entraîné des lésions corporelles).

Chaque case du tableau reprend un effectif observé cumulant les informations allant de 2009 à 2018. L'organisation de ce jeu de données renvoie à une juxtaposition de différents tableaux de contingence (cf. section A de l'annexe 1), de façon totalement indépendante entre eux. Cette organisation structurelle justifiera une préférence méthodologique allant vers l'utilisation d'AFC (cf. section B de l'annexe 1).

Les variables finalement retenues sont les suivantes (voir aussi les fiches métadonnées Statbel ou Walstat) :

- L'**état de la chaussée** (cinq niveaux initiaux : sec/propre, humide, verglas/neige (conditions hivernales), sale et non-renseigné) ;
- Le moment de l'accident ou **période** (cinq niveaux initiaux : jour dans la semaine, nuit dans la semaine, jour dans le week-end, nuit dans le week-end et non-renseigné. Le jour est défini comme la période comprise entre 6h et 21h59. La nuit est définie comme la période comprise entre 22h et 5h59) ;
- La **luminosité** (cinq niveaux initiaux : plein jour, aube/crépuscule, nuit avec éclairage public allumé, nuit sans éclairage ou éclairage éteint et non-renseigné) ;
- Le **type de route** (quatre niveaux initiaux : autoroute, route régionale, route communale, non-renseigné) ;
- La **vitesse maximale autorisée** sur le lieu de l'accident (six niveaux initiaux : 30km/h, 50km/h, 70km/h, 90km/h, 120km/h, non-renseigné) ;
- Le **type de collision** (neuf niveaux initiaux : en chaîne (trois conducteur(rice)s ou plus), frontale, par l'arrière, par le côté, avec un piéton, contre un obstacle situé sur la chaussée, contre un obstacle situé hors de la chaussée, un seul usager sans obstacle, non-renseigné) ;
- Certaines **conditions particulières de circulation** (sept niveaux initiaux : mauvais état de la route (ornières, inondations…), signalisation défectueuse, éclairage défectueux ou insuffisant, encombrement de la circulation (file, accident préalable…), forte descente (7% de pente ou plus), virage serré, entrave à la visibilité).

Les six premières variables catégorielles ci-dessus se conforment parfaitement à ce que l'on attend d'un tableau de contingence : si l'on ajoute les effectifs de toutes les modalités pour chaque variable, nous retrouvons à chaque fois le nombre total d'accidents observés dans les communes entre 2009 et 2018. Ceci n'est pas le cas pour la septième et dernière variable traitant des conditions particulières de circulation.

Afin de rechercher de l'information utile, tous les niveaux non-renseignés ont été supprimés de l'analyse.

D'autres variables ont été retirées. Il s'agit par exemple du nombre de conducteur(rice)s impliqués dans un accident qui sont sous influence de l'alcool ou testés positifs à un test d'alcoolémie. Cette variable semblait sensible dans le sens où tous les conducteur(rice)s (ou piétons) ne sont pas systématiquement soumis à un test d'alcoolémie.

[2] Et plus particulièrement par Yvan Casteels et Nathalie Focant, qui sont ici grandement remerciés.

Enfin, le but de cet article, au-delà de faire un point méthodologique sur comment traiter les variables catégorielles, est de pouvoir mettre en lumière des causes structurelles aux accidents de la route. Et, éventuellement, de pouvoir déterminer des leviers d'action pertinents au niveau local. Ainsi, les facteurs directement liés à des erreurs humaines de conduite (feu rouge grillé, franchissement de ligne blanche, non-respect de la priorité, non-respect de la distance de sécurité…) ou directement liés à des défauts techniques du véhicule impliqué (pneus lisses, éclairage déficient…) semblent moins pertinents et ont été également retirés.

Notons aussi une rupture dans la collecte des données signifiée par Statbel[3] concernant la variable type de collision : « jusque novembre 2017, la police enregistrait une collision en chaîne à partir de quatre véhicules impliqués. Depuis, une collision en chaîne est enregistrée à partir de trois véhicules impliqués ».

2.2. DONNÉES MULTIVARIÉES, PROFILS ÉLARGIS DES ACCIDENTS DE LA ROUTE

Dans ce deuxième cas, les données ont été recueillies directement sur Statbel (be.STAT) avec une finalité différente (finalité davantage policière pour un établissement des responsabilités) par rapport au jeu de données précédent. Au lieu de juxtaposer différents tableaux de contingence qui prennent en compte un seul facteur catégoriel à la fois, l'idée est de pouvoir cumuler l'information issue de **plusieurs facteurs simultanément** et, ainsi, de pouvoir créer un **profil pour chaque accident ayant eu lieu dans une commune wallonne entre 2009 et 2018**. Cette organisation structurelle justifiera ainsi une préférence méthodologique allant vers l'utilisation d'ACM (cf. section C de l'annexe 1).

En conséquence, la matrice des données ne comporte plus les 262 communes wallonnes en ligne. Elle comportera une ligne par accident corporel constaté, ce qui a pour conséquence une explosion de la dimensionnalité de l'objet matriciel sur lequel nous travaillerons. Pour amortir cette explosion, et pour essayer d'obtenir davantage de possibilités d'analyse, les plus stables et fiables possible, il a été décidé de séparer les accidents corporels mortels (selon l'encodage « personnes décédées sur place ou endéans les 30 jours ») et les accidents corporels ayant causé des blessés (graves ou légers).

Ce caractère cumulatif exigé ici provoque une baisse de la quantité des variables simultanément disponibles pour tous les accidents. Pour chaque accident corporel ayant eu lieu dans une commune wallonne entre 2009 et 2018, mortel ou ayant causé des blessés, les variables suivantes sont au final retenues :

- La **luminosité** (cinq niveaux initiaux : plein jour, aube/crépuscule, nuit avec éclairage public allumé, nuit sans éclairage ou éclairage éteint et non-renseigné) ;
- Le **type de route** (trois niveaux initiaux : autoroute, route régionale ou communale, non-renseigné) ;
- L'occurrence de l'accident en **agglomération** ou pas (trois niveaux initiaux : en agglomération, hors agglomération, non-renseigné) ;
- Le **type de collision** (neuf niveaux initiaux : en chaîne (trois conducteur(rice)s ou plus), frontale, par l'arrière, par le côté, avec un piéton, contre un obstacle situé sur la chaussée, contre un obstacle situé hors de la chaussée, un seul usager sans obstacle, non-renseigné).

En retirant les niveaux non-renseignés et en veillant à ne garder que des profils entièrement complets, la matrice des accidents mortels a pour dimensions (3 456 × 5). Notons qu'il n'y a que cinq colonnes, le nom de la commune et les quatre variables ci-dessus, dans la mesure où la transformation des variables catégorielles en objets binaires pour la construction du tableau disjonctif complet (cf. tous les détails dans la section C de l'annexe 1 et plus particulièrement le tableau A2), puis de la table de Burt, a lieu implicitement lors de la modélisation de l'ACM. La matrice des accidents ayant causé des blessés graves ou légers a quant à elle pour dimensions (120 207 × 5).

[3] https://bestat.statbel.fgov.be/bestat/crosstable.xhtml

CHAPITRE 3

ANALYSES NUMÉRIQUES

Ce chapitre vise à présenter les principaux résultats des analyses statistiques sur les causes enregistrées des accidents de la route en Wallonie entre 2009 et 2018.

Dans un premier temps (cf. section 3.1), une brève introduction aux modèles factoriels d'analyse des correspondances est proposée. L'idée est de proposer au lecteur une compréhension intuitive de ces méthodes et des principes sous-jacents. Pour plus de détails théoriques, l'**annexe 1** de cette publication contient un petit *vade-mecum* sur la mise en pratique d'une AFC (Analyse Factorielle des Correspondances simples) et d'une ACM (Analyse factorielle des Correspondances Multiples)

Ensuite, dans la section 3.2, les analyses se porteront précisément sur des AFC simples sur la base du premier jeu de données. L'effet des facteurs disponibles, un par un, sera étudié et les communes particulièrement impactées par tel ou tel facteur seront mises en évidence via des représentations graphiques.

Dans un troisième temps (cf. section 3.3), les principaux résultats des ACM appliquées sur des profils complets d'accidents de la route en Wallonie mettront en évidence des interactions, associations, entre plusieurs facteurs. Des représentations cartographiques de la Wallonie mettront en évidence les communes particulièrement impactées, de façon relative, par telle ou telle modalité (cf. section 3.3.2).

Il est important de bien noter que les données sont au préalable centrées et réduites[4] avant toute analyse, ceci afin de contrôler les effets d'échelle et les disparités liés aux tailles des différentes communes (que ce soit en superficie, en nombre d'habitants, en nombre de kilomètres d'autoroute ou de voiries en général, etc.).

Enfin, dans la section 3.3.5, un prolongement des Analyses des Correspondances Multiples sera évoqué avec une ébauche de *clustering* des communes, afin d'appréhender la question d'un regroupement de communes sur la base de l'étude des accidents de la route qui y ont eu lieu entre 2009 et 2018.

3.1. INTRODUCTION INTUITIVE AUX ANALYSES FACTORIELLES DES CORRESPONDANCES

« L'histoire des mathématiques et celle des connaissances humaines ont évolué en parallèle au fil du temps. À chaque étape de l'histoire de l'humanité, les progrès scientifiques et technologiques auraient été impossibles sans l'avancée préalable des données mathématiques sur lesquelles ils étaient fondés. » Francesco Martin Casalderrey, *La mystification des sens*, RBA (2019).

À cet égard, parmi l'arsenal des modèles statistiques multivariés, la technique de l'Analyse Factorielle des Correspondances (AFC) a été particulièrement utile à l'évolution des sciences humaines, avec une mission d'exploitation initiale de bases de données d'ampleur.

Félicité des Nétumières[5] disait à cet égard :

« [L']objectif premier [de l'analyse factorielle] est de permettre au chercheur d'appréhender le plus simplement possible la masse de données dont il dispose dans son enquête et d'en extraire les informations pertinentes. Pour reprendre une expression consacrée, l'analyse factorielle est « un aveu d'ignorance », un « radar tourné vers le brouillard » qui « sert avant tout à dépeindre à grands traits les dimensions les plus importantes d'une variation dans un nouveau champ de recherche ». C'est donc, en premier lieu, lors de la phase exploratoire des données, que l'analyse factorielle se présente comme un outil particulièrement utile ».

Comme son nom l'indique, l'AFC est une analyse factorielle, ou en axes principaux. Ce type de technique s'accommode sans problème de milliers voire de millions d'observations potentielles. Tels des orpailleurs devant brasser des tonnes de terre pour trouver LA pépite, les *dataminers* vont utiliser ces techniques pour tamiser les données et en ex-

[4] Technique qui consiste à diviser chaque observation par la moyenne correspondante et à diviser le résultat par l'écart-type de la série.
[5] « Méthode de régression et analyse factorielle », Histoire et Mesure 1997, pp.271-298.

traire l'information principale, ou le signal. Précisons tout de même que ces techniques ne sont pour la plupart que descriptives, dans la mesure où l'explication des liaisons ou leurs causalités ne sont pas immédiates.

Le point de départ d'une telle analyse des données se présente toujours de la même façon : un tableau d'observations. Celui-ci comporte autant de lignes qu'il y a d'individus (ou unités statistiques). En colonnes figurent les valeurs des variables aléatoires prises par ces unités statistiques. Les variables sont bien souvent quantitatives, auquel cas le tableau comprend autant de colonnes qu'il existe de variables observées. C'est un peu plus compliqué lorsque les critères sont qualitatifs ou catégoriels, comme dans le cadre de ce travail. L'organisation des données s'en trouve alors souvent changée (cf. chapitre 2 sur les données disponibles).

Les analyses factorielles permettent de détecter des proximités entre variables, entre individus et entre variables × individus, mettant à jour des liens d'attraction ou au contraire des « répulsions ». Le terme « correspondances » signifie que sont étudiées les liaisons entre deux variables nominales, par opposition à la corrélation qui s'intéresse aux liens entre variables numériques. Les deux buts principaux d'une analyse factorielle sont ainsi, d'une part, de détecter ces proximités et, d'autre part, de faire apparaître un certain nombre de variables synthétiques abstraites décorrélées entre elles (dans le but d'une réduction de dimensionnalité). On parle alors d'axes principaux ou de dimensions principales. Ce second attrait ne constitue toutefois pas le but majeur de l'AFC.

Les analyses factorielles conduisent aussi à identifier facilement des valeurs aberrantes, qu'il est alors possible d'éliminer pour reconduire l'analyse. Enfin, et surtout, elles permettent de hiérarchiser l'importance des critères ou des facteurs éventuellement explicatifs.

Dès le départ, nous disposons donc d'un nombre important d'observations. Partons dans un premier temps du cas où elles sont caractérisées par deux variables. Ces observations sont synthétisées dans un tableau de contingence dans lequel chaque case figure le nombre d'unités statistiques qui présentent telle modalité d'une première variable placée en ligne et telle modalité d'une autre variable présentée en colonne. Quand il y a peu de modalités, un simple coup d'œil pourrait suffire pour estimer s'il y a liaison ou pas, et une brève présentation de pourcentages suffirait. Afin d'évaluer plus précisément s'il existe des liens, il faut cependant procéder à un test statistique d'indépendance, qui permet d'objectiver et de mesurer le constat.

En revanche, lorsque les modalités sont plus nombreuses (types d'articles en magasin, points de vente d'une grande enseigne, tranches de dix minutes dans la journée…), un dispositif plus adapté s'avère nécessaire. C'est ainsi que l'on passe des statistiques descriptives à l'analyse factorielle des données. Que faire de toutes ces données ? Les passer aux rayons X pour découvrir l'invisible, c'est-à-dire les liens qui existent au sein d'une pléthore de modalités… Supposons que les deux variables soient des produits alimentaires et des points de vente franchisés : peut-être découvrirez-vous que les magasins en zones rurales sont davantage associés aux consommations de pâté, que les fromages à pâte molle sont moins appréciés dans telle région, que telle région est plutôt « bière » et telle autre plutôt « vin », etc. Et vous adapterez l'effort publicitaire en évitant les dépenses inutiles. Toutes ces informations sont envisageables en une seule étude.

Contrairement à l'ACP (Analyse en Composantes Principales), on ne peut pas visualiser le nuage représentatif des individus à l'aide d'un système de coordonnées cartésiennes puisque la population est définie par des critères nominaux. Mais comme toute analyse factorielle qui se respecte, l'analyse des correspondances permet de visualiser des liaisons sur un ou deux plans factoriels puisque l'on va définir une métrique adéquate.

Il est possible de croiser plus de deux variables simultanément mais il faut alors utiliser la procédure de l'ACM (Analyse des Correspondances Multiples) et non l'AFC.

Pour une vision élargie, davantage théorique et mathématique, le lecteur est invité à se référer à l'**annexe 1**, dans laquelle toutes les considérations statistiques liées aux variables catégorielles sont

détaillées, du test d'indépendance entre deux variables aux manipulations matricielles liées à une ACM.

La section 3.2 et les suivantes s'orientent directement vers l'application de ces modèles aux données qui nous intéressent : les causes enregistrées des accidents de la route en Wallonie entre 2009 et 2018 (selon le prisme des bases de données présentées au chapitre 2).

3.2. SORTIES DES AFC SUR LES MODALITÉS DES ACCIDENTS DE LA ROUTE EN WALLONIE (PÉRIODE 2009-2018)

Les différents facteurs de causes supposées d'accidents de la route sont ici analysés, un par un, via les tableaux de contingences décrits dans la section A de l'Annexe 1. Les AFC mises en œuvre fournissent les résultats suivants.

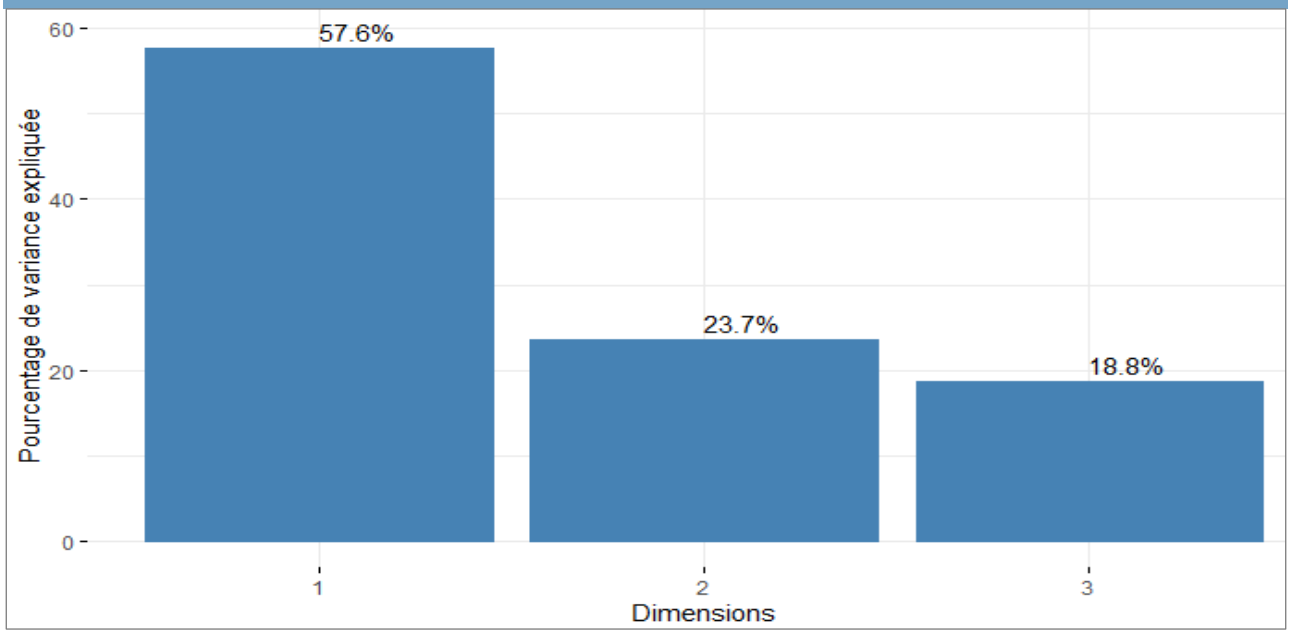
3.2.1 Facteur « état de la chaussée »

Pour ce facteur, qui a quatre niveaux renseignés (chaussée propre, sale, humide et verglas/neige), l'analyse des valeurs propres obtenues par l'AFC permet de connaître le pourcentage de variance expliquée par le modèle (cf. graphique 1) et l'importance, ainsi que le nombre, de dimensions principales à retenir. Ainsi, trois dimensions suffisent naturellement pour capter l'intégralité de l'information. Les deux premiers axes permettent de capter 81,3% de l'information à eux seuls.

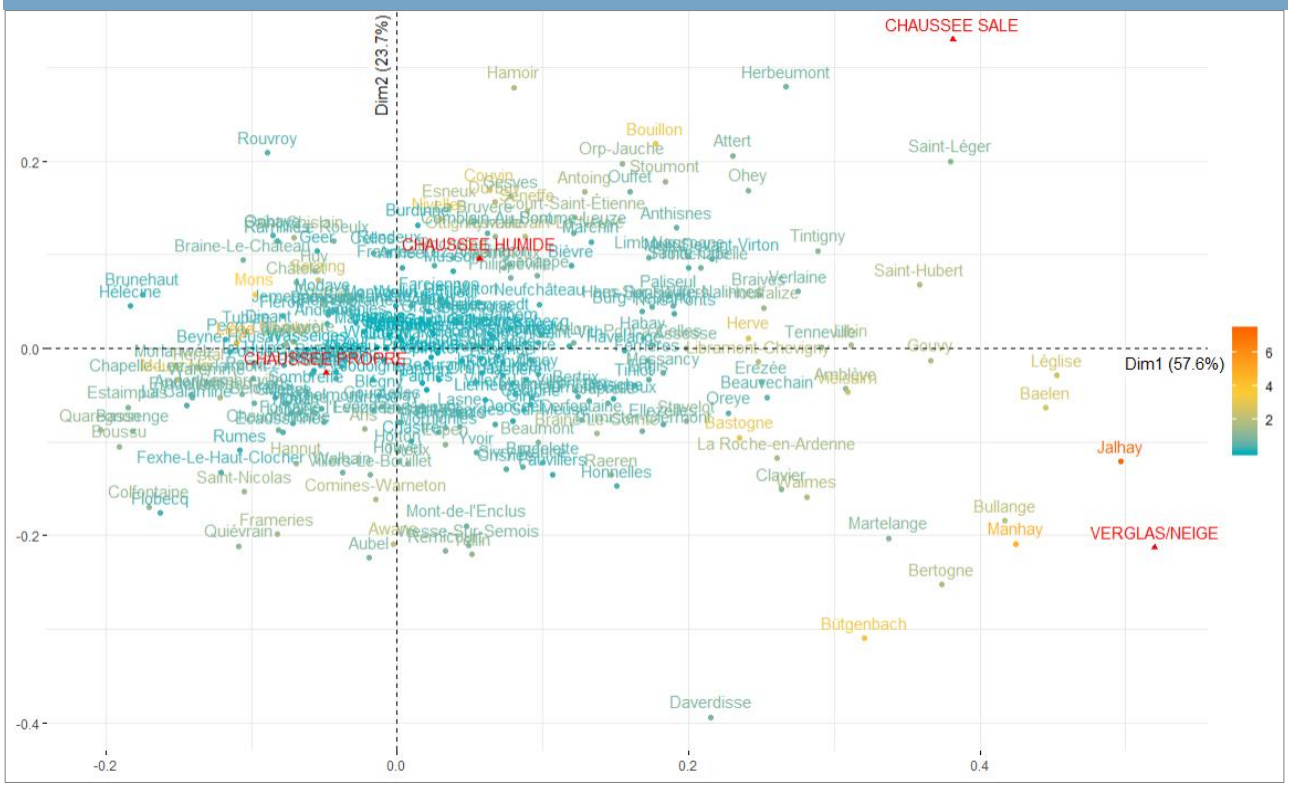
Le graphique 2 montre ensuite la projection des profils ligne (les communes wallonnes) et des profils colonne (les modalités de la variable considérée) dans le **plan formé par les deux premiers axes principaux de l'AFC**. Pour rappel, ce graphique est construit à partir de l'axe 1 non trivial, repris en abscisse, et l'axe 2 en ordonnée. Ce graphique, comme les suivants, colore également les communes par rapport à leurs contributions (en pourcentage cumulé) dans ces deux premiers axes principaux. Un exemple de tableau numérique complet (coordonnées, cosinus carrés, contributions) est visible en annexe 2.

La modalité « chaussée propre » est très proche de l'origine, ce qui est normal dans le sens où ce n'est pas un facteur expliquant un accident de la route. La modalité « chaussée humide » est peu éloignée de l'origine. Ceci est également logique. La pluie étant un facteur assez également distribué sur le territoire wallon, elle n'apparaîtra pas comme un facteur fortement discriminant. Les deux autres modalités, par contre, permettent d'extraire de l'information pertinente et d'interpréter des associations positives.

Graphique 1 : Dimensions de l'AFC et variance expliquée (variable « état de la chaussée »)



Graphique 2 : Carte des profils ligne et colonne dans le plan formé par les deux premiers axes principaux de l'AFC (variable « état de la chaussée »)



Tout d'abord, on peut montrer une association forte positive (une surreprésentation) entre la modalité « chaussée sale » et les communes de Herbeumont et de Saint-Léger (voir aussi Attert et Ohey). Cela veut dire que, entre 2009 et 2018, il y a eu une proportion anormalement élevée (comparativement à ailleurs, effet saisonnier des récoltes agricoles ?) d'accidents de la route sur une chaussée jugée sale par la police lors de l'encodage de ces accidents.

Enfin, la modalité « verglas/neige », liée directement à des conditions hivernales difficiles, a joué un rôle important dans l'analyse puisqu'elle a grandement étiré le nuage de points vers la droite (elle a étiré le premier axe principal, qui explique à lui seul 57,6% de l'information). Ainsi, les communes qui se rapprochent le plus de cette modalité sont concernées par une seconde association forte positive. Il s'agit des communes de Jalhay (en tout premier lieu, Jalhay étant aussi la commune ayant les plus grandes contributions), Manhay, Bullange, Bertogne, Martelange, Baelen et Léglise. Sans surprise, il s'agit de communes proches des Hautes-Fagnes ou de communes de la province du Luxembourg. Cependant, si certaines communes ressortent très fortement (Jalhay, Manhay et Bullange) par rapport à d'autres, cela signifie que le traitement des routes enneigées ou verglacées pourrait y être sans doute questionné, afin de mieux comprendre ce qui pourrait être envisagé en termes de gestion du risque. Ici, il faut toutefois faire abstraction de potentielles configurations particulières très localisées (routes plus sinueuses et/ou pentues ?).

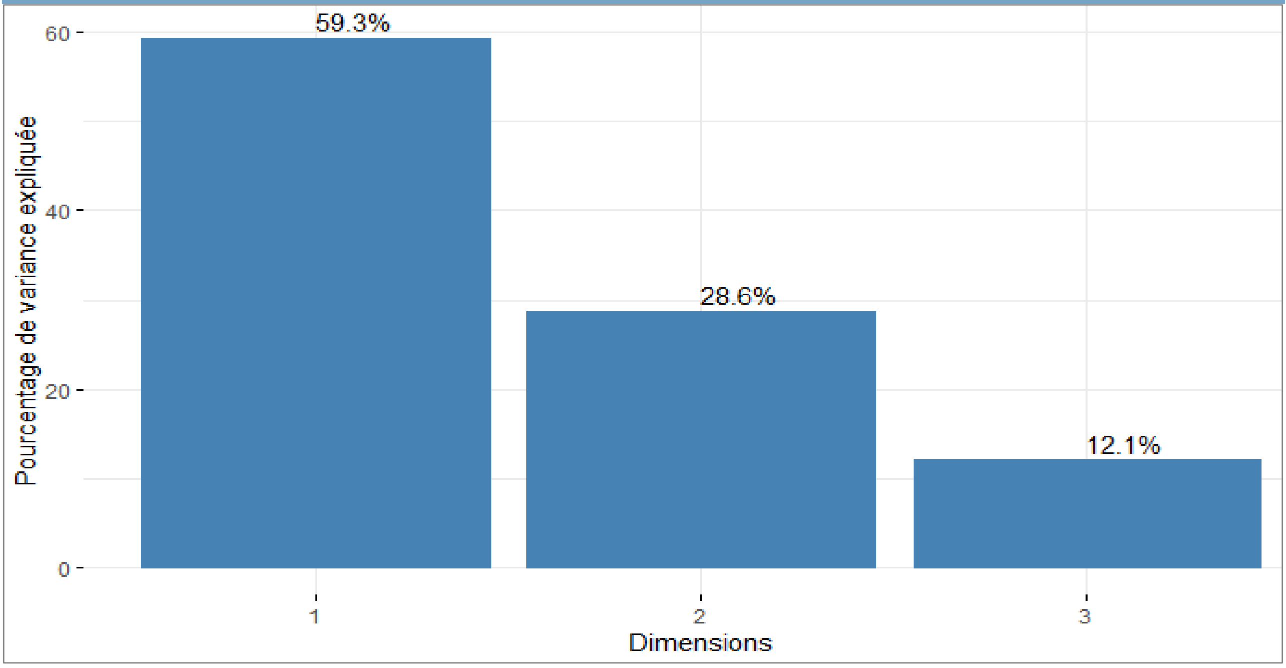
Si une analyse multivariée plus vaste était possible, en incluant des données complémentaires, d'autres conditions devraient être prises en compte pour affiner ces constats, comme la densité du trafic ou le respect des vitesses de circulation dans des conditions hivernales difficiles.

3.2.2 Facteur « période de la semaine »

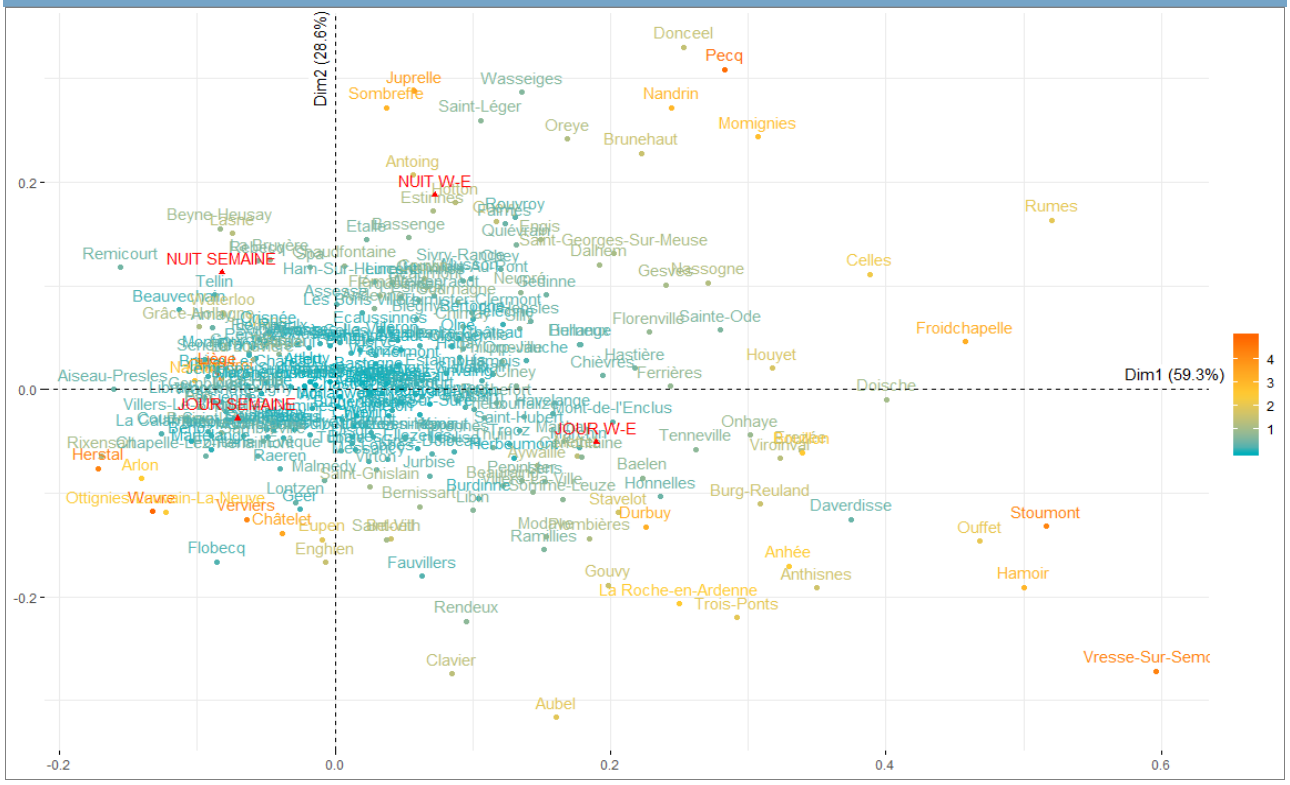
Cette variable comprend quatre niveaux renseignés (jour et nuit en semaine, jour et nuit lors du week-end).

Pour cette variable, les deux premières dimensions expliquent 87,9% de l'information (cf. graphique 3).

Graphique 3 : Dimensions de l'AFC et variance expliquée (variable « période de la semaine »)



Graphique 4 : Carte des profils ligne et colonne dans le plan formé par les deux premiers axes principaux de l'AFC (variable « période de la semaine »)



La carte des projections de l'AFC est visible sur le graphique 4. De par la disposition des modalités (en rouge), on y voit que l'axe 1 (59,3%) oppose la semaine au week-end et que l'axe 2 (28,6%) oppose la nuit au jour. Le nuage de points est très largement plus éparpillé dans sa partie droite (qui concerne les week-ends) et c'est également de ce côté que se trouvent les communes qui sont les plus éloignées de l'origine.

Par conséquent, les communes situées le plus à droite sont celles qui subissent de façon anormale un effet week-end important (une surreprésentation importante des accidents de la route qui ont lieu le week-end, de façon générale). Il s'agit surtout des communes de Vresse-sur-Semois, Stoumont, Rumes, Hamoir, Ouffet et Froidchapelle. Cela pourrait simplement signifier que ces communes sont largement désertées en semaine ou, par opposition, largement sujettes à un surinvestissement le week-end (sorties de motards par exemple).

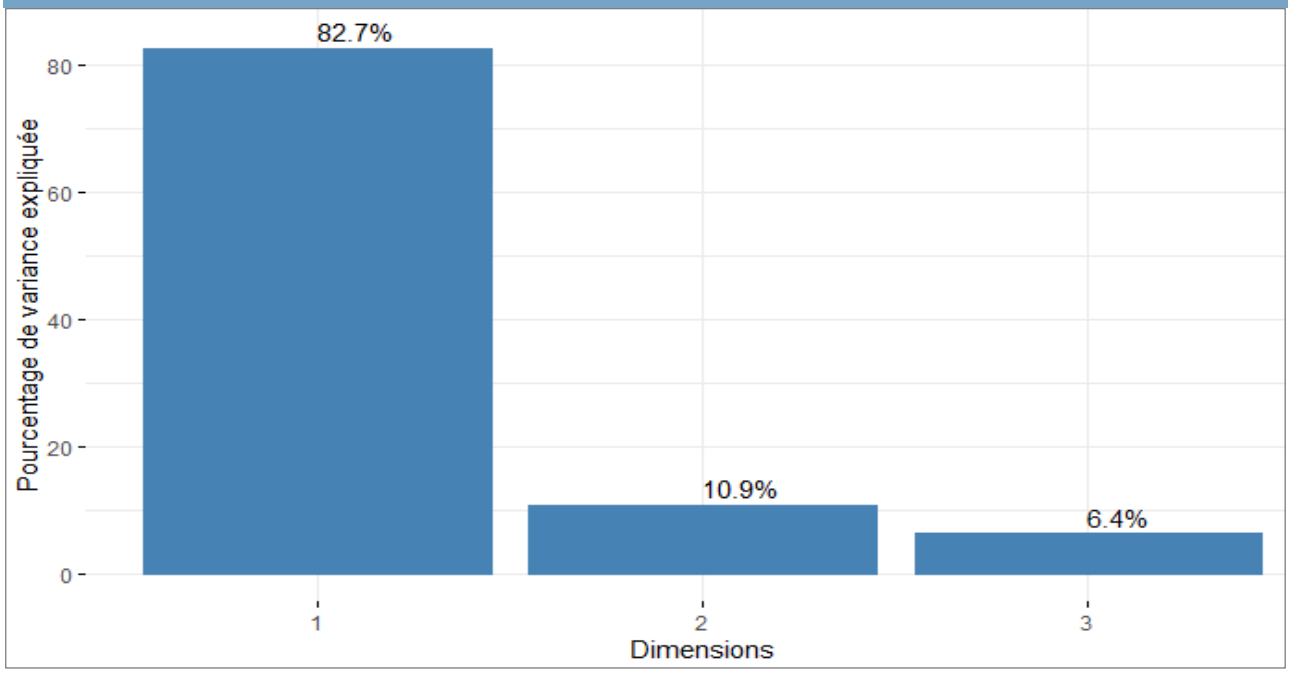
Un autre groupement de communes se situe en haut du graphique et concerne une dizaine de communes (Pecq, Donceel, Nandrin, Momignies, Brunehaut, Oreye, Wasseiges, Saint-Léger, Juprelle et Sombreffe). Ces dernières sont caractérisées par une surreprésentation significative des accidents de la route survenus la nuit lors des week-ends. Cela pourrait s'expliquer par le biais des activités festives, parfois avec des afflux frontaliers (discothèques, boîtes de nuit…).

Ici, ces résultats pourraient être approfondis afin de voir si des campagnes de préventions (autour de l'alcoolémie ou du respect de la vitesse autorisée selon le type de voirie) pourraient être renforcées ou améliorées pour éventuellement mieux canaliser les sorties nocturnes de fin de semaine dans ces communes.

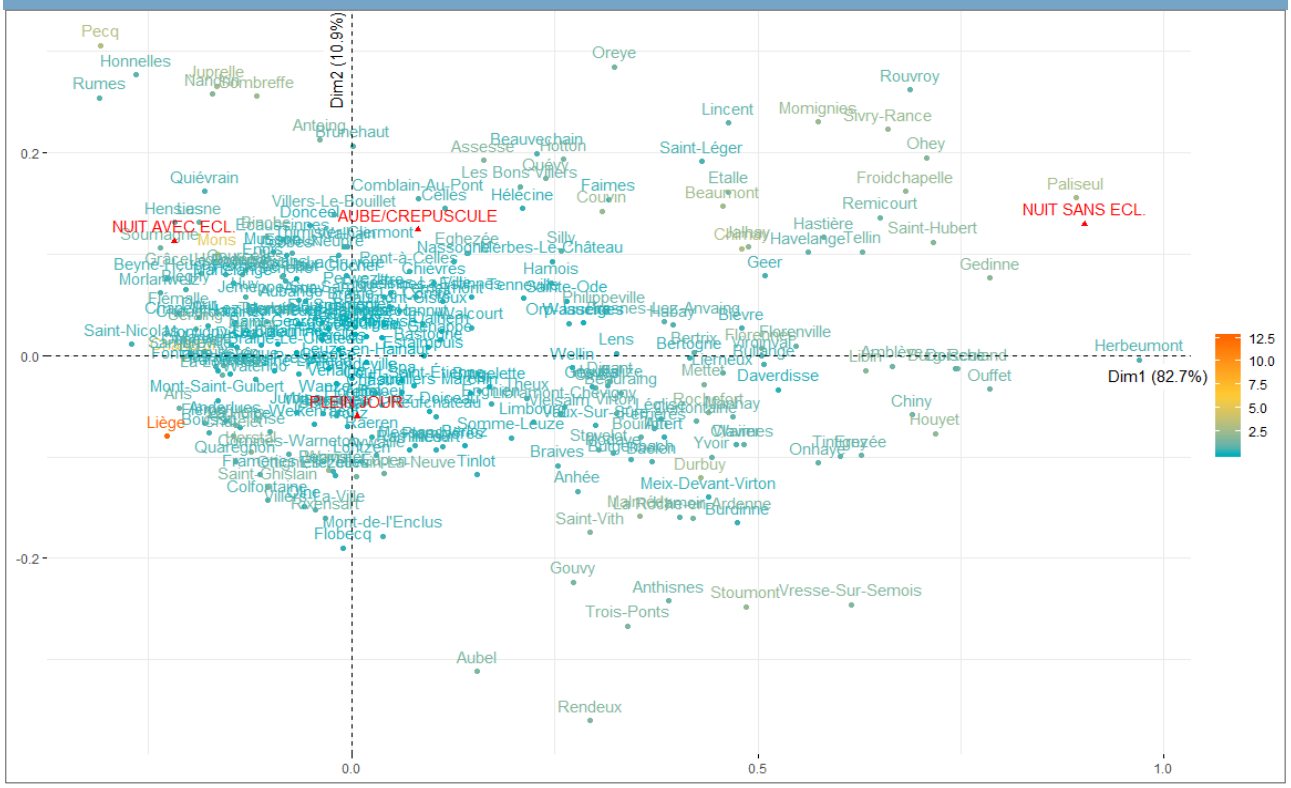
3.2.3 Facteur « luminosité »

Pour rappel, cette variable catégorielle comporte quatre niveaux renseignés. *A priori*, au moins un de ces niveaux, celui de plein jour, ne devrait pas être pertinent en tant que tel pour expliquer la survenue d'un accident. Pour cette variable, les deux premières dimensions parviennent à expliquer 93,6% de la variabilité totale, ce qui est conséquent (cf. graphique 5). La carte des profils ligne et colonne est montrée sur le graphique 6.

Graphique 5 : Dimensions de l'AFC et variance expliquée (variable « luminosité »)



Graphique 6 : Carte des profils ligne et colonne dans le plan formé par les deux premiers axes principaux de l'AFC (variable « luminosité »)



Comme on pouvait s'y attendre, la modalité « plein jour » est très proche de l'origine, et donc peu pertinente pour expliquer des discriminations au niveau de l'occurrence des accidents. Il en est de même avec la modalité « aube/crépuscule » qui ne permet pas vraiment d'élargir le nuage de points. De façon tout à fait intuitive, c'est le niveau « nuit sans éclairage » (pas d'éclairage du tout ou éclairage public éteint) qui est le plus explicatif, car il crée des surreprésentations (des associations fortes positives) dans plusieurs communes.

Ainsi, sur la base de l'historique des accidents ayant eu lieu entre 2009 et 2018, les communes projetées le plus à droite du graphique, près de cette dernière modalité, devraient sans doute envisager de questionner l'état de leurs signalisations (au sol ou panneaux), de questionner l'intérêt éventuel d'installer de nouveaux éclairages publics ou de les allumer plus souvent ou à meilleur escient. Il s'agit principalement des communes de : Herbeumont, Paliseul, Gedinne, Ouffet, Saint-Hubert et Ohey.

Il est intéressant également de regarder le petit groupe de communes qui se dégage en haut à gauche du graphique (Pecq, Honnelles, Rumes, Sombreffe, Juprelle et Nandrin). S'il l'on s'en tient à ce que l'AFC nous permet d'interpréter, ces communes semblent surreprésentées en lien avec la modalité la plus proche qui est « nuit avec éclairage ». Cette modalité n'est sans doute pas totalement pertinente en tant que telle, dans la mesure où elle ne permet pas réellement de proposer un levier d'action immédiat et précis. Par contre, rappelons-nous que la plupart de ces communes (Pecq, Juprelle, Sombreffe, Nandrin) étaient déjà apparues dans la section 3.2.2. Elles avaient été identifiées pour une surreprésentation significative des accidents de la route survenus la nuit lors des week-ends. Les deux modalités sont donc corrélées entre elles de façon cohérente. Cela montre aussi que d'autres pistes que l'éclairage doivent être prioritaires dans ces communes (prévention autour de l'alcool, limitations de vitesse localisées…).

Au-delà d'une apparente cohérence entre les deux analyses (qui sont pourtant, rappelons-le, indépendantes puisque les données ne sont pas croisées entre facteurs et sont issues de tableaux de contingence séparés), cela montre aussi que l'éclairage n'est pas un facteur principal pour les accidents qui ont lieu dans ces communes. On en revient alors aux campagnes de prévention mentionnées précédemment.

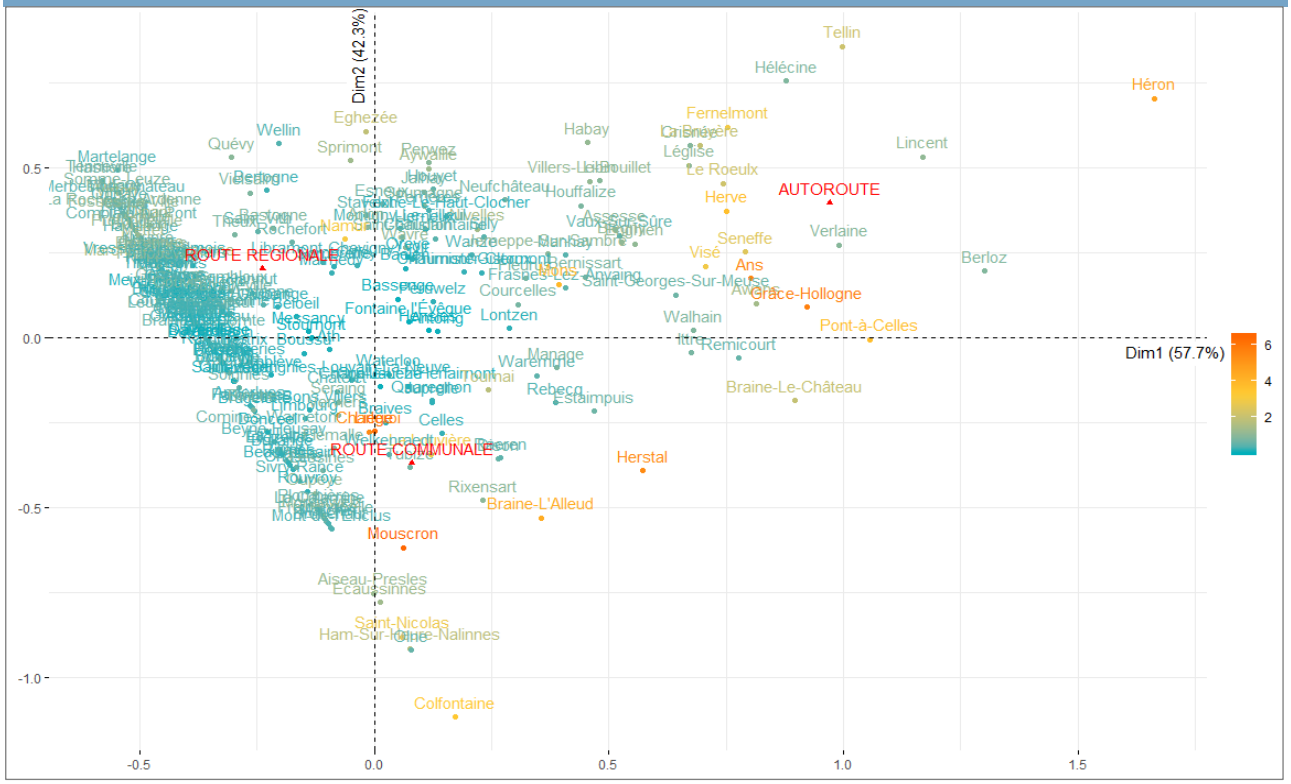
3.2.4 Facteur « type de route »

Ici, l'occurrence d'un accident est mise en lien avec le type de route de cet accident (autoroute, route régionale ou route communale). Il est évident que le réseau autoroutier ne concerne pas toutes les communes wallonnes et que, par conséquent, de nombreuses cases du tableau de contingence restent vides, car non observables. Un accident sur autoroute est en effet structurellement impossible dans la commune de Chimay ou à Comines-Warneton. Par convenance, ces vides sont ici remplacés par des zéros[6].

Cette variable catégorielle est entièrement captée par les deux premières dimensions principales de l'AFC, avec respectivement 57,7% et 42,3% de variance expliquée. De fait, le facteur ayant trois modalités, le plan généré par les deux premiers axes exprime forcément la totalité de l'information, sans aucune perte.

La carte des projections des communes et des trois modalités est reprise dans le graphique 7 ci-après.

Graphique 7 : Carte des profils ligne et colonne dans le plan formé par les deux premiers axes principaux de l'AFC (variable « type de route »)



[6] Pour pouvoir différencier sans perte d'information un « vrai zéro » (évènement observable mais non observé) d'un « faux zéro » (évènement non observable, ou impossible, encodé comme zéro), nous verrons comment faire dans la section 3.3.2, dédiée aux représentations multivariées.

Le fait d'avoir massivement remplacé des cas impossibles en zéros artificiels a créé un biais, qui se manifeste graphiquement par la diagonale sur la gauche. Ainsi, la modalité « autoroute » permet de largement étaler le nuage de points vers la droite. Les modalités « route régionale » et « route communale » jouent davantage sur la deuxième dimension, toutes choses égales par ailleurs.

Sachant que nous travaillons ici sur des données biaisées, avec toutes les précautions nécessaires, il apparaît toutefois deux sortes d'associations fortes positives synonymes de cas de surreprésentations :

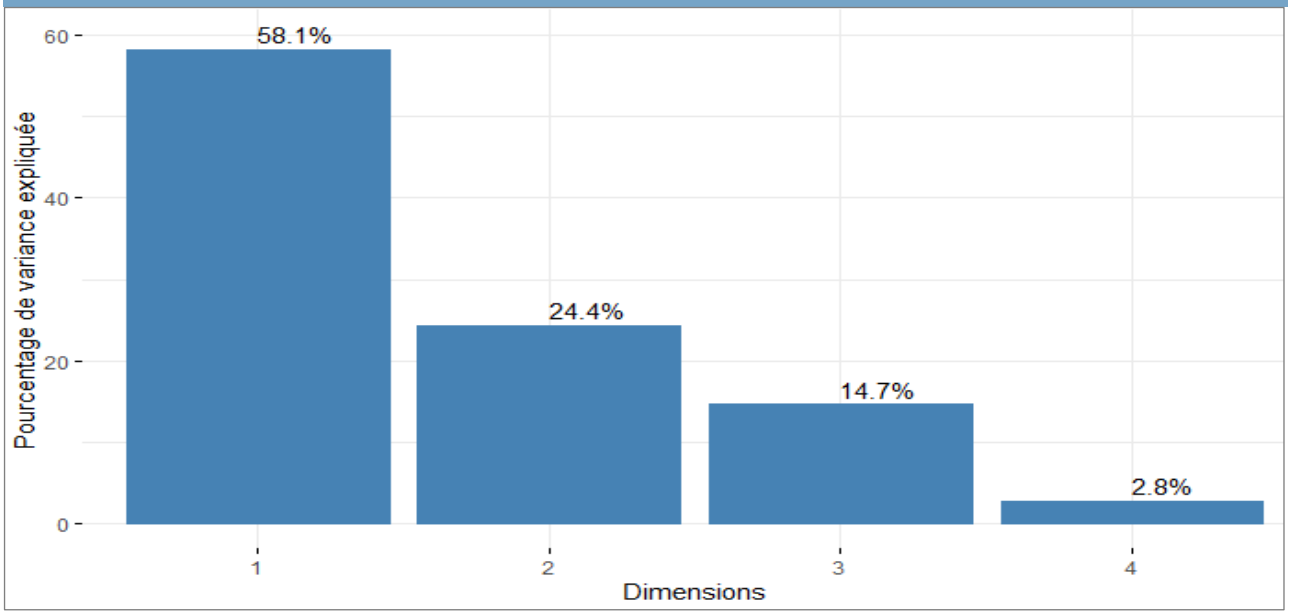
- Les communes situées le plus à droite sont surreprésentées par le nombre relatif d'accidents autoroutiers sur leur territoire (Héron, Lincent, Berloz, Tellin, Hélécine, Verlaine et Pont-à-Celles). Une attention particulière devrait donc être portée dans ces communes aux conditions de trafic sur autoroute (questionnements possibles sur la signalisation, radars…).
- Les communes projetées tout en bas sont surreprésentées par le nombre relatif d'accidents sur les routes communales de leur territoire (Colfontaine, Ham-sur-Heure-Nalinnes, Olne, Saint-Nicolas, Écaussinnes, Aiseau-Presles, Mouscron et Braine-L'Alleud). Les six premières communes citées se trouvent sur la diagonale créée par le biais expliqué plus tôt. Cela veut dire, vu qu'il n'y a pas d'autoroute sur ces six communes, que le nombre d'accidents sur route communale y prend significativement et largement le pas sur le nombre d'accidents sur route régionale. Et que, donc, un levier pourrait y être actionné précisément sur ce type de route.
- Pour Mouscron ou Braine-L'Alleud (et aussi Herstal dans une moindre mesure), les trois types de route sont en revanche présents. Ainsi, la prédominance relative des accidents sur route communale s'y constate comparativement aux accidents sur route régionale, mais aussi comparativement à ceux sur autoroute. La conclusion reste cependant identique avec des mesures potentielles à prendre au niveau du réseau communal.

3.2.5 Facteur « vitesse maximale autorisée »

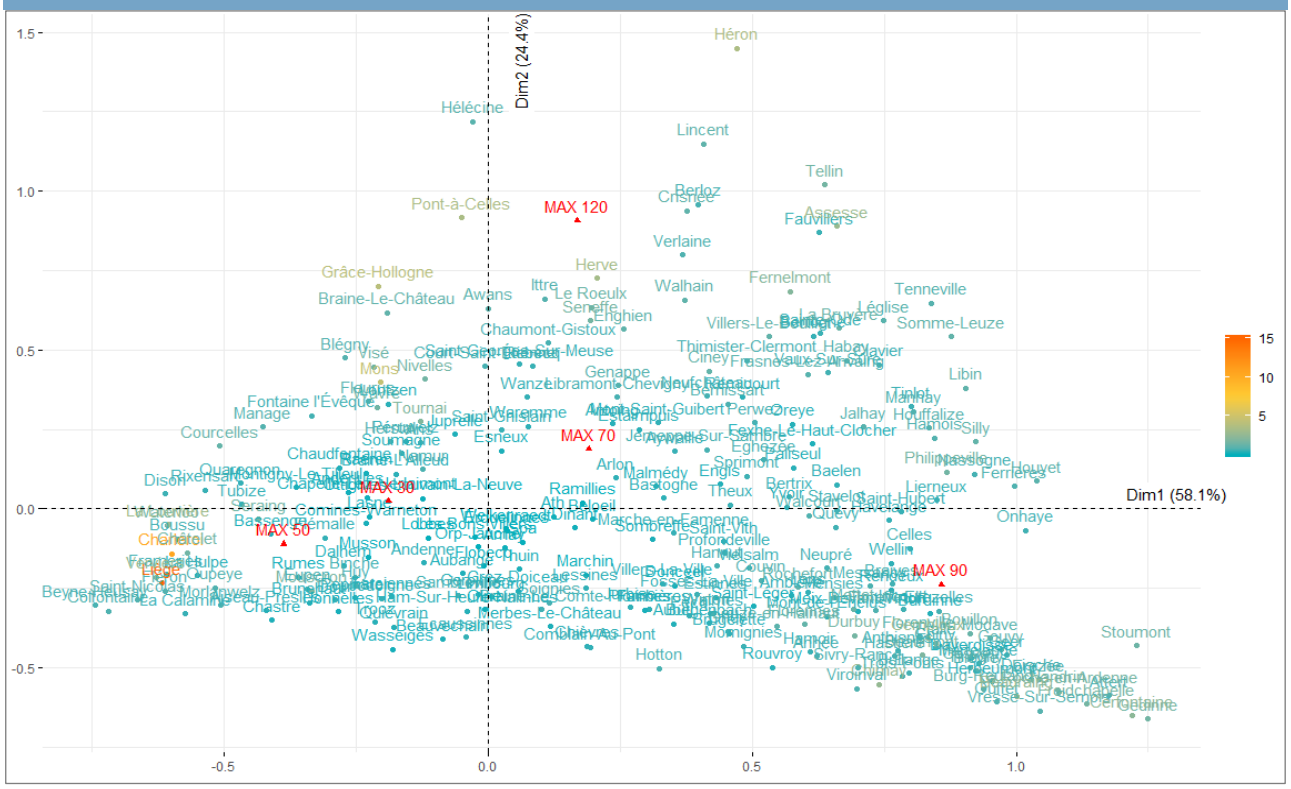
Cette variable comporte cinq niveaux renseignés : 30km/h, 50km/h, 70km/h, 90km/h et 120km/h. En cas d'intersection, une matrice permet de déterminer une vitesse unique sur le lieu de l'accident. Comme pour la variable précédente, un biais est observé dans le sens où toutes les communes wallonnes ne sont pas sujettes à l'existence de limitations de vitesse à 120km/h (directement en lien avec la présence de tronçons autoroutiers).

Pour ce facteur, quatre dimensions issues de l'AFC sont nécessaires pour capter 100% de l'information, et 82,5% de l'inertie est expliquée par les deux premiers axes (cf. graphique 8).

Graphique 8 : Dimensions de l'AFC et variance expliquée (variable « vitesse maximale autorisée »)



Graphique 9 : Carte des profils ligne et colonne dans le plan formé par les deux premiers axes principaux de l'AFC (variable « vitesse maximale autorisée »)



La carte des projections des communes et des modalités des limitations de vitesse (cf. graphique 9) est quelque peu plus difficile à lire que les précédentes dans le sens où le nuage de points est plus homogène. Remarquons que plus la vitesse autorisée augmente, plus la modalité correspondante est éloignée de l'origine et joue un rôle influent et discriminant dans le modèle. Ainsi, les limitations à 30, 50 et 70km/h ne permettent pas de tirer de conclusions pertinentes ou de mettre à jour des associations fortes (car trop proches de l'origine du graphique). La limitation à 90km/h permet de tirer le nuage de points vers la droite et joue un rôle important dans la première dimension. Cependant, même étiré, le nuage de point reste homogène dans cette direction et aucune commune ne se détache réellement des autres.

Enfin, la limitation à 120km/h permet de tirer le nuage de points vers le haut, jouant ainsi un rôle majeur dans la construction de la deuxième dimension principale. Ici, un groupe de communes apparaît, formant une association forte positive (une surreprésentation) avec cette limitation de vitesse : Héron, Hélécine, Lincent, Tellin, Berloz et Pont-à-Celles.

Sans réelle surprise, mais avec beaucoup de cohérence entre les modèles, il s'agit du même groupe de communes qui avait été mis en avant dans la section 3.2.4 via la modalité « autoroute ». Malgré une certaine redondance, ces résultats sont aussi confirmatoires et appuient le besoin d'une attention particulière aux conditions de trafic sur autoroute dans ces communes.

Pour aller plus loin dans cette analyse, il serait intéressant de mettre en parallèle la vitesse autorisée avec la densité de circulation sur les différents types de réseaux, la vitesse moyenne sur ceux-ci ou le rapport du nombre de kilomètres d'autoroute sur la commune par rapport aux voiries communales ou régionales (afin de créer des pondérations).

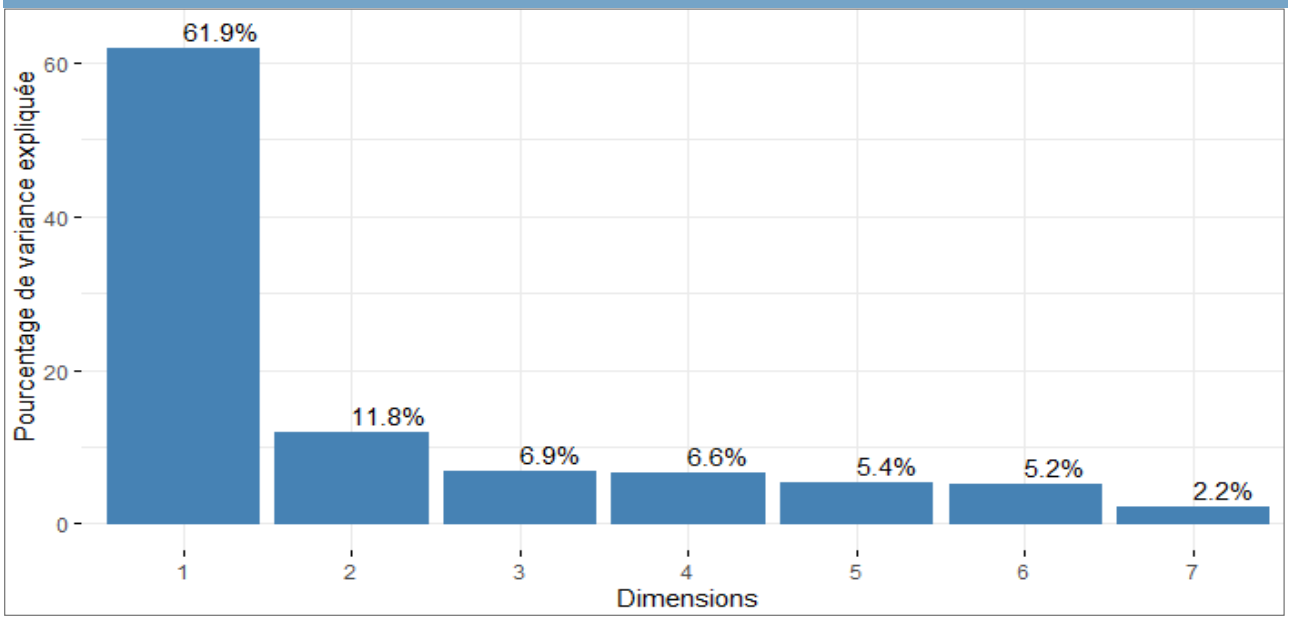
3.2.6 Facteur « type de collision »

Pour ce facteur, rappelons que huit niveaux renseignés différents sont repris selon les conditions de l'accident (en chaîne (trois conducteur(rice)s ou

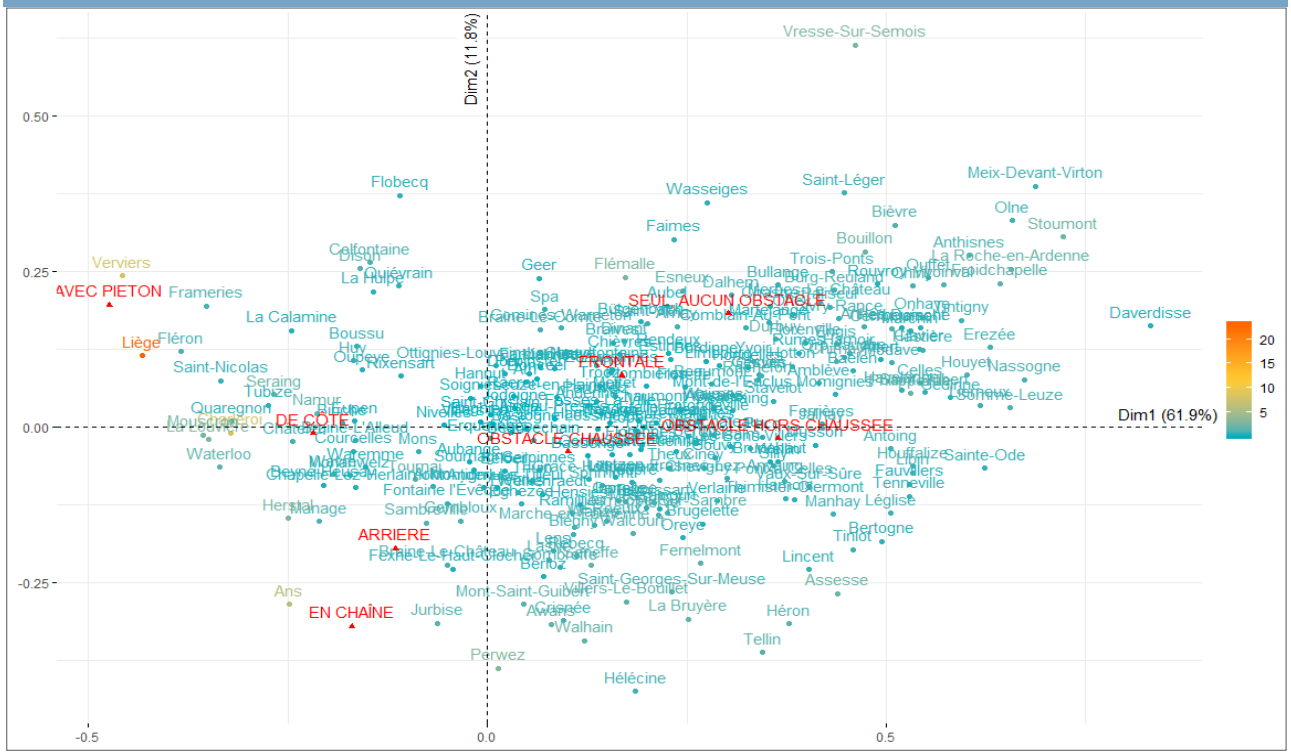
plus), frontale, par l'arrière, par le côté, avec un piéton, contre un obstacle situé sur la chaussée, contre un obstacle situé hors de la chaussée, un seul usager sans obstacle). Ceci va permettre d'obtenir un profil type de collision dans les communes wallonnes, sur la base de l'historique des accidents qui ont eu lieu entre 2009 et 2018.

L'AFC sur la base de cette variable révèle l'existence de sept dimensions pour expliquer 100% de l'information, les deux premières captant 73,7% de l'inertie totale (cf. graphique 10).

Graphique 10 : Dimensions de l'AFC et variance expliquée (variable « type de collision »)



Graphique 11 : Carte des profils ligne et colonne dans le plan formé par les deux premiers axes principaux de l'AFC (variable « type de collision »)



La carte des projections des profils ligne et colonne est représentée dans le graphique 11 et permet de mettre en évidence plusieurs associations fortes positives.

Les modalités « obstacle chaussée », « frontale » et « de côté » sont les moins discriminantes, car les plus proches de l'origine.

La modalité « obstacle hors chaussée » étire le nuage de points vers la droite, sans faire apparaître un sous-groupe évident de communes.

Les modalités « arrière » et « en chaîne » étirent le nuage de points vers le bas à gauche. Ces deux modalités sont proches entre elles, ce qui signifie qu'elles ont des individus en commun. Cela paraît logique : une collision en chaîne peut être souvent vue comme une succession de chocs par l'arrière. Ces deux modalités mettent en avant un petit groupe de communes pour lesquelles elles sont en surreprésentation : Ans, Jurbise, Perwez et Hélécine.

La modalité « avec piéton » étire le nuage de points dans le cadran supérieur à gauche, entraînant clairement dans son sillon un petit groupe distinct de communes : Liège (avec une très forte contribution cumulée dans les deux premières dimensions), Verviers, Fléron, Frameries et Saint-Nicolas. Ces communes pourraient donc questionner le bien-fondé de mieux signaler ou sécuriser leurs passages pour piétons ou de mener une campagne de prévention dans ce sens. Un effet « centre-ville » est aussi sans doute à questionner.

Enfin, la modalité « seul, aucun obstacle » étire le nuage de points vers le cadran supérieur à droite. Ce type d'accidents, avec un conducteur seul sur la route, est difficilement interprétable ici dans le sens où la cause supposée n'est pas directe via le facteur considéré, et peut prendre le chemin d'autres considérations non abordées dans ce travail (par exemple les problématiques liées à l'alcoolémie, la prise de drogues ou de médicaments, une vitesse largement excessive, etc.). Ainsi, certaines communes se retrouvent projetées très haut dans ce cadran sans que l'on puisse proposer d'explications ou de leviers d'action précis : Vresse-sur-Semois, Saint-Léger, Meix-Devant-Virton et Wasseiges principalement. Notons tout de même que Vresse-sur-Semois, Wasseiges et Saint-Léger étaient apparues en surreprésentation en ce qui concerne les accidents ayant lieu pendant le week-end (cf. section 3.2.2).

3.2.7 Autres conditions particulières de circulation

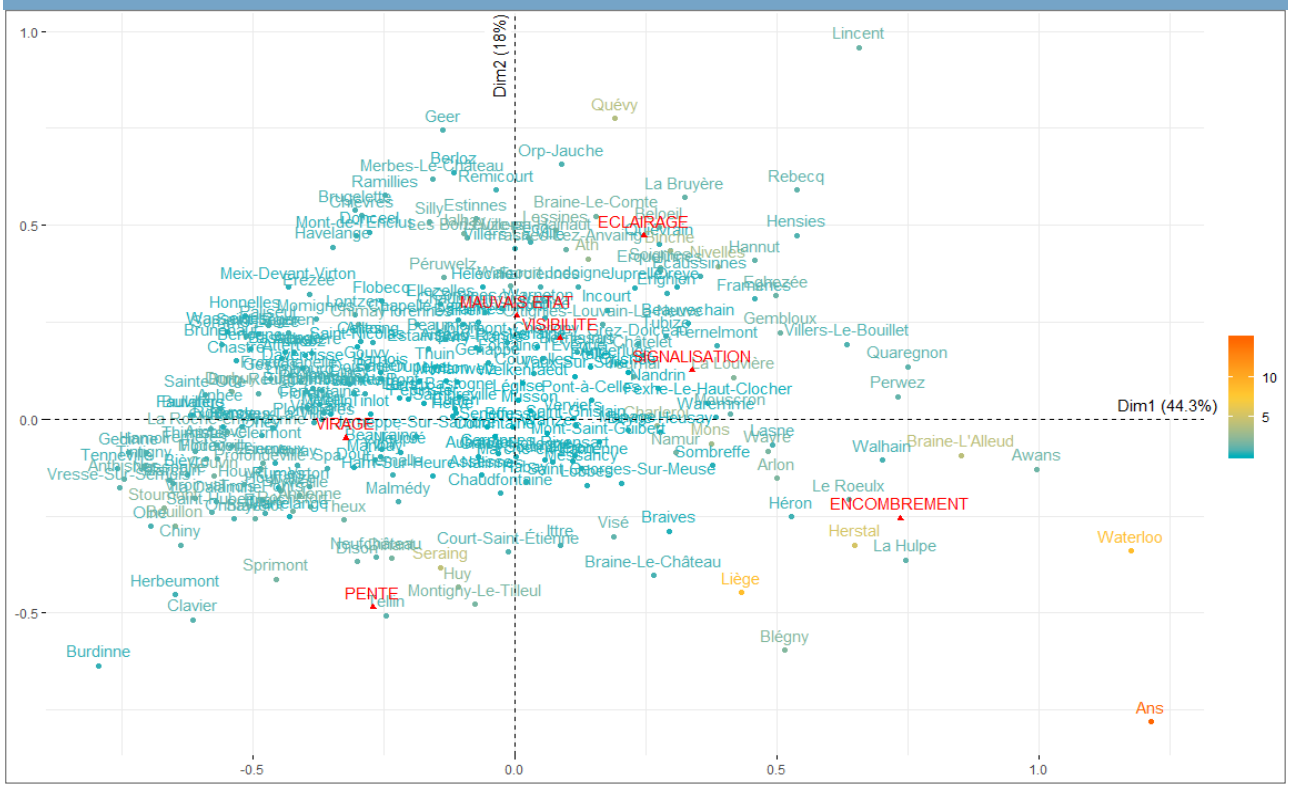
Dans cette section, certaines conditions particulières de circulation, liées à des défauts d'équipement ou à des défauts de perception, sont aussi prises en compte. Sept circonstances sont considérées :

- Mauvais état de la route (ornières, verglas…) ;
- Signalisation défectueuse ;
- Éclairage défectueux ou insuffisant ;
- Encombrement de la circulation, file, accident préalable ;
- Forte descente (pente de 7% ou plus) ;
- Virage serré ;
- Entrave à la visibilité (relief, obstacle fixe, véhicule immobilisé…).

Comme signalé plus tôt, il n'est plus question ici de tableaux de contingence parfaits, dans le sens où la somme des accidents repris dans ces sept modalités n'est plus égale au total des accidents recensés (et ceci pour l'ensemble des communes). Il y a donc une forte perte d'information, des profils calculés sur la base d'effectifs différents et davantage de précautions à adopter lors des interprétations.

La carte des projections obtenues est montrée dans le graphique 12, construit sur la base de 62,3% de variance expliquée par les deux premiers axes (pourcentage plus faible).

Graphique 12 : Carte des profils ligne et colonne dans le plan formé par les deux premiers axes principaux de l'AFC (sur base des conditions particulières de circulation)



Avec toutes les précautions nécessaires, cette carte AFC souligne surtout l'importance des modalités « éclairage », « pente » et « encombrement ».

Les accidents corporels liés à un éclairage défectueux ou insuffisant sont surreprésentés dans les communes suivantes : Lincent, Quévy, Rebecq et La Bruyère (en haut de la carte). Lincent apparaissait auparavant via une surreprésentation des accidents sur autoroute (cf. sections 3.2.4 et 3.2.5). En recoupant les informations, il est peut-être pertinent de se demander si les tronçons d'autoroute présents sur la commune de Lincent sont suffisamment éclairés.

Les accidents corporels liés à une forte descente, une forte pente, sont surreprésentés dans les communes suivantes : Burdinne, Clavier, Herbeumont, Sprimont, Tellin, Montigny-le-Tilleul et Huy (en bas à gauche). Une réflexion sur la signalisation de ce danger pourrait être un levier d'amélioration. Tout comme le fait de veiller au respect de ces signalisations avec, par exemple, une réflexion sur l'installation potentielle de nouveaux radars dans la zone. Notons que Tellin apparaissait déjà pour les tronçons autoroutiers (cf. sections 3.2.4 et 3.2.5), Herbeumont pour les chaussées sales (cf. section 3.2.1).

Les accidents corporels liés à un encombrement de la chaussée (files, embouteillages) sont surreprésentés dans les communes suivantes : Ans (avec de fortes contributions), Waterloo, Blégny, La Hulpe, Herstal, Awans et Liège (le tout en bas à droite). La commune d'Ans, très largement isolée sur la carte, apparaissait déjà pour les accidents liés à une collision en chaîne ou par l'arrière (cf. section 3.2.6). C'est cohérent avec la présence de files de véhicules. Les embouteillages, ou encombrements de façon générale, semblent y être un problème majeur.

3.3. SORTIES DES ACM SUR DES PROFILS COMPLETS D'ACCIDENTS DE LA ROUTE EN WALLONIE (PÉRIODE 2009-2018)

Contrairement à la section précédente, nous allons considérer ici des résultats issus de l'analyse simultanée de plusieurs facteurs, ou variables catégorielles, ayant chacun plusieurs modalités (ou niveaux). Il est très difficile d'obtenir des profils élargis de tous les accidents de la route, un par un. Certaines variables sont régulièrement manquantes, d'autres sont facultatives lorsque la police remplit le constat sur le lieu de l'accident et d'autres encore ne permettent pas d'établir des leviers concrets, ou des actions, à mettre en œuvre. C'est pour cela que les profils des accidents ici étudiés comportent quatre informations catégorielles (cf. section 2.2), et malheureusement pas davantage : la **luminosité**, le **type de route**, le **type de collision** et si l'accident a eu lieu en **agglomération ou hors agglomération**.

Dans cette section, une différenciation est faite entre les accidents corporels mortels (décès dans les 30 jours) et les accidents corporels n'ayant entraîné que des blessés. Des analyses parallèles seront ainsi menées et comparées pour ces deux cas de figure.

La section 3.3.1 est dédiée à la possibilité d'analyses issues d'ACM commune par commune. Quelques résultats illustratifs seront montrés et interprétés, tout en sachant que ces résultats sont stockés et disponibles pour l'ensemble des communes wallonnes[7].

La section 3.3.2 est quant à elle dédiée à des représentations plus globales des modalités des variables, à l'échelle de la Wallonie. Des cartes seront ainsi proposées et permettront de mettre en évidence des cas de surreprésentation, des profils types problématiques et, en conséquence, des pistes de réflexion pour proposer des mesures adéquates.

3.3.1 Représentations multivariées des causes d'accidents au niveau communal

Les analyses multivariées de type ACM permettent donc de tenir compte simultanément de plusieurs facteurs catégoriels. Un premier niveau d'analyse permet de comprendre ce qu'il se passe entre ces facteurs au niveau de chaque commune, une à une. Le but est alors de visualiser directement quelle(s) modalité(s) prédomine(nt), dans le sens où une discrimination est observée.

Concrètement, l'intérêt est de lire et de comprendre les contributions de ces modalités dans le plan formé par les deux premières dimensions principales de l'ACM. Les modalités proches entre elles auront beaucoup d'accidents en commun (causes supposées qui se cumulent ou se complètent). Des modalités éloignées auront peu de cas d'accidents en commun. Les modalités proches de l'origine du plan auront un pouvoir explicatif modéré, dans le sens où elles n'expliqueront pas les particularités des accidents de la commune considérée. Elles ne mettront pas à jour les cas de surreprésentation qui nous intéressent particulièrement dans cette publication. Les modalités les plus éloignées de l'origine auront ainsi une grande importance pour comprendre les causes supposées récurrentes des accidents dans chaque commune.

Comme expliqué plus haut, il est naturellement impossible de représenter toutes les communes ici. Nous nous contenterons donc, de façon arbitraire, d'illustrations basées sur deux communes qui sont ressorties, à un moment ou à un autre, dans la section précédente : une commune urbaine, Liège (cf. graphique 13 et annexe 3 pour des chiffres détaillés), et une commune rurale, Herbeumont (cf. graphique 14 et annexe 4 pour des chiffres détaillés).

Les résultats liés à toute autre commune wallonne peuvent être communiqués sur demande.

Dans la section précédente, portant sur les résultats des AFC, la commune de Liège était significativement apparue pour des accidents ayant impliqué une collision avec des piétons (cf. section 3.2.6 et graphique 11). Ici, la modalité « avec un piéton » est très discrète dans le cadran inférieur gauche : elle est à la fois proche de l'origine et caractérisée par une contribution cumulée faible dans les deux premiers axes (rappelons que la coloration des points, tout comme dans la section 3.2, est en lien direct avec la somme de leurs contributions dans les deux premières dimensions formées, ici, par l'ACM). Ce n'est pas contradictoire ni erroné ; cela

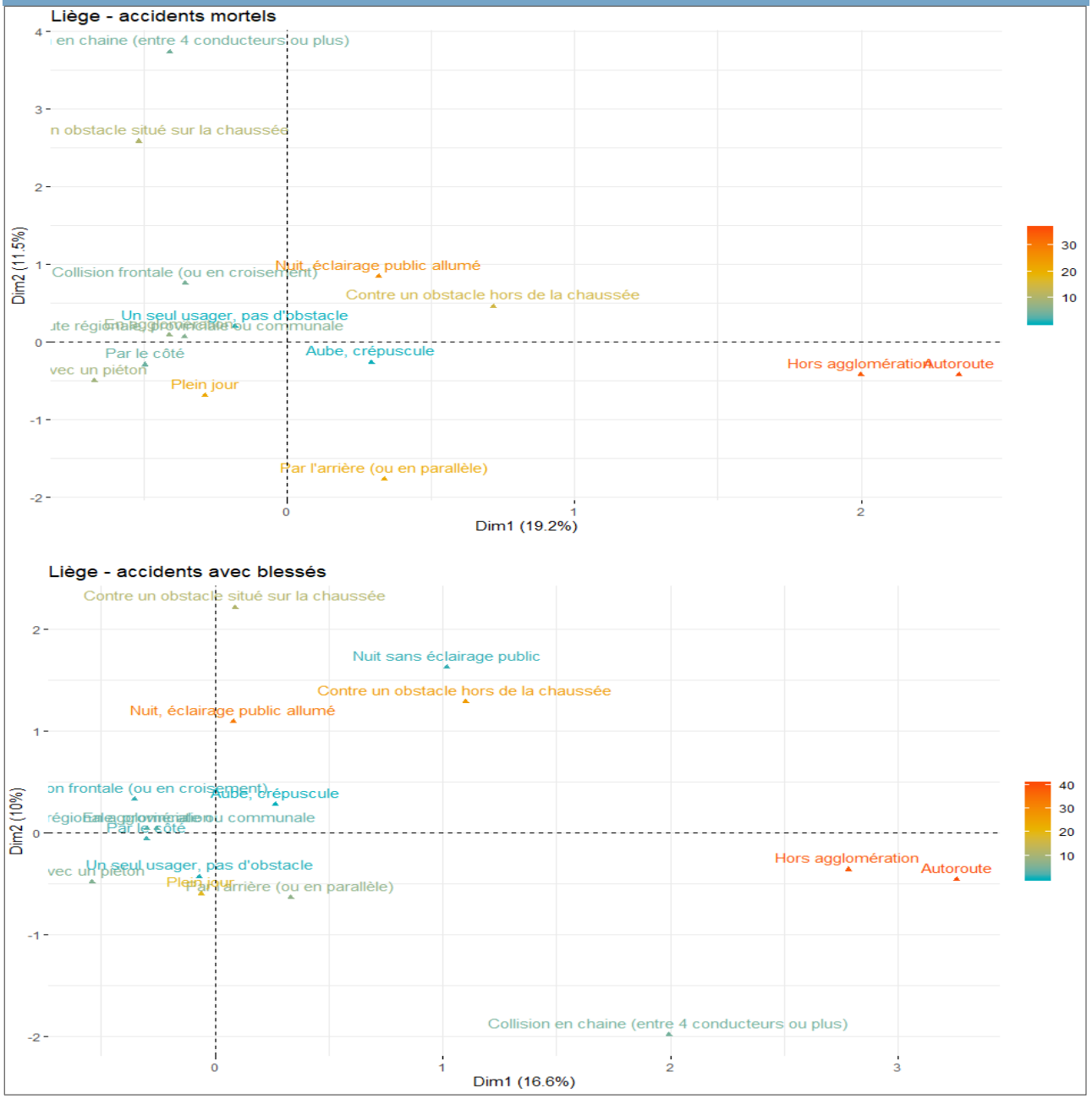
[7] Ces résultats peuvent être communiqués sur demande.

renvoie à l'opposition fondamentale entre des **considérations absolues** et des **considérations relatives**.

En effet, le graphique 13 doit être interprété comme une photographie des accidents survenus à Liège entre 2009 et 2018. Il y a eu davantage d'accidents à Liège entre 2009 et 2018 qui ont eu lieu sur autoroute, qui ont eu lieu hors agglomération, qui ont eu lieu la nuit avec éclairage allumé (surtout pour les accidents non mortels) ou qui ont été causés par des collisions en chaîne.

Ce sont en effet les modalités qui se dégagent le plus, celles qui sont le plus éloignées de l'origine du graphe. Dans une photographie des accidents propres à Liège (survenus entre 2009 et 2018), il y a peu d'accidents liés à une collision avec un piéton, de façon absolue. Ils y sont par exemple minoritaires par rapport aux accidents survenus sur autoroute qui, en principe, n'impliquent pas de piétons.

Graphique 13 : Projections des modalités des facteurs, commune de Liège. En haut, pour les accidents corporels mortels ; en bas, pour les accidents corporels avec blessés



Graphique 14 : Projections des modalités des facteurs, commune de Herbeumont. En haut, pour les accidents corporels mortels ; en bas, pour les accidents corporels avec blessés

Mais ces accidents avec piétons à Liège ont été évalués par l'AFC comme étant en surreprésentation dans le graphique 11. Cette surreprésentation est à mettre en relation avec les notions de projection de profils ligne et de profils colonne (comme on le voit dans la section B de l'annexe 1 théorique). Elle est donc, par construction, liée à une pondération entre toutes les communes. La conséquence de tout ceci est que l'on peut dire que, de façon relative maintenant, les accidents impliquant une collision avec un piéton à Liège entre 2009 et 2018 sont en surreprésentation par rapport à un profil

moyen de ces mêmes accidents dans toutes les autres communes.

Pour Herbeumont, le graphique 14 nous propose une photographie des accidents mortels d'une part, et des accidents corporels avec blessés d'autre part. On y constate tout d'abord que le nombre de modalités représentées est différent dans les deux cas. Logiquement, certaines modalités sont totalement absentes dans les deux cas (en premier lieu « autoroute » qui n'est pas d'application dans cette commune). Ce nombre différent de modalités est à mettre en lien avec le très petit nombre d'accidents survenus à Herbeumont entre 2009 et 2018 : 42 accidents corporels en tout, dont quatre seulement se sont avérés mortels.

Par conséquent, la carte des accidents dans la partie supérieure du graphique 14 n'est construite que sur la base de ces quatre seuls accidents. Très peu d'accidents signifie automatiquement très peu de causes d'accident. Des modalités disparaissent donc. Et le pouvoir informatif d'une telle carte s'en trouve immédiatement très limité. Nous démontrons ici par l'absurde que ce type d'analyses absolues est tributaire de la quantité d'informations à disposition, ces analyses étant plus robustes dans certains cas (Liège) que dans d'autres (Herbeumont).

La partie inférieure du graphique 14 est cependant davantage informative, car basée sur plus d'accidents. Apparemment, les accidents avec blessés à Herbeumont ont eu lieu le plus souvent la nuit (ou aube et crépuscule), en agglomération, ont impliqué un choc avec un piéton ou avec un obstacle situé sur la chaussée.

Dans la section 3.2.3 (cf. graphique 6) sur l'analyse du facteur « luminosité », Herbeumont apparaissait clairement parmi les communes impactées par les accidents de nuit sans éclairage. Ceci est donc largement confirmé ici.

3.3.2 Discussion : analyse par commune *versus* perte de généralisation

Si l'on se contente des analyses vues dans la section 3.3.1, commune par commune, le principal avantage sera de disposer de photographies précises des causes d'accidents corporels, qu'il ne faudrait interpréter qu'en termes absolus. Ces analyses pourront être confirmatoires d'observations tirées d'AFC précédentes. Mais elles pourront aussi nous rendre plus perplexe, comme ce fut le cas pour Liège.

Un autre problème est la perte d'information importante lorsque l'analyse se porte sur des communes qui sont peu concernées, avec peu d'accidents. Pour Herbeumont, on a ainsi vu des photographies dans lesquelles de nombreuses modalités manquent, faute d'avoir été constatées réellement. De plus, construire des cartes sur la base de quatre accidents n'a pas beaucoup de sens.

Comment généraliser l'apport des ACM ? Comment en tirer de l'information non pas absolue (et exclusive à une seule commune) mais plutôt générale et parlante pour toutes les communes à la fois, que l'on pourrait alors directement comparer ? Une perspective est de travailler de nouveau sur les contributions dans les deux principaux axes construits par le modèle. La somme des contributions à ces deux axes a déjà servi à colorer les points dans la plupart des graphiques précédents. Ces sommes sont disponibles et calculées pour toutes les modalités des variables considérées, et ceci dans l'ensemble des communes wallonnes.

La solution est alors toute trouvée : construire des cartes géographiques afin de visualiser des « points chauds » pour une modalité donnée. C'est ce que l'on va découvrir dans la section 3.3.3 suivante.

3.3.3 Représentations cartographiques générales des contributions des modalités des facteurs

Dans cette partie dédiée aux analyses de type ACM, 16 modalités renseignées différentes (issues de quatre variables catégorielles, cf. section 2.2) sont étudiées simultanément. Cela laisse la possibilité de construire jusqu'à 16 cartes de contributions cumulées, et même 32 si l'on continue de différencier les accidents corporels mortels et ceux avec blessés.

Le but ici n'est pas de montrer toutes ces cartes. Certaines ne mettront pas en évidence de « points chauds » pertinents (typiquement la modalité

« plein jour »), certaines encore ne permettent pas de déterminer directement un levier d'action potentiel (accidents avec conducteur seul sans obstacle, accidents avec choc sur le côté, accidents de nuit avec éclairage public allumé…) et d'autres seront par construction incomplètes. En effet, si l'on prend la variable « type de route » et ses deux modalités (autoroute *versus* route régionale ou communale), nous obtiendrons une carte pertinente du réseau autoroutier wallon dans laquelle nous pourrons mettre en avant des cas sensibles liés aux accidents survenus sur autoroute (cf. cartes 1 et 2). Mais en se focalisant sur l'autre niveau de la variable (route régionale ou communale), la carte serait largement incomplète car elle ne ferait apparaître à nouveau que les communes concernées par le réseau autoroutier (qui traverse leur territoire).

Concrètement, pour les communes qui ne sont pas traversées par une autoroute, tous les accidents ont eu lieu d'office sur une route régionale ou communale. Ce n'est donc plus un facteur favorable ou non, c'est une certitude, une évidence. Et cela n'a ainsi rien d'informatif.

Cette carte permet donc de visualiser directement les points chauds liés à la modalité « autoroute », sur la base des contributions cumulées (première et deuxième composantes principales) de la modalité dans le processus de l'ACM.

En visualisant les cartes 1 et 2, construites respectivement sur base des accidents avec blessés et des accidents mortels, il faut être conscient que les calculs sous-jacents y sont différents, tout comme le nombre d'accidents considérés (davantage d'accidents corporels que mortels). Ceci explique dans quelle mesure les points chauds peuvent différer d'une carte à l'autre. Pour la carte 1, les plus gros scores concernent les communes de Beyne-Heusay (avec 46.39%), Saint-Nicolas, Liège, Celles, Yvoir et Boussu. Pour la carte 2 (accidents mortels), ils concernent Flémalle (43.59%), Malmedy, Bertrix, Mouscron et Rochefort. Liège et Boussu suivent à nouveau de près, Héron apparaît aussi rapidement (commune déjà identifiée dans le graphique 7).

En recoupant ainsi les deux cartes, il est donc envisageable de détecter des « territoires noirs » de façon globale, pour tous les types d'accidents corporels confondus. Ce type d'analyse serait alors à mettre en parallèle avec d'autres travaux comme le Baromètre de la sécurité routière de l'Institut VIAS[8].

De la même façon, les autres modalités considérées peuvent être cartographiées à l'échelle de la Wallonie. Les cartes basées sur les accidents avec blessés sont plus informatives, car construites sur la base de plus d'accidents recensés. Les cartes suivantes basées sur les accidents mortels seront donc dorénavant disponibles en annexe 5. Tout d'abord, les cartes 3 et 4 montrent les contributions cumulées des ACM pour les modalités « agglomération » et « hors agglomération ».

Pour les accidents constatés en agglomération, les scores les plus élevés concernent les communes de Somme-Leuze (avec 38.15%), Nassogne, La Roche-en-Ardenne, Floreffe et Tinlot.

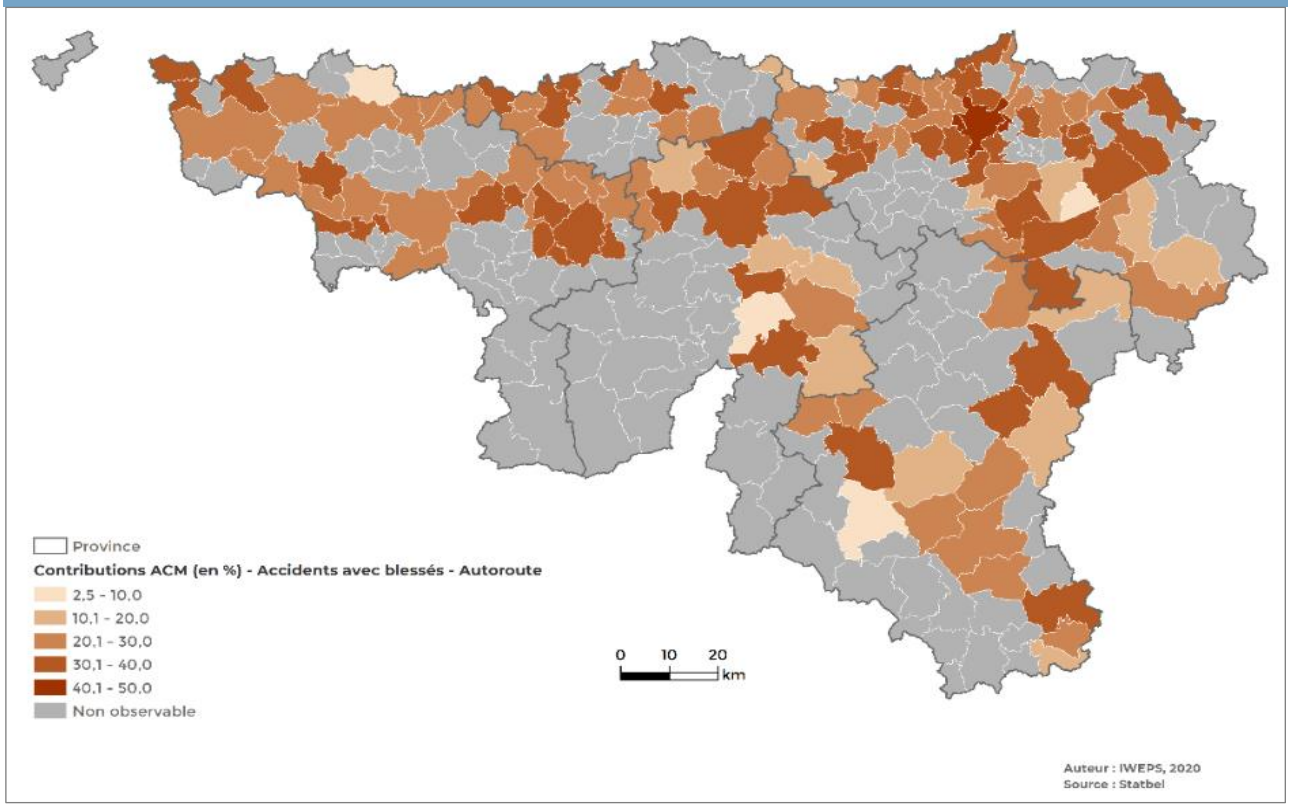
Pour ceux constatés hors agglomération, les scores les plus importants se trouvent à Beyne-Heusay (avec 48.50%), Saint-Nicolas, Colfontaine, Frameries, Binche et Liège.

On retrouve Beyne-Heusay, Saint-Nicolas et Liège. On recoupe donc, de façon assez naturelle, les accidents sur autoroute et les accidents hors agglomération pour ces communes. On peut parler ici de problème récurrent pour ce petit *cluster* de communes de l'agglomération liégeoise.

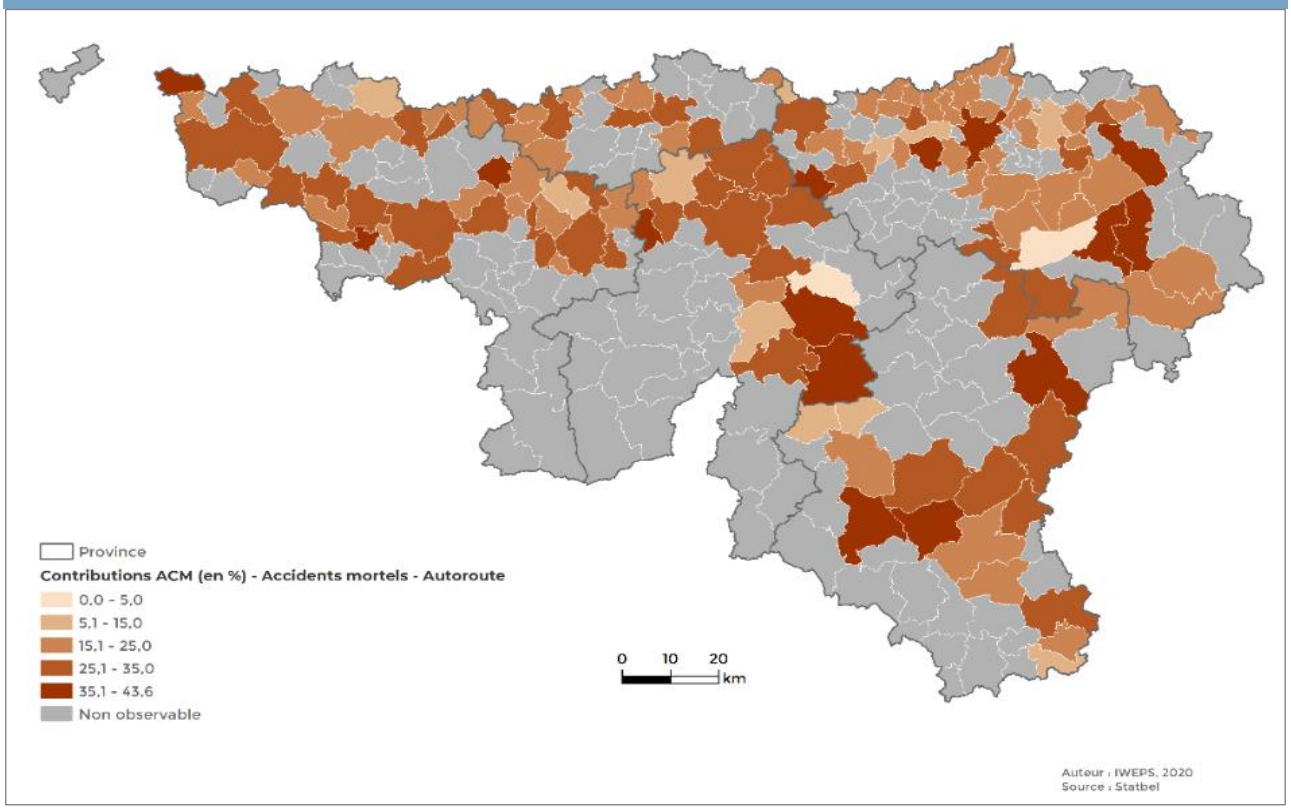
Pour ce qui est de la variable luminosité, il est intéressant de pouvoir visualiser les communes par rapport aux modalités « aube, crépuscule » (cf. carte 5) et « nuit sans éclairage public » (cf. carte 6). Ce sont en effet les modalités qui pourraient le mieux impliquer un levier d'action en cas de surreprésentation manifeste.

[8] https://www.vias.be/fr/recherche/publications/?sort=score%20desc&years[]=2020&themes[]=Accidents

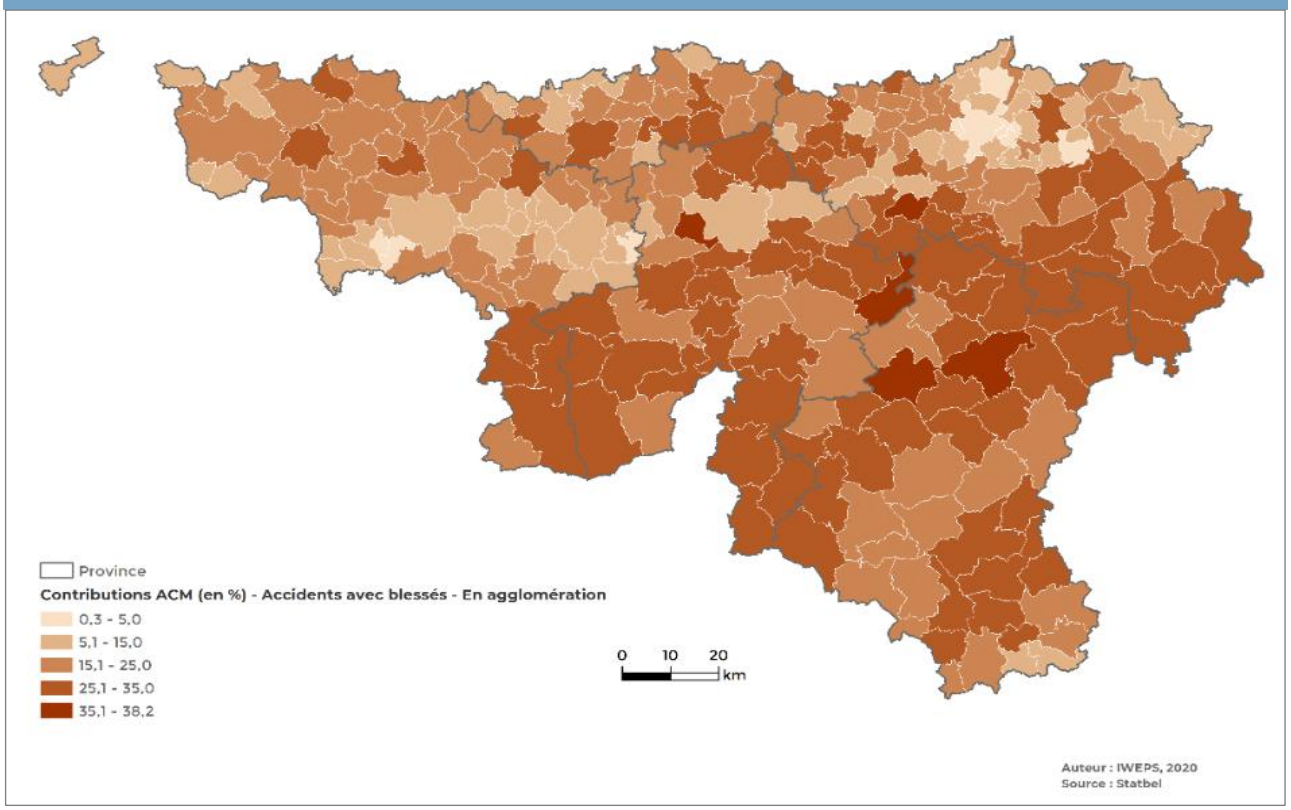
Carte 1 : Représentation des contributions ACM cumulées – Accidents avec blessés – Modalité « autoroute »



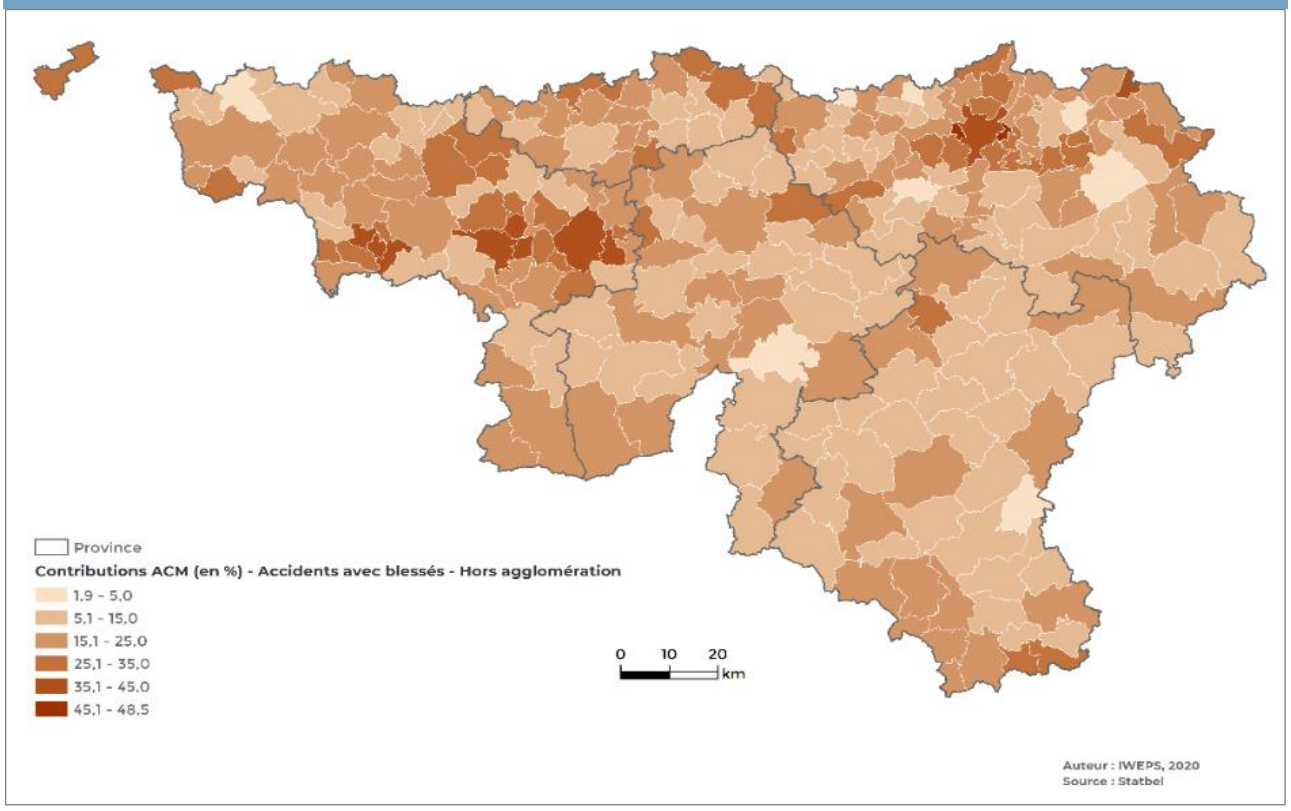
Carte 2 : Représentation des contributions ACM cumulées – Accidents mortels – Modalité « autoroute »



Carte 3 : Représentation des contributions ACM cumulées – Accidents avec blessés – Modalité « agglomération »



Carte 4 : Représentation des contributions ACM cumulées – Accidents avec blessés – Modalité « hors agglomération »



La carte 5 (« aube, crépuscule ») fait apparaître une série de communes avec un score relativement élevé : Berloz (avec 46.43%), Trois-Ponts, Chapelle-lez-Herlaimont, Tintigny et Les Bons Villers.

La carte 6 (« nuit sans éclairage public ») fait quant à elle apparaître les communes suivantes : Fernelmont (48.67%), Mont-de-L'Enclus, Pont-à-Celles, Herbeumont, Paliseul, Attert, Gedinne… Notons la cohérence avec certaines des communes clairement identifiées au préalable dans le graphique 6 plus haut.

Une dernière série de représentations cartographiques peut également être riche en informations pertinentes. Il s'agit des types de collision, et plus particulièrement les accidents ayant impliqué un choc avec des piétons ou avec un obstacle situé sur la chaussée (de nouveau, ces deux catégories sont celles qui peuvent éventuellement impliquer un levier d'action plus évident). La carte 7 concerne ainsi les accidents corporels avec blessés liés à la présence active d'un piéton.

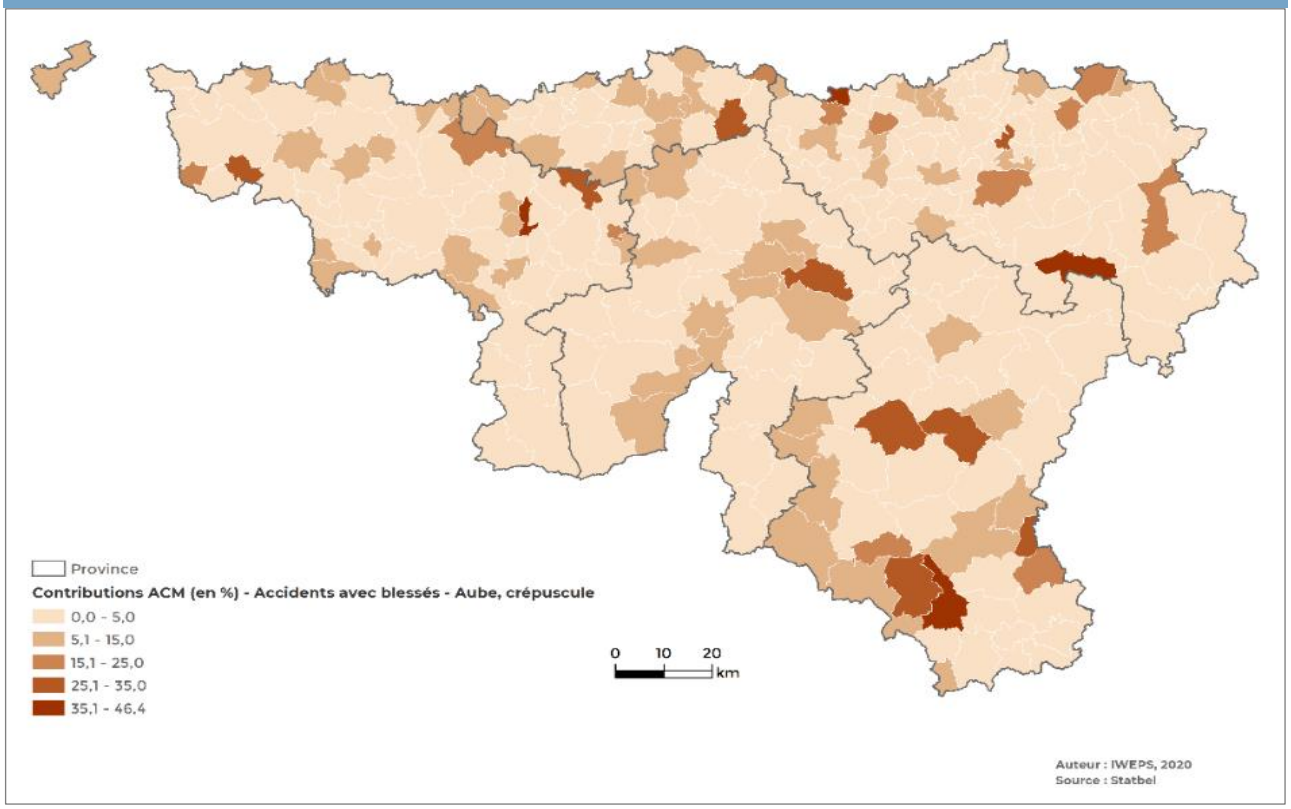
Sur cette carte 7, certaines communes wallonnes se distinguent, en premier lieu les communes de Floreffe (39.94%) et de Binche, pour lesquelles une attention toute particulière devrait être portée sur la protection des piétons. Suivent ensuite Gembloux, Martelange, Braine-le-Comte, Walcourt…

Notons qu'on ne retrouve pas ici les communes mises en avant dans le graphique 11 (Liège, Verviers, Frameries…). Ceci est la preuve, une nouvelle fois, que les enseignements issus d'analyses statistiques univariées ne convergent pas forcément vers les enseignements issus d'analyses statistiques multivariées, comme ici.

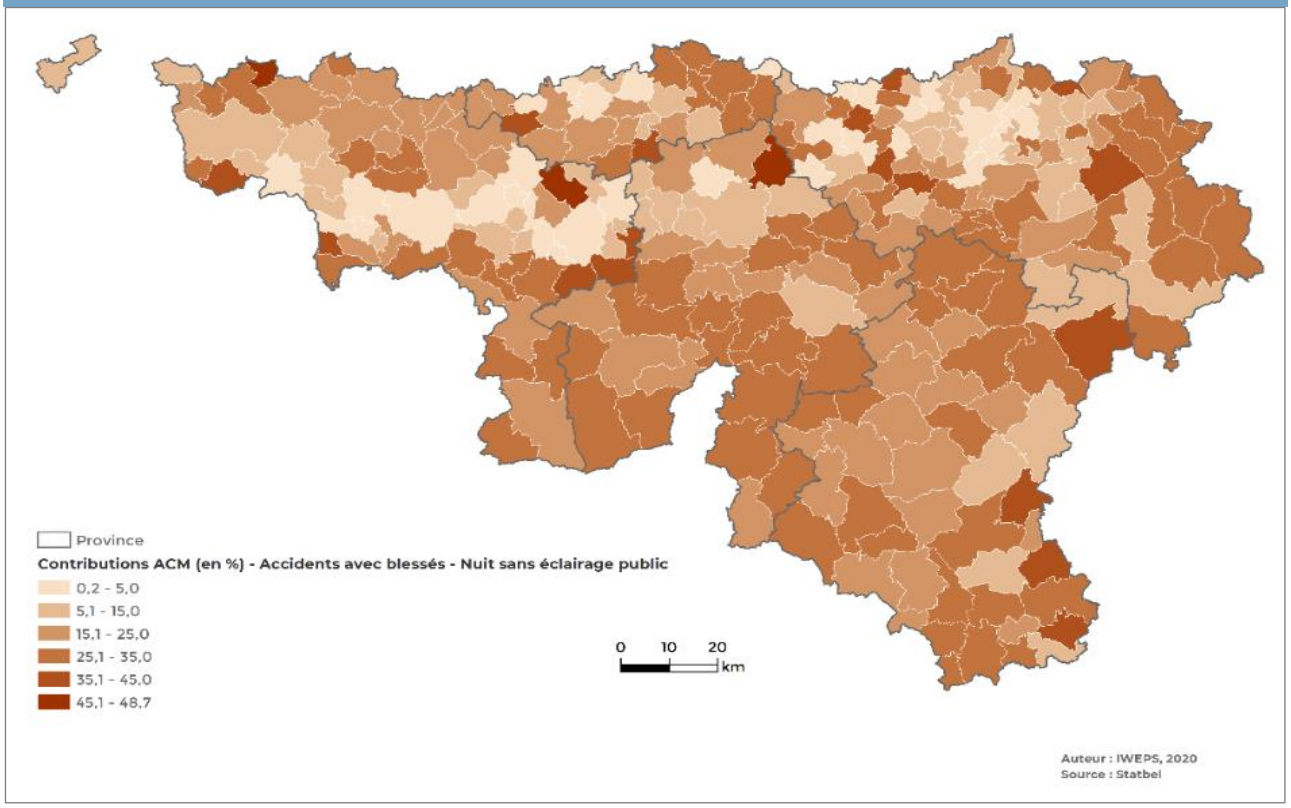
Enfin, la carte 8 représente les communes selon les contributions cumulées de la modalité « contre un obstacle situé sur la chaussée » dans les deux premières composantes de l'ACM (sur la base des accidents avec blessés).

Les communes de Juprelle (40.31%), Marchin, La Calamine, Tintigny et Seneffe arrivent en tête. Ces communes sont ainsi relativement plus sujettes à ce type d'accidents.

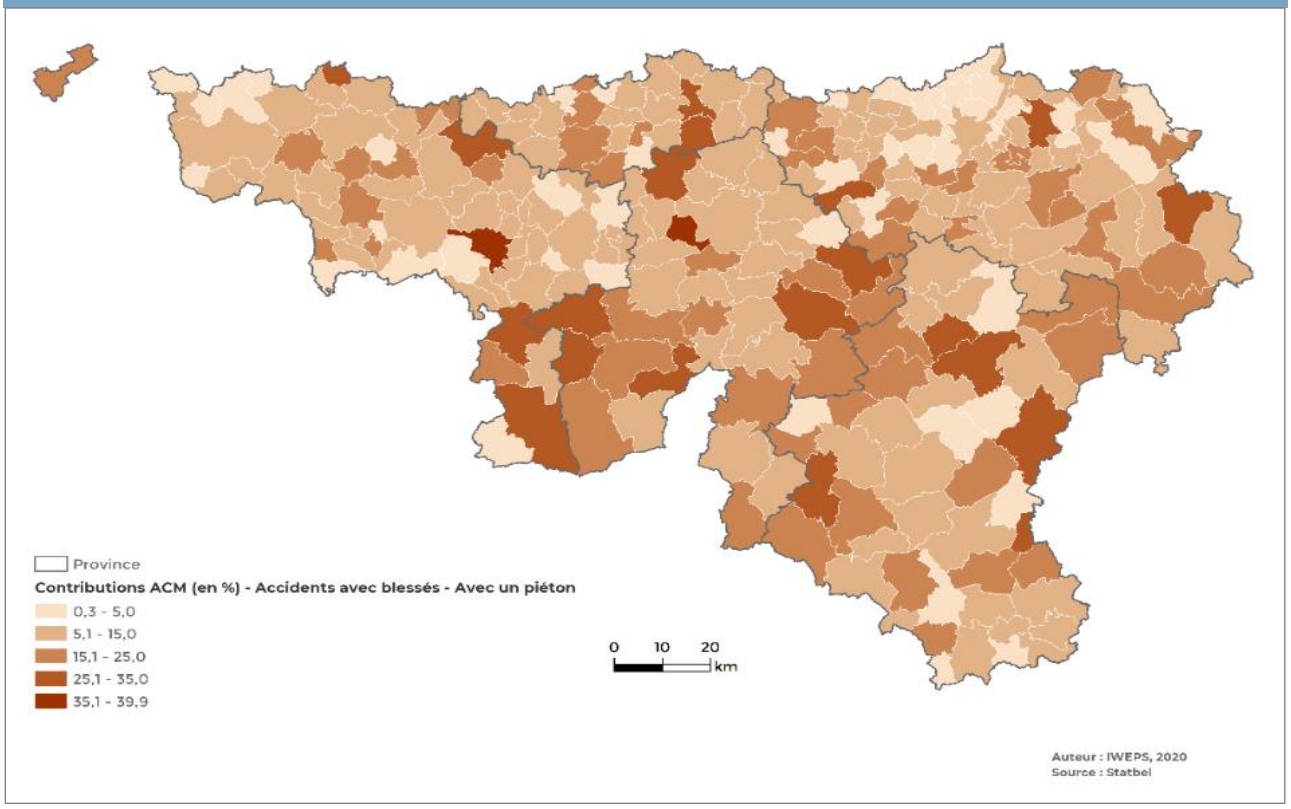
Carte 5 : Représentation des contributions ACM cumulées – Accidents avec blessés – Modalité « aube, crépuscule »



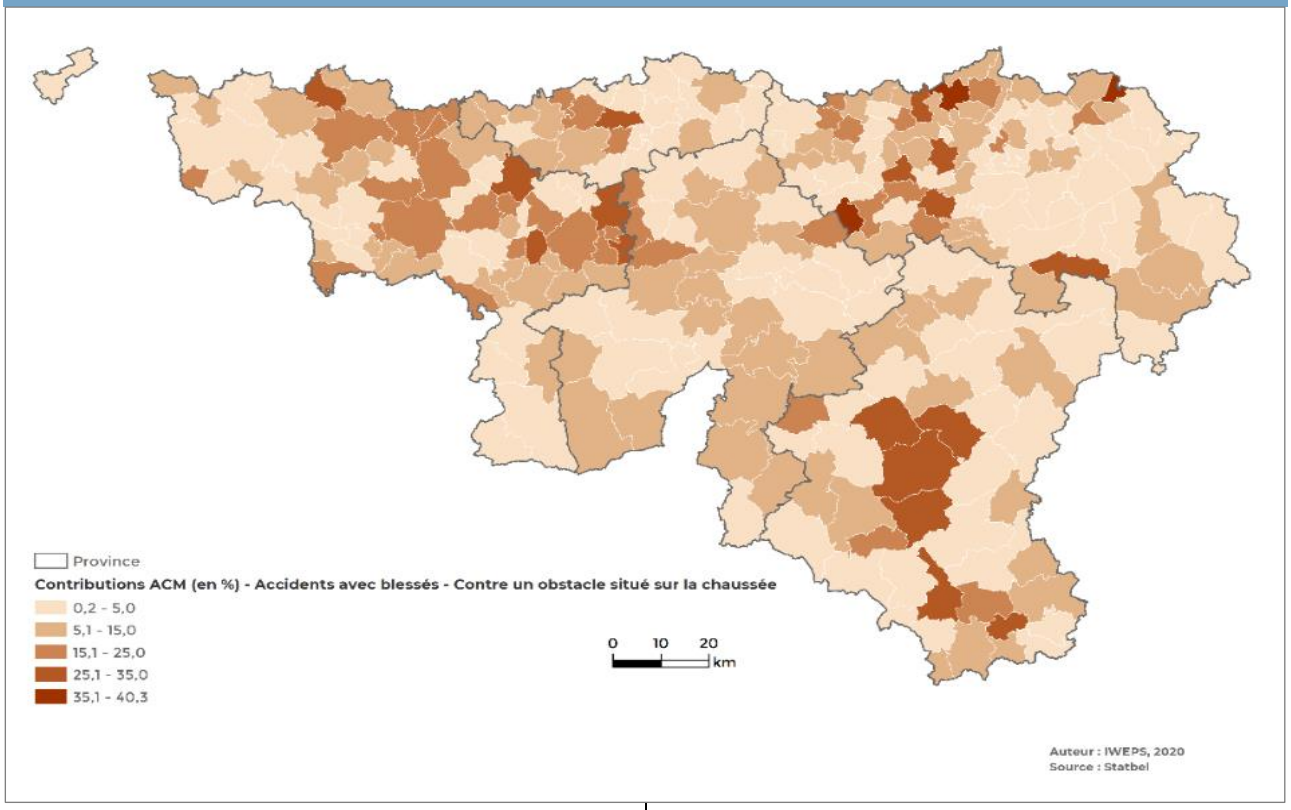
Carte 6 : Représentation des contributions ACM cumulées – Accidents avec blessés – Modalité « nuit sans éclairage public »



Carte 7 : Représentation des contributions ACM cumulées – Accidents avec blessés – Modalité « avec un piéton »



Carte 8 : Représentation des contributions ACM cumulées – Accidents avec blessés – Modalité « contre un obstacle situé sur la chaussée »



3.3.4 Tableau récapitulatif des ACM commune par commune

Pour conclure cette partie sur les contributions des facteurs d'accidentalité aux ACM, un tableau récapitulatif semble utile pour synthétiser les informations obtenues. Le tableau 1 reprend pour cela les contributions cumulées dans les deux principaux axes des ACM, ces mêmes contributions qui ont permis de construire les cartes de la section 3.3.3. Pour rappel, ces contributions ont été calculées sur la base des accidents corporels ayant causé des blessés en Wallonie entre 2009 et 2018.

Les facteurs les plus pertinents ont été conservés et ont été renommés pour l'occasion afin que le tableau soit lisible :

- « Autoroute » = AUTO
- « En agglomération » = AGGL
- « Hors agglomération » = H-AGGL
- « Aube, crépuscule » = AUBE
- « Nuit sans éclairage public » = NSEP
- « Nuit, éclairage public allumé » = NAEP
- « Avec un piéton » = PIETON
- « Contre un obstacle situé sur la chaussée » = OBS-CH.

Les contributions sont donc reprises dans ce tableau, accompagnées d'un jeu de couleurs pouvant favoriser une lecture rapide des problèmes rencontrés. Les communes ayant les dix contributions les plus élevées pour un facteur donné seront signalées en rouge, les communes classées entre la 11e et la 30e place seront en orange. Les cinq communes qui ont les contributions les plus élevées sont de plus signalées en gras.

Ce tableau 1, ainsi que tout résultat disponible dans cette publication, peut servir à l'avenir à la construction ou à l'amélioration d'indicateurs communaux (composites ou non) en lien avec la thématique de la sécurité routière. Ces résultats sont avant tout une invitation à aller plus loin dans l'analyse de la sécurité routière en Wallonie, ils ont pour but d'attirer l'attention et de poser des questions pour comprendre pourquoi telle ou telle commune ressort pour l'un ou l'autre facteur.

Tableau 1 : Récapitulatif des principales contributions cumulées commune par commune

COMMUNE	AUTO	AGGL	H-AGGL	AUBE	NSEP	NAEP	PIETON	OBS-CH
Aiseau-Presles	NA	4,214	17,730	8,094	35,182	25,668	7,013	25,461
Amay	NA	14,760	23,448	8,189	42,556	18,907	17,798	0,540
Amblève	15,914	33,086	13,373	2,667	33,048	10,369	21,504	10,661
Andenne	36,146	14,800	25,811	0,232	11,738	19,587	10,447	2,489
Anderlues	NA	15,372	35,704	2,616	9,972	28,667	9,538	18,706
Anhée	NA	27,537	19,072	3,534	18,893	39,495	14,399	8,187
Ans	23,020	15,321	21,467	7,014	6,056	29,299	3,959	6,780
Anthisnes	NA	25,355	14,728	0,766	29,978	34,061	11,397	27,481
Antoing	25,456	19,783	14,421	25,809	7,705	18,627	7,053	6,423
Arlon	36,498	18,860	18,719	1,891	28,852	2,578	10,860	7,055
Assesse	19,734	28,771	5,046	6,134	33,776	6,825	10,646	3,276
Ath	22,143	15,433	16,281	1,220	18,232	36,384	8,144	21,943
Attert	NA	27,859	5,785	21,318	44,661	0,142	22,329	9,873
Aubange	18,659	13,237	29,230	1,508	14,173	38,144	8,870	4,781
Aubel	NA	21,689	18,099	0,131	40,234	25,886	14,081	6,108
Awans	26,578	17,542	13,002	13,799	4,345	23,925	3,337	26,631
Aywaille	35,130	18,909	11,307	0,130	30,143	3,486	6,674	0,978
Baelen	36,315	24,842	8,281	2,625	31,463	0,384	0,414	1,872
Bassenge	33,251	10,268	29,258	0,019	15,307	28,496	2,060	7,292
Bastogne	13,968	21,328	18,051	2,879	12,902	35,693	25,590	0,290
Beaumont	NA	33,045	12,806	0,091	24,540	26,958	27,411	3,805
Beauraing	NA	28,771	14,722	0,007	29,785	29,484	20,480	7,656
Beauvechain	NA	12,426	30,902	14,886	31,358	21,449	7,098	3,613
Beloeil	30,015	24,102	20,226	2,099	9,314	27,046	12,850	12,124
Berloz	14,236	19,320	4,830	46,434	1,320	8,996	4,797	19,504
Bernissart	29,178	21,978	10,780	2,059	8,844	22,359	6,765	0,389
Bertogne	32,375	28,386	10,890	7,688	17,841	5,097	4,573	0,671
Bertrix	6,604	24,028	17,523	0,089	27,124	34,183	17,211	9,877
Beyne-Heusay	46,393	0,326	48,499	4,852	0,194	24,398	0,988	2,502
Bièvre	NA	29,145	15,780	1,784	32,059	39,452	10,797	11,510
Binche	NA	11,968	40,692	2,907	13,163	21,747	35,995	1,138
Blégny	28,973	13,710	24,260	0,060	3,121	34,945	4,991	0,157
Bouillon	NA	34,235	11,214	5,038	28,627	26,186	23,744	3,299
Boussu	39,172	5,307	36,550	0,581	3,005	30,446	6,254	3,831
Braine-l'Alleud	33,676	13,870	22,051	4,966	16,595	14,098	8,401	8,148
Braine-le-Château	24,603	19,703	16,948	1,064	2,646	26,584	6,223	12,402
Braine-le-Comte	NA	18,668	25,679	17,202	15,050	15,555	30,810	6,285
Braives	38,450	32,275	10,252	7,522	1,495	30,274	21,552	5,032
Brugelette	NA	17,567	14,167	7,033	24,467	34,559	2,151	8,918
Brunehaut	NA	12,615	28,066	1,914	35,930	34,398	5,718	1,397
Bullange	NA	26,446	12,910	2,494	30,577	40,910	12,452	2,387

COMMUNE	AUTO	AGGL	H-AGGL	AUBE	NSEP	NAEP	PIETON	OBS-CH
Burdinne	NA	30,653	13,314	1,273	31,906	27,480	16,294	9,187
Burg-Reuland	NA	31,949	12,119	2,371	31,711	33,734	13,832	1,855
Butgenbach	NA	20,375	15,416	0,927	34,005	36,620	30,715	4,446
Celles	39,969	7,503	1,948	2,366	26,125	34,028	2,723	1,678
Cerfontaine	NA	34,491	10,901	3,990	25,065	27,695	25,632	11,341
Chapelle-lez-Herlaimont	28,384	14,181	15,881	39,505	3,231	4,425	7,882	2,547
Charleroi	37,344	5,042	35,063	0,990	1,104	35,309	6,987	21,786
Chastre	NA	6,685	28,464	1,227	35,135	33,503	4,557	0,787
Châtelet	37,600	5,079	35,227	2,427	7,843	26,353	5,034	20,630
Chaudfontaine	31,461	14,259	25,197	0,440	1,433	30,258	8,113	3,859
Chaumont-Gistoux	30,910	19,474	13,319	11,221	7,591	15,535	5,480	1,467
Chièvres	NA	22,635	22,517	5,459	27,666	29,851	15,132	8,543
Chimay	NA	27,634	15,372	2,804	20,545	31,434	26,435	1,024
Chiny	NA	23,127	16,968	25,774	21,064	20,969	21,574	1,316
Ciney	26,594	24,944	14,793	5,575	13,045	20,651	26,386	0,549
Clavier	NA	29,911	9,330	1,670	16,823	37,261	17,595	5,777
Colfontaine	NA	1,322	42,836	5,013	13,051	34,299	18,161	8,620
Comblain-au-Pont	14,780	17,317	23,232	4,525	15,983	20,076	11,654	4,991
Comines-Warneton	NA	12,877	27,935	10,463	12,591	35,985	20,739	1,049
Courcelles	31,910	9,248	30,040	3,924	22,759	17,632	9,449	22,827
Court-Saint-Etienne	NA	22,172	21,053	2,671	24,406	41,859	9,746	16,929
Couvin	NA	28,100	15,542	1,620	25,110	30,857	18,551	14,344
Crisnée	22,377	24,067	4,640	11,250	10,505	24,203	3,467	13,504
Dalhem	NA	13,985	24,322	7,532	26,961	36,962	5,023	8,157
Daverdisse	NA	27,976	13,662	11,810	17,315	37,421	17,613	3,235
Dinant	3,407	20,467	24,851	3,045	25,084	36,552	12,115	6,073
Dison	30,281	10,200	27,967	0,277	26,469	10,904	4,637	3,362
Doische	NA	25,644	9,354	5,512	23,397	37,103	31,262	4,308
Donceel	NA	19,587	15,856	20,238	19,813	32,867	8,494	3,226
Dour	NA	10,131	27,376	1,437	22,334	31,588	13,127	3,117
Durbuy	NA	27,793	16,382	2,756	30,624	29,009	13,378	1,998
Écaussinnes	NA	16,811	28,915	0,575	23,919	37,218	20,193	3,817
Eghezée	30,351	27,542	13,094	0,044	16,053	27,309	12,695	1,790
Ellezelles	NA	31,824	6,569	6,585	21,956	24,720	13,255	25,933
Enghien	25,146	18,400	11,621	5,947	19,128	20,036	19,893	19,062
Engis	NA	25,524	17,610	0,687	16,623	32,675	13,622	32,989
Erezée	NA	26,254	7,540	1,990	33,855	44,469	10,678	5,497
Erquelinnes	NA	18,351	23,026	5,698	21,779	37,898	12,214	16,471
Esneux	30,158	16,431	17,966	3,573	2,555	26,910	15,843	6,558
Estaimpuis	30,107	15,556	12,162	0,337	20,077	37,833	6,051	2,519
Estinnes	NA	22,083	14,380	5,567	31,106	32,594	4,986	0,720
Etalle	NA	31,898	14,446	2,207	26,698	35,238	8,102	20,988
Eupen	NA	7,855	33,578	1,469	26,892	32,057	21,413	2,289

COMMUNE	AUTO	AGGL	H-AGGL	AUBE	NSEP	NAEP	PIETON	OBS-CH
Faimes	NA	13,563	20,171	0,336	38,897	30,936	5,564	15,369
Farciennes	NA	11,643	27,505	22,653	12,620	18,425	10,272	23,446
Fauvillers	NA	32,492	4,513	8,001	44,368	9,330	3,046	11,741
Fernelmont	26,919	25,983	8,246	1,533	48,675	6,716	6,025	9,531
Ferrières	23,362	31,230	5,626	1,338	30,007	10,306	14,921	12,991
Fexhe-le-Haut-Clocher	30,553	22,075	13,942	3,800	0,525	32,856	10,298	20,444
Flémalle	36,809	7,624	29,675	1,968	7,912	31,486	3,831	7,832
Fléron	NA	1,705	28,644	26,762	11,298	30,818	9,704	20,962
Fleurus	29,371	15,016	21,141	0,643	2,726	27,769	4,151	26,061
Flobecq	NA	19,120	16,570	9,546	26,381	30,044	26,173	14,911
Floreffe	NA	35,226	12,080	1,917	14,979	15,720	39,938	3,478
Florennes	NA	24,057	16,571	0,001	30,487	34,316	18,908	1,571
Florenville	NA	22,531	16,528	7,177	22,279	33,561	11,812	1,259
Fontaine-l'Evêque	30,328	12,888	25,136	3,410	1,182	35,145	10,061	26,774
Fosses-la-Ville	NA	22,181	15,936	9,454	21,524	34,067	14,416	15,115
Frameries	NA	4,925	41,613	0,115	15,367	35,344	8,857	7,557
Frasnes-lez-Anvaing	28,346	22,604	5,670	4,081	18,673	31,395	9,928	5,873
Froidchapelle	NA	31,149	11,483	3,453	28,673	37,621	5,628	7,842
Gedinne	NA	29,374	10,482	0,998	32,738	40,528	13,439	7,237
Geer	NA	26,710	7,672	17,945	31,847	22,314	10,122	20,831
Gembloux	11,278	24,064	15,241	10,410	18,701	10,929	33,233	2,894
Genappe	NA	33,109	14,546	2,644	20,397	17,977	24,407	5,656
Gerpinnes	NA	8,553	12,295	1,265	38,082	34,176	4,060	8,514
Gesves	NA	26,326	12,536	7,377	30,435	31,773	11,849	13,157
Gouvy	NA	29,066	11,715	4,763	38,396	25,399	16,141	0,669
Grâce-Hollogne	24,800	17,559	20,897	0,110	5,802	27,999	5,279	9,386
Grez-Doiceau	NA	19,722	18,106	1,077	32,382	36,947	7,657	3,203
Habay	25,782	28,567	9,983	0,170	7,887	36,109	22,700	0,319
Ham-sur-Heure-Nalinnes	NA	26,223	20,359	2,426	28,433	36,006	15,723	5,349
Hamoir	16,491	25,765	6,838	26,858	22,329	11,371	15,195	0,407
Hamois	NA	11,966	30,550	0,310	35,892	33,724	8,171	7,284
Hannut	28,829	23,041	21,106	0,405	17,963	25,800	18,946	0,840
Hastière	NA	26,986	16,192	13,248	26,727	18,901	11,140	0,259
Havelange	NA	34,012	11,255	0,114	27,476	28,289	25,977	2,194
Hélécine	17,959	23,514	11,113	17,643	4,225	15,329	5,040	3,800
Hensies	30,375	21,478	9,885	3,656	0,416	28,056	7,811	2,318
Herbeumont	NA	24,599	13,666	19,890	29,879	22,073	8,195	17,138
Héron	17,940	26,408	7,331	1,408	3,374	27,691	12,505	0,528
Herstal	30,127	11,422	25,186	2,154	5,628	31,264	4,196	14,659
Herve	22,487	25,252	10,541	1,774	8,834	19,240	25,375	1,189
Honnelles	NA	12,399	23,742	10,449	33,723	11,730	2,957	21,361
Hotton	NA	19,671	25,171	0,666	29,847	38,184	8,271	5,981
Houffalize	30,301	26,351	6,981	4,745	33,234	7,351	14,854	5,977

COMMUNE	AUTO	AGGL	H-AGGL	AUBE	NSEP	NAEP	PIETON	OBS-CH
Houyet	34,245	26,517	4,614	0,491	28,593	8,567	7,027	9,150
Huy	NA	11,763	32,244	2,696	18,530	31,543	25,445	2,942
Incourt	NA	33,051	10,727	7,455	30,376	17,153	29,795	2,171
Ittre	26,880	28,363	8,657	3,528	38,091	8,896	6,476	0,779
Jalhay	37,305	25,368	4,839	1,885	37,479	4,631	6,111	1,059
Jemeppe-sur-Sambre	31,936	17,761	12,043	0,764	13,077	21,685	5,605	4,699
Jodoigne	NA	15,724	25,027	1,380	29,739	36,333	11,744	9,897
Juprelle	32,008	13,246	20,205	1,118	7,676	31,348	1,554	40,308
Jurbise	NA	22,207	16,691	0,778	26,299	27,069	9,429	19,381
La Bruyère	24,965	29,324	8,669	0,789	3,501	27,269	7,624	14,502
La Calamine	NA	6,210	36,803	5,200	34,832	18,654	12,447	37,338
La Hulpe	NA	6,987	29,762	13,285	12,979	29,349	18,494	1,770
La Louvière	34,700	6,186	34,822	0,010	2,525	36,337	7,744	24,004
La Roche-en-Ardenne	NA	35,950	12,429	1,809	25,641	25,320	28,134	1,233
Lasne	NA	22,973	24,863	1,679	3,829	26,994	15,030	23,441
Le Roeulx	25,870	22,676	8,393	2,882	12,075	18,716	5,826	3,046
Léglise	25,505	28,266	5,670	11,180	26,650	7,334	9,088	2,701
Lens	NA	29,796	19,602	3,093	30,604	31,137	15,421	3,032
Les Bons Villers	22,101	15,242	16,380	34,821	10,304	13,725	9,141	4,108
Lessines	8,751	23,280	21,693	0,702	17,943	32,695	12,440	11,643
Leuze-en-Hainaut	NA	25,032	15,601	9,057	12,895	30,729	24,039	0,818
Libin	31,462	30,136	5,989	0,441	22,308	13,924	14,445	1,430
Libramont-Chevigny	17,811	21,788	18,945	0,947	24,411	21,742	11,366	29,731
Liège	40,252	4,164	38,749	0,304	1,263	33,167	5,985	11,506
Lierneux	32,507	26,485	11,310	0,111	9,002	29,781	7,200	8,856
Limbourg	NA	17,379	24,120	0,428	21,052	36,457	13,629	1,812
Lincent	17,862	25,435	10,026	6,476	9,013	21,848	14,700	2,449
Lobbes	NA	15,273	16,368	10,973	26,266	29,398	11,731	7,374
Lontzen	33,652	17,650	16,578	1,813	19,122	16,843	20,902	10,855
Malmedy	17,209	24,733	14,076	0,089	14,884	37,907	14,624	2,815
Manage	31,277	11,194	27,385	7,241	2,150	28,675	8,754	20,558
Manhay	29,641	29,291	8,615	1,472	31,573	12,419	2,276	4,409
Marche-en-Famenne	NA	20,457	22,156	1,245	18,023	35,073	15,421	12,816
Marchin	NA	20,250	23,095	1,736	15,827	36,580	9,531	38,651
Martelange	NA	27,787	10,016	32,906	16,774	8,120	31,307	0,982
Meix-devant-Virton	NA	29,093	16,403	3,077	26,064	31,422	17,625	1,236
Merbes-le-Château	NA	21,214	22,330	3,743	25,191	38,295	7,909	8,357
Messancy	24,240	17,038	6,864	2,033	40,030	22,359	6,597	2,574
Mettet	NA	27,464	11,037	1,824	34,773	32,583	14,141	7,664
Modave	NA	27,587	11,556	3,588	32,627	28,189	0,325	18,947
Momignies	NA	20,886	17,189	4,729	26,520	36,019	2,739	2,927
Mons	29,192	13,535	23,564	0,171	2,528	30,585	5,492	21,165
Mont-de-l'Enclus	NA	19,049	5,556	6,686	47,889	19,763	1,552	3,310

COMMUNE	AUTO	AGGL	H-AGGL	AUBE	NSEP	NAEP	PIETON	OBS-CH
Mont-Saint-Guibert	35,135	10,251	29,868	4,276	2,981	24,674	4,383	9,844
Montigny-le-Tilleul	NA	32,351	17,315	3,974	10,669	26,070	15,918	0,843
Morlanwelz	NA	7,078	35,856	10,989	14,627	21,082	11,937	11,503
Mouscron	33,256	6,560	29,873	1,373	8,527	29,740	3,807	11,622
Musson	NA	14,419	32,531	3,950	30,278	38,254	4,355	11,262
Namur	33,619	12,768	23,530	0,036	10,290	24,702	8,948	12,735
Nandrin	NA	7,210	2,435	4,170	37,273	27,579	11,330	19,075
Nassogne	NA	37,827	10,666	1,932	21,040	18,612	22,233	1,822
Neufchâteau	22,093	21,540	9,588	1,965	21,538	15,031	9,127	26,523
Neupré	NA	21,766	13,355	8,498	21,334	28,434	21,546	1,697
Nivelles	26,188	18,129	15,878	9,149	20,978	8,082	14,423	6,712
Ohey	NA	20,094	20,540	0,754	28,718	37,813	4,503	18,863
Olne	NA	24,189	21,131	2,269	26,222	35,797	20,176	2,605
Onhaye	NA	33,019	9,752	8,899	30,409	32,322	19,283	6,179
Oreye	38,308	25,395	17,330	2,165	36,644	11,005	4,076	8,045
Orp-Jauche	NA	15,585	29,484	0,873	32,358	37,077	6,898	2,690
Ottignies-Louvain-la-Neuve	29,896	16,141	20,246	0,859	18,479	18,009	8,637	29,931
Ouffet	NA	27,935	15,100	10,579	22,072	30,821	7,140	24,866
Oupeye	NA	4,440	32,222	3,897	25,932	27,690	2,865	19,093
Paliseul	NA	27,353	14,556	12,794	24,106	24,804	26,354	10,447
Pecq	NA	16,273	10,931	0,847	34,339	23,284	1,541	7,220
Pepinster	NA	11,376	27,669	0,065	23,507	38,607	10,176	14,382
Péruwelz	28,747	17,614	22,225	0,370	4,588	29,227	12,151	2,971
Perwez	26,046	28,332	8,889	1,918	11,515	29,028	28,766	7,777
Philippeville	NA	33,762	9,936	0,628	16,838	32,873	23,930	4,477
Plombières	NA	20,764	17,809	15,337	18,626	28,257	22,368	8,287
Pont-à-Celles	22,589	23,015	14,609	0,646	45,942	6,160	4,658	0,753
Profondeville	NA	32,537	14,511	2,586	21,122	28,081	21,126	10,101
Quaregnon	33,478	8,443	28,320	1,758	23,952	14,813	3,222	21,446
Quévy	25,361	16,722	6,154	0,418	26,976	28,718	3,969	9,110
Quiévrain	NA	10,573	25,719	9,075	35,433	14,641	19,023	5,870
Raeren	35,028	14,011	18,818	0,295	33,766	7,611	3,233	2,596
Ramillies	NA	17,575	13,059	25,791	28,284	25,781	5,070	4,633
Rebecq	29,358	15,168	11,292	8,465	19,865	36,562	5,632	12,688
Remicourt	32,652	21,825	10,765	0,594	34,233	16,173	6,817	2,276
Rendeux	NA	29,296	10,728	14,010	23,172	14,895	27,664	2,065
Rixensart	36,984	6,004	34,311	1,106	5,518	32,910	3,964	2,609
Rochefort	19,882	23,962	15,651	2,142	30,180	17,162	17,542	7,198
Rouvroy	NA	19,357	18,487	5,151	27,599	25,882	0,592	5,784
Rumes	NA	11,949	23,174	15,683	27,601	24,241	3,091	15,814
Saint-Georges-sur-Meuse	NA	26,816	9,527	32,610	27,975	7,061	4,982	27,994
Saint-Ghislain	29,697	17,331	19,266	3,230	0,935	26,905	3,449	6,206

COMMUNE	AUTO	AGGL	H-AGGL	AUBE	NSEP	NAEP	PIETON	OBS-CH
Saint-Hubert	29,298	22,024	15,500	0,403	3,042	30,765	21,187	1,677
Saint-Léger (Virton)	NA	27,010	14,822	29,403	17,860	16,324	13,965	28,594
Saint-Nicolas (Liège)	NA	25,388	14,946	2,245	24,350	37,928	11,136	28,710
Saint-Vith	41,091	2,162	46,640	1,542	7,087	27,049	4,613	4,871
Sainte-Ode	21,617	27,698	16,167	1,041	13,844	31,182	23,713	14,835
Sambreville	29,777	13,407	26,018	0,203	14,262	23,920	11,364	24,073
Seneffe	21,242	26,733	9,758	2,272	0,848	31,305	7,276	30,587
Seraing	35,182	5,366	32,541	0,397	10,638	24,048	4,238	28,592
Silly	25,589	23,619	5,115	2,028	16,943	25,429	7,117	19,245
Sivry-Rance	NA	27,035	15,187	2,208	26,551	34,211	21,034	3,627
Soignies	NA	18,878	28,508	0,088	18,101	36,971	11,367	15,860
Sombreffe	27,493	18,757	16,223	7,268	8,112	22,422	6,156	22,588
Somme-Leuze	NA	38,151	8,454	3,836	20,735	16,557	21,491	3,327
Soumagne	33,240	13,330	24,655	2,056	1,807	27,008	5,132	10,052
Spa	2,546	21,898	24,774	2,232	8,826	30,746	21,640	2,337
Sprimont	28,763	20,655	10,361	17,390	18,991	7,638	11,927	3,390
Stavelot	22,399	27,248	9,489	1,780	31,792	8,582	17,182	4,702
Stoumont	31,811	28,733	7,137	3,718	22,257	38,937	12,214	0,914
Tellin	24,424	33,502	7,614	3,502	25,352	8,560	19,465	2,297
Tenneville	NA	30,552	7,667	0,044	23,682	37,074	5,242	8,763
Theux	16,641	25,099	16,720	0,147	8,755	37,535	18,677	0,161
Thimister-Clermont	28,801	14,222	2,571	23,695	14,441	13,542	3,923	0,658
Thuin	NA	17,529	21,890	1,263	27,628	36,396	12,128	5,218
Tinlot	NA	35,179	7,575	2,678	10,901	37,666	16,967	0,866
Tintigny	NA	28,349	16,247	38,041	28,006	13,258	4,334	34,019
Tournai	29,373	15,777	17,538	3,315	6,759	35,267	7,468	3,137
Trois-Ponts	NA	28,103	15,284	41,466	15,173	3,212	10,318	30,994
Trooz	NA	11,107	31,272	5,279	11,119	32,441	15,544	0,846
Tubize	33,123	6,864	19,533	5,821	24,002	29,590	5,352	8,112
Vaux-sur-Sûre	23,998	23,669	7,830	1,347	12,841	22,598	16,611	0,863
Verlaine	22,586	23,099	13,352	7,272	29,265	16,163	8,932	0,692
Verviers	37,179	4,624	34,120	4,531	11,836	20,527	6,771	6,379
Vielsalm	15,477	25,651	18,184	3,752	12,879	35,066	18,951	2,474
Villers-la-Ville	NA	20,173	20,537	5,426	34,285	26,795	16,265	0,855
Villers-Le-Bouillet	31,391	29,659	10,052	0,397	0,361	28,846	5,585	8,730
Viroinval	NA	20,883	18,069	11,156	26,454	27,112	13,643	8,836
Virton	NA	19,784	22,938	2,712	32,648	35,666	8,315	8,183
Visé	23,633	17,494	18,092	0,524	7,489	25,045	5,049	6,149
Vresse-sur-Semois	NA	31,527	13,034	3,039	24,252	32,136	24,140	2,562
Waimes	NA	25,993	17,194	16,377	25,695	24,366	8,930	7,553
Walcourt	NA	32,446	14,148	4,677	15,329	13,747	28,697	1,062
Walhain	28,397	26,747	9,130	5,534	17,149	18,454	5,596	4,250
Wanze	31,102	15,157	18,197	0,376	11,444	34,812	3,888	3,558

COMMUNE	AUTO	AGGL	H-AGGL	AUBE	NSEP	NAEP	PIETON	OBS-CH
Waremme	28,459	19,297	16,765	4,708	3,880	29,233	13,153	10,408
Wasseiges	NA	14,109	30,039	3,701	32,215	36,053	2,747	3,630
Waterloo	35,004	7,043	32,525	1,686	2,186	34,114	3,092	17,710
Wavre	27,932	16,527	20,767	7,252	1,519	34,331	12,138	3,466
Welkenraedt	26,551	17,701	16,774	0,125	9,727	31,392	7,364	22,922
Wellin	25,464	22,227	11,630	8,048	27,750	15,351	3,828	19,343
Yvoir	39,442	23,864	16,020	5,936	23,697	10,650	7,796	4,230

3.3.5 Perspective suivante : *clustering* des communes

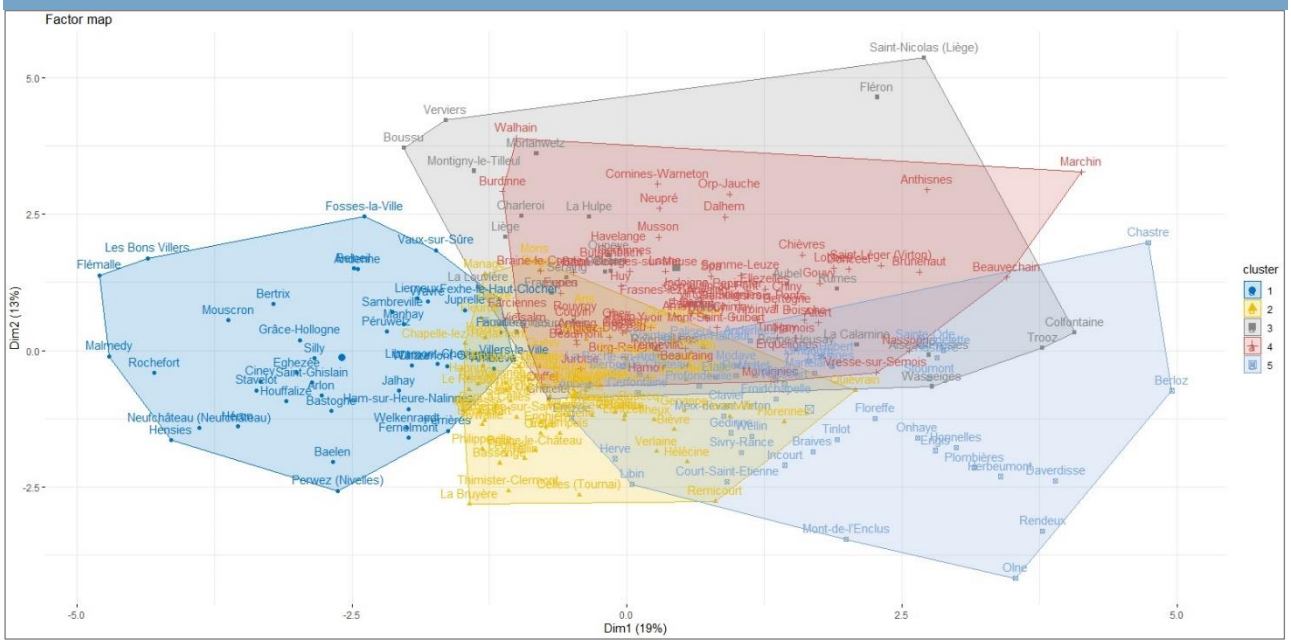
Une fois toutes ces informations compilées, via des analyses AFC et/ou ACM, un prolongement logique pourrait impliquer la recherche de *clusters* naturels de communes. Ce type d'analyse vise à mettre en avant des regroupements de communes qui auraient des profils semblables, ou ressemblants, du point de vue de la topologie et des causes supposées des accidents de la route qui y ont eu lieu.

Cet article n'a pas vocation à se consacrer pleinement à une telle recherche de *clustering* optimal. Le choix de la méthodologie employée, le choix du nombre de *clusters* à retenir et à interpréter, l'analyse de dendrogrammes ou l'éventuel ajout de contraintes territoriales de contiguïté sont autant de champs d'investigation qui peuvent rapidement devenir complexes et qui peuvent chacun faire l'objet d'une recherche séparée. De même, ces analyses de *clustering* peuvent être mises en œuvre directement sur les données brutes, sur la base des résultats d'une ACP (Analyse en Composantes Principales) préalable, d'une AFC ou d'une ACM.

À titre de simple illustration (cf. graphique 15), voici une partition obtenue sur la base des quatre premières composantes principales d'une Analyse en Composantes Principales (ACP) préalable, en utilisant un algorithme hiérarchique de Ward (voir Murtagh et Legendre, 2011) et en retenant arbitrairement cinq *clusters*.

Notons clairement que les choix finaux de l'algorithme de *clustering* et du nombre de *clusters* à retenir et à interpréter peuvent être considérés comme des sujets d'étude à part entière, à la fois selon les critères d'arrêt de l'algorithme mis en œuvre, le niveau de considération d'éventuelles données aberrantes et les objectifs propres à l'analyse.

Graphique 15 : Exemple de *clustering* basé sur les quatre premières composantes d'une ACP et projection sur les deux premiers axes



CHAPITRE 4

PISTES D'AIDE À LA PRISE DE DÉCISION

Avant de conclure, les méthodes statistiques mises en pratique dans les sections précédentes et les résultats obtenus peuvent participer à l'élaboration de pistes de réflexion en combinaison, et en complémentarité, avec d'autres approches de terrain et d'autres actions de prévention/sensibilisation déjà entreprises[9]. L'objectif est alors de proposer ces pistes qui peuvent aider à une prise de décision au niveau local (selon les leviers d'action régulièrement évoqués dans ce travail).

Ainsi, en regroupant les résultats présentés dans les sections 3.2 et 3.3 (ainsi que dans le tableau 1), nous pouvons relever que :

- les communes d'Herbeumont, Saint-Léger, Attert et Ohey ressortent de manière significative dans des contextes d'accidents sur chaussée sale ;
- les communes de Jalhay, Manhay, Bullange, Bertogne et Martelange ressortent de manière significative dans des contextes de conditions hivernales (chaussées enneigées ou verglacées). Selon les configurations, les leviers possibles peuvent impliquer un questionnement sur le traitement des chaussées en tant que tel, sur une signalisation adaptée, sur un plus grand contrôle des vitesses dans des conditions hivernales difficiles ou sur une prévention générale plus poussée ;
- les communes de Vresse-sur-Semois, Stoumont, Rumes et Hamoir ressortent de manière significative dans des contextes d'accidents ayant lieu les week-ends. Ici, les leviers iraient vers une maîtrise accrue des effets « week-end » les plus significatifs et vers une régulation d'un trafic manifestement beaucoup plus intense le week-end, avec probablement davantage de contrôles, de radars et/ou de tests d'alcoolémie ;
- les communes de Pecq, Donceel, Nandrin, Momignies, Brunehaut, Wasseiges et Juprelle ressortent de manière significative pour des contextes d'accidents ayant un effet cumulé week-end + nuit ;
- les communes de Herbeumont, Paliseul, Gedinne, Ouffet, Saint-Hubert, Fernelmont, Mont-de-l'Enclus, Fauvillers, Pont-à-Celles, Attert, Amay et Aubel ressortent de manière significative pour des contextes d'accidents liés à une circulation de nuit. Des leviers possibles impliqueraient un questionnement sur l'état général des éclairages (plus d'éclairage ou amélioration des éclairages existants), sur la signalisation au sol ou par davantage de contrôles ;
- les communes de Beyne-Heusay, Liège, Saint-Nicolas, Celles, Yvoir, Boussu, Saint-Vith, Héron, Lincent et Berloz ressortent de manière significative pour des contextes d'accidents liés aux tronçons autoroutiers ;
- les communes de Floreffe, Binche, Gembloux, Martelange, Braine-le-Comte, Liège, Verviers, Fléron, Frameries, Doische et Walcourt ressortent de manière significative pour des contextes d'accidents impliquant des piétons.

Rappelons une dernière fois qu'il s'agit ici de pistes de réflexion et de questionnement. Les diagnostics finaux ne peuvent être émis, confirmés ou infirmés que par des agents de terrain pouvant être davantage au fait des particularités locales.

[9] Par exemple ici : http://www.tousconcernes.be/campagnes/

CHAPITRE 5

CONCLUSIONS

Lorsqu'un accident de la circulation, mortel ou non, survient, il est évidemment important d'encoder ses caractéristiques, ses circonstances ou ses indices de causes supposées. Ces informations recueillies par la police locale sont par nature imparfaites et régulièrement incomplètes, et aboutissent à la création dans un fichier de données de variables catégorielles (pouvant contenir un nombre variable de modalités, des non-observations, etc.).

Le premier objectif de cette publication était de fournir un état des lieux, une sorte de rapide *vade-mecum*, afin de comprendre comment traiter statistiquement de telles variables catégorielles, de façon tout à fait générale. Partant de la notion d'indépendance entre variables catégorielles, des méthodes univariées (AFC) et multivariées (ACM) ont été montrées, décrites et interprétées (principalement dans l'annexe 1).

L'application directe de ces méthodes aux données d'accidents de la route (mortels ou non, au niveau des communes wallonnes, pour les accidents recensés entre 2009 et 2018) a fourni des sorties graphiques diverses : représentations univariées, images instantanées dans une commune à la fois, représentations cartographiques des contributions des modalités des variables… Tous ces *outputs* sont donc disponibles et fournissent tant des informations absolues que relatives.

La partie analytique de ce travail nous a montré que les analyses univariées (AFC) et multivariées (ACM) pouvaient dans certains cas aller dans le même sens, confirmant ainsi des cas particuliers pour lesquels des leviers d'action pourraient être mis en pratique pour améliorer la sécurité des usagers. Dans d'autres cas, on a vu que les communes mises en avant pour un même facteur, pour une même cause supposée, peuvent être largement différentes selon que l'on travaille en univarié ou en multivarié. Le point de vue dépend donc très clairement des données sur lesquelles on travaille : si l'on travaille sur plusieurs données simultanément, l'influence d'une variable en particulier sera forcément altérée.

De façon concrète, l'ensemble de la partie analytique (cf. chapitre 3) a contribué à étudier ce qu'il se passe au niveau communal du point de vue de nombreux facteurs qui caractérisent les accidents constatés : l'état de la chaussée, le type de route (autoroute, route communale…), la luminosité (dont l'absence ou le défaut d'éclairage la nuit), le type de collision (dont avec un piéton), mais aussi la période de la semaine ou la vitesse maximale autorisée sur le lieu de l'accident. Toutes ces variables sont détaillées dans le chapitre 2. Ainsi, pour chacun de ces facteurs, de façon univariée d'abord puis en multivarié pour certains d'entre eux, le lecteur a l'occasion de comprendre la théorie et de visualiser quelles communes se distinguent dans chacun des cas. Naturellement, certaines constatations peuvent ensuite être reprises à un niveau décisionnel afin de mettre en œuvre des actions concrètes et adéquates.

À titre de perspectives supplémentaires, il serait important de pouvoir disposer de plus de variables simultanément au niveau multivarié (pour les ACM). Disposer de variables avancées de façon simultanée n'est pas la même chose que de disposer de plusieurs tableaux de contingence (qui sont par nature indépendants entre eux *a priori*). Pour cela, il faudrait que les techniques d'encodage des accidents sur le terrain soient plus riches. Cela soulève évidemment d'autres problématiques comme les outils d'encodage disponibles, les outils de géolocalisation, le temps nécessaire à l'encodage que doit prendre l'agent sur le terrain ou encore le niveau de subjectivité dans l'encodage des circonstances de l'accident.

Toujours pour aller plus loin dans l'analyse globale liée à la sécurité routière, il sera intéressant de pouvoir établir et construire des indicateurs communaux, probablement composites, sur la base des différents résultats présents dans ce travail (aussi bien au niveau des AFC que des ACM). À ce titre, le tableau 1, qui récapitule les principales contributions cumulées des ACM commune par commune, pourra servir de base de réflexion.

En outre, une éventuelle analyse qui permettrait de mettre en avant des facteurs qui puissent différencier et discriminer les accidents mortels des accidents avec blessés pourrait amener un plus. Cela reviendrait à considérer la gravité des accidents comme un nouveau facteur bien particulier, qui s'expliquerait par rapport aux autres facteurs.

Enfin, comme brièvement abordé dans la section 3.3.5, une autre perspective intéressante est de récupérer les résultats statistiques des méthodes abordées afin d'en construire un *clustering* de communes et ainsi de profiler les communes selon leurs ressemblances.

CHAPITRE 6

RÉFÉRENCES

Benzecri J. P., 1973. L'Analyse des correspondances. Paris : Dunod.

Brunet S., O'Dorchai S. et Reginster I., 2020. Où en est la Wallonie par rapport aux objectifs de développement durable ? Bilan des progrès. IWEPS.

Escofier B., 1979. Une représentation des variables dans l'analyse des correspondances multiples. Revue de Statistique Appliquée, 27(4), pp. 37-47.

Greenacre M. et Blasius J., 2006. Multiple correspondence analysis and related methods. CRC press.

Greenacre M., 2017. Correspondence analysis in practice. CRC press.

Härdle W. K. et Hlavka Z., 2015. Correspondence analysis. Multivariate statistics, Springer (pp. 259-280).

Heckler C. E., 2005. Applied multivariate statistical analysis.

La Rocca S. et Masuy A., 2019. Rapport méthodologique de l'enquête MOBWAL 2017. Rapport de Recherche n°36, IWEPS.

Le Roux B. et Rouanet H., 2004. Geometric data analysis : from correspondence analysis to structured data analysis. Springer Science & Business Media.

Masuy A., Juprelle J. et La Rocca S., 2018. MOBWAL et GPSWAL, deux enquêtes complémentaires pour mieux appréhender les comportements de mobilité de la population wallonne, WP n°24, IWEPS.

Murtagh, F., & Legendre, P. (2011). Ward's hierarchical clustering method: Clustering criterion and agglomerative algorithm. arXiv preprint arXiv:1111.6285.

Fiches métadonnées Statbel et Walstat :

https://statbel.fgov.be/fr/themes/mobilite/circulation/accidents-de-la-circulation#documents

https://walstat.iweps.be/fichiers/metadonnees/meta-217400.pdf

CHAPITRE 7

ANNEXES

ANNEXE 1 : CADRE MÉTHODOLOGIQUE ÉLARGI

D'un point de vue strictement méthodologique, ce travail est orienté vers la prise en main et l'analyse statistique de variables catégorielles possédant au moins deux catégories (ou niveaux). De façon la plus intuitive possible, et sans rentrer dans trop de détails mathématiques (qui seront disponibles via le renvoi vers quelques références), les problématiques sont présentées par ordre croissant de complexité dans les sections suivantes. À chaque fois, un lien sera fait entre l'outil envisagé et les questions auxquelles il est susceptible de répondre.

A. Variables catégorielles et principe d'indépendance

Il est toujours bon de rappeler en tout premier lieu ce qu'est concrètement une **variable catégorielle**, même si cela peut paraître évident. En statistique, une variable catégorielle, qui est qualitative par construction et qui est aussi régulièrement appelée **facteur** (ex. : le facteur luminosité), est une variable qui prend pour valeur des **modalités**, **catégories** ou **niveaux** (ex. : le niveau « nuit sans éclairage »), par opposition aux variables quantitatives qui mesurent sur chaque individu une quantité continue ou discrète (ex. : le nombre de kilomètres parcourus). Dans le cadre de ce travail, les variables exploitées, décrites dans le chapitre 2, seront nominales. Cela signifie qu'il n'y aura pas de notion d'ordre ou d'échelle de valeur parmi les niveaux de ces variables. Dans le cas contraire, nous parlerions de variable ordinale. Notons que seule la variable « vitesse maximale autorisée » peut avoir une connotation ordinale, connotation toutefois nuancée si l'on considère que cette vitesse est directement liée à des types de voirie différents.

Il faut noter que, dans le cadre précis d'utilisation des analyses factorielles des correspondances, le nom des individus (ici le nom des communes wallonnes) peut être davantage qu'un simple identifiant. Il peut tout aussi bien jouer le rôle d'une variable catégorielle de fond. Autrement dit, dans ce contexte, les traditionnelles lignes et colonnes de la table statistique sont parfaitement interchangeables, et il n'y a pas de notion stricte de variables *versus* individus.

Nous passons rapidement sur les analyses d'une seule variable catégorielle à la fois, ou analyses univariées, qui ne permettent pas d'établir de quelconques conclusions prédictives. L'intérêt des analyses univariées repose en grande partie sur la visualisation des distributions des variables catégorielles prises une à une. Le choix entre différentes sorties graphiques se fera alors, par exemple, sur la base du nombre de niveaux (ou catégories) de la variable concernée. Les diagrammes en barres ou les diagrammes circulaires (ou diagrammes en secteurs, camemberts ou tartes…) seront alors d'application. Numériquement, le mode mettra à jour la catégorie la plus représentée et jouera le rôle de la tendance centrale de la distribution, pour autant qu'il y ait une ébauche de symétrie autour de cette tendance centrale.

La problématique devient davantage intéressante dès lors que l'on ajoute une seconde variable catégorielle. Le champ des analyses possibles devient plus fourni et permettra idéalement de mettre à jour des liens entre ces deux variables. Existe-t-il un lien, ou une relation, entre ces deux variables ? Est-ce qu'elles interagissent ? Comment interagissent-elles ? Quelle est la force de cette interaction ? Voilà des exemples de questions auxquelles il sera possible de répondre.

Très rapidement, ce questionnement sur un possible lien entre deux variables catégorielles va impliquer la notion d'**indépendance**. En statistique et en théorie des probabilités, l'indépendance est une notion qui qualifie de manière intuitive des évènements aléatoires n'ayant aucune influence l'un sur l'autre. Par exemple, dans un jeu non biaisé, la valeur d'un premier lancé de dé n'a pas d'influence sur la valeur d'un second lancé de dé. De même, le résultat de ce lancé sera soit pair, soit impair. Les deux évènements « avoir un résultat pair » et « avoir un résultat impair » sont donc indépendants. Évidemment, l'indépendance ou non de deux évènements, ou entre deux variables catégorielles, n'est pas toujours facile à établir. La définition mathématique de l'indépendance de deux évènements A et B est la suivante : $P(A \cap B) = P(A).P(B)$. La probabilité jointe de A et B simultanément est le produit des probabilités de A et de B

séparément. Si la probabilité de *B* n'est ni nulle ni égale à un, une autre reformulation dit que *A* et *B* sont indépendants si : $P(A|B) = P(A)$. Avec l'apparition de cette probabilité conditionnelle, cela signifie que connaître *B a priori* n'influe en rien sur l'occurrence ou pas de l'évènement *A*.

Dans ces conditions, où l'indépendance entre deux variables n'a souvent rien d'évident, le recours à un test statistique formel devient nécessaire afin de s'assurer des deux égalités ci-dessus. Tester l'indépendance entre deux variables catégorielles revient alors à mesurer l'écart entre ce que l'on observe réellement dans les données et ce que l'on s'attend à observer dans une situation théorique d'indépendance. Et c'est ainsi qu'apparaissent les notions d'**effectifs observés** et d'**effectifs attendus**. Pour mieux visualiser, soient *A* et *B* deux variables qualitatives ayant respectivement *I* et *J* modalités. Soit *N* le nombre d'individus sur lesquels *A* et *B* ont été observées. On note n_{ij} le nombre d'individus qui possèdent à la fois la modalité *i* de la variable *A* et la modalité *j* de la variable *B*. Ceci permet de remplir le tableau A1 suivant. Ce tableau, qui croise deux variables catégorielles, s'appelle communément **tableau de contingence**.

Tableau A1 : Effectifs observés

	B_1	...	B_j	...	B_J	total
A_1	n_{11}	...	n_{1j}	...	n_{1J}	$\boldsymbol{n_{1\bullet}}$
⋮	⋮		⋮		⋮	⋮
A_i	n_{i1}	...	n_{ij}	...	n_{iJ}	$\boldsymbol{n_{i\bullet}}$
⋮	⋮		⋮		⋮	⋮
A_I	n_{I1}	...	n_{Ij}	...	n_{IJ}	$\boldsymbol{n_{I\bullet}}$
total	$\boldsymbol{n_{\bullet 1}}$	...	$\boldsymbol{n_{\bullet j}}$	...	$\boldsymbol{n_{\bullet J}}$	$N = \boldsymbol{n_{\bullet\bullet}}$

Il est à noter que si l'on divise chacune des cases de ce tableau de contingence par *N*, le nombre total d'individus considérés, l'étude sur la base d'effectifs passe sans perte d'information à une étude sur la base de fréquences (transformables en des pourcentages). Le grand total en bas à droite deviendrait alors 1 (ou 100%).

Les cases de ce tableau de contingence ainsi définies, les totaux en ligne (pour chacune des modalités de *A*) et les totaux en colonnes (pour chacune des modalités de *B*) vont servir à calculer les effectifs attendus sous hypothèse d'indépendance mentionnés plus haut. Il faut alors imaginer un second tableau de mêmes dimensions que le tableau A1 dans lequel chaque case, chaque effectif attendu donc, sera calculée comme suit, pour tout *i* et pour tout *j* :

$$e_{ij} = \frac{n_{i\bullet} n_{\bullet j}}{n_{\bullet\bullet}}$$

Dorénavant, chaque effectif observé n_{ij} peut être directement mis en comparaison avec son pendant attendu e_{ij}. Intuitivement, il semble logique que l'hypothèse d'indépendance entre les deux variables ne sera pas rejetée si tous les n_{ij} sont suffisamment proches des (ou, dans un cas très extrême, égaux à) e_{ij}. Dès lors, par symétrie, un lien entre les deux variables sera mis à jour si certains écarts sont jugés significatifs entre les observés et

les attendus, mettant de ce fait à mal l'hypothèse d'indépendance. Ces intuitions sont confirmées par la statistique de test généralement utilisée dans le cas d'un **test d'indépendance** (dont l'hypothèse nulle est l'indépendance et l'hypothèse alternative est l'existence d'un lien) qui est la suivante :

$$Q = \sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\left(n_{ij}-e_{ij}\right)^2}{e_{ij}} \sim \chi^2_{(I-1)(J-1)}$$

Si des écarts existent, cette statistique Q, qui suit une loi chi-carrée à $(I-1)(J-1)$ degrés de liberté sous hypothèse nulle, va rapidement augmenter (d'autant plus dans la mesure où les écarts y sont mis au carré). Et plus Q est grande (ou, alternativement, plus la p-valeur correspondante à Q est petite), plus il sera évident que la décision à prendre ira dans le sens d'un rejet de l'hypothèse d'indépendance. De ce fait, un lien entre les deux variables catégorielles sera mis à jour.

Bien que potentiellement très intéressante en soi, cette conclusion peut paraître insuffisante, voire frustrante. À ce niveau de l'étude, on sait éventuellement qu'il existe un lien et… c'est tout. Pour quantifier la force de ce lien, il existe certaines mesures bien connues comme le *V* de Cramer qui permettent d'aller plus loin. Mais il est sans doute encore plus intéressant d'observer et de comprendre la contribution de chaque modalité à la décision que l'on vient de prendre. Pourquoi avons-nous rejeté l'hypothèse nulle d'indépendance ? Sur quelles bases ? À cause de quels individus en particulier ? Qui ont quelles caractéristiques ? Voilà quelques exemples de questions qui pourront être abordés par une Analyse Factorielle des Correspondances (AFC).

B. Analyse factorielle des correspondances simples (AFC)

Une **Analyse Factorielle des Correspondances (AFC)** simple peut être mise en œuvre directement sur la base d'un tableau croisé, ou de contingence, dès lors que l'on croise deux variables catégorielles. Pour rappel, l'une de ces deux variables peut être l'identifiant des individus dans la mesure où l'on peut s'affranchir de la traditionnelle distinction lignes *versus* colonnes. Les extensions, généralisations, utilisations particulières de l'AFC dans diverses disciplines sont tellement nombreuses qu'établir la liste des approches n'est plus un exercice raisonnable. Cependant, le schéma de dualité lignes/colonnes permet de faire un lien commun entre toutes ces approches possibles.

L'objectif principal est ici de rechercher la meilleure représentation simultanée de deux ensembles constituant les lignes et les colonnes d'une table de contingence. Ces lignes et ces colonnes sont ainsi deux partitions distinctes des mêmes *N* individus et doivent jouer des rôles parfaitement symétriques. On raisonne en termes de **profils (profils ligne ou profils colonne) :** les modalités d'une même variable sont alors comparables entre elles et les proximités entre les points s'interprètent comme des similitudes de profils.

B.1. Dualité des profils et distances entre points

Sans entrer totalement dans un cours théorique (voir Benzecri, 1973, Heckler, 2005, Härdle et Hlavka, 2015 ou Greenacre, 2017), il est important de bien cerner cette notion de profil et de visualiser à quel point l'AFC s'appuie non plus sur les variables initiales en tant que telles mais sur leurs modalités (ou niveaux) respectives. Partons des fréquences conjointes f_{ij} et marginales $f_{i\bullet}$ et $f_{\bullet j}$ qui se calculent comme ceci :

$$f_{ij} = \frac{n_{ij}}{N} \qquad f_{i\bullet} = \frac{n_{i\bullet}}{N} \qquad f_{\bullet j} = \frac{n_{\bullet j}}{N}$$

Dans le nuage de points des *I* lignes, chaque point se retrouve avec une coordonnée :

$$\left[\frac{f_{ij}}{f_{i\bullet}}\right]$$

Chaque point est ainsi affecté du poids $f_{i\bullet}$, sa fréquence relative. Il est montré que le centre de gravité de ce point correspond à la fréquence marginale des colonnes, tel que :

$$\sum_{i=1}^{I} f_{i\bullet}\frac{f_{ij}}{f_{i\bullet}} = f_{\bullet j}$$

Le raisonnement est rigoureusement le même pour les points des *J* colonnes, de façon à obtenir :

$$\sum_{j=1}^{J} f_{\bullet j} \frac{f_{ij}}{f_{\bullet j}} = f_{i\bullet}$$

C'est cette dualité entre les lignes et les colonnes, cette relation étroite (on parle de relations quasi barycentriques), qui permet de poser un critère d'ajustement pour l'AFC et de comprendre les distances qui peuvent exister entre les points.

En effet, le critère cible d'une AFC vise à maximiser la somme pondérée des carrés des distances entre les points et le centre de gravité du nuage. Il faut pour cela définir cette notion de **distance**. La distance euclidienne entre deux points *i* et *i'* exprimée sur le tableau des effectifs initiaux traduit simplement la différence d'effectifs entre ces deux modalités de la variable *A*. Par contre, la distance euclidienne entre deux points *i* et *i'* exprimée sur le tableau des profils ligne traduit la **ressemblance** ou la **différence** entre les deux modalités de la variable *A* :

$$d^2(i,i') = \sum_{j=1}^{J} \left(\frac{f_{ij}}{f_{i\bullet}} - \frac{f_{i'j}}{f_{i'\bullet}} \right)^2$$

Cette distance favorise les colonnes qui ont un poids $f_{\bullet j}$ important. Il est ainsi bon de pondérer chaque écart par l'inverse du poids de la colonne. On obtient une nouvelle distance :

$$d^2(i,i') = \sum_{j=1}^{J} \frac{1}{f_{\bullet j}} \left(\frac{f_{ij}}{f_{i\bullet}} - \frac{f_{i'j}}{f_{i'\bullet}} \right)^2$$

Afin de généraliser ces distances à l'ensemble d'un tableau de contingence, il faut dorénavant passer en écritures matricielles. On note **F** le tableau complet des fréquences f_{ij}, $\mathbf{D_I}$ et $\mathbf{D_J}$ les matrices diagonales telles que $\mathbf{D_I} = Diag(f_{1\bullet}, \dots, f_{I\bullet})$ et $\mathbf{D_J} = Diag(f_{\bullet 1}, \dots, f_{\bullet J})$. Soit aussi **L** le tableau :

$$\mathbf{L} = \mathbf{D_I}^{-1}\mathbf{F} = \left[\frac{f_{ij}}{f_{i\bullet}} \right]$$

Un tableau centré serait obtenu comme ceci : $\mathbf{L_0} = \mathbf{L} - \mathbf{1}_{IJ}\mathbf{D_J}$

Ainsi, à chaque point est associée sa fréquence relative comme poids $f_{i\bullet}$ ($\mathbf{D_I}$). Et pour ne pas favoriser les colonnes ayant un poids $f_{\bullet j}$ trop important, il faut retranscrire au niveau matriciel l'étape vue plus haut, c'est-à-dire une pondération de l'écart entre deux points i et i' par l'inverse de ce poids ($\mathbf{D_J}^{-1}$).

B.2. Recherche des axes principaux (composantes principales)

À ce niveau de l'AFC, une fois les principaux éléments définis, la méthodologie devient projectionniste à la façon d'une Analyse en Composantes Principales (ACP). Il va ainsi s'agir de proposer des axes qui permettront successivement d'ordonner les lignes et les colonnes. En d'autres mots, les coordonnées des lignes (colonnes) permettent une première ordination des lignes (colonnes) sur l'axe 1. On identifiera alors grâce à ces ordinations des premières associations interprétables entre lignes et colonnes. L'axe 2 et les suivants proposeront des nouvelles ordinations et d'autres associations, etc.

Comme dans une ACP classique, les axes obtenus seront orthogonaux. Cela signifie que les structures successivement identifiées sont indépendantes entre elles (il n'existe pas d'informations redondantes entre les axes successifs).

De façon pratique, et sans de nouveau rentrer dans trop de détails mathématiques, les matrices définies dans la section B.1 sont combinées afin de former une seule nouvelle matrice **H**, telle que :

$$\mathbf{H} = \mathbf{D_J}^{-1/2}\mathbf{L_0}^T\mathbf{D_I}\mathbf{L_0}\mathbf{D_J}^{-1/2}$$

H est la matrice qui sera soumise au processus de décomposition spectrale lors du déroulement de l'AFC. Comme dans toute méthode de projection qui vise à transformer des éléments liés entre eux (ou corrélés) en nouvelles variables artificielles indépendantes, l'AFC implique la construction de **vecteurs propres** et de **valeurs propres**.

Différentes méthodes de décomposition, ou de diagonalisation, existent dans la pratique, certaines plus intuitives que d'autres. Leur but commun est de mettre à jour une suite de valeurs propres non nulles et d'en conserver les *K* premières, dans un ordre décroissant, dans $\Lambda_K = Diag(\lambda_1, \dots, \lambda_K)$, ainsi

que de conserver en colonne les *K* premiers vecteurs propres associés, orthonormés, dans une matrice $\mathbf{U}_K$.

$\mathbf{U}_K$ possède *J* lignes et *K* colonnes et vérifie l'égalité $\mathbf{U}_K^T\mathbf{U}_K = \mathbf{Id}_K$.

Sur la base des vecteurs et valeurs propres, on doit retrouver maintenant la propriété qui veut que $\mathbf{H} = \mathbf{U}_K\mathbf{\Lambda}_K\mathbf{U}_K^T$.

Finalement, les coordonnées des colonnes se trouvent dans : $\mathbf{D}_\mathbf{J}^{-1/2}\mathbf{U}_K\mathbf{\Lambda}_K^{1/2}$.

Et les coordonnées des lignes se trouvent dans : $\mathbf{L_0}\mathbf{D}_J^{-1/2}\mathbf{U}_K$.

L'examen des valeurs propres permet de déterminer le nombre d'**axes principaux** à retenir, sans perte d'information utile et sans accorder trop d'importance au bruit. Les valeurs propres correspondent à la quantité d'information (l'inertie) retenue par chacun des axes correspondants. Il n'y a pas de règle générale pour choisir le nombre de dimensions à conserver pour interpréter les résultats. Selon les domaines, une inertie totale de 90% sera exigée (en médecine par exemple) alors qu'un pourcentage bien inférieur sera acceptable par ailleurs.

Techniquement, et momentanément sans prétention pédagogique, les coordonnées des lignes fournies par cette procédure générale sont donc celles de l'analyse d'inertie du nuage des *I* distributions conditionnelles par ligne, pondéré et centré par la distribution marginale des lignes utilisant la métrique des inverses des poids des colonnes. Il s'en suit par symétrie que les coordonnées des colonnes sont celles de l'analyse d'inertie du nuage des *J* distributions conditionnelles par colonne, pondéré et centré par la distribution marginale des colonnes utilisant la métrique des inverses des poids des lignes.

Enfin, de façon comparable à une ACP (mais pas totalement puisque l'ACP se construit au niveau des variables alors que l'AFC se construit sur la base des modalités des variables catégorielles), notons enfin que chaque axe (ou composante principale) conservé à la fin d'un processus d'AFC s'écrit comme une combinaison linéaire des éléments mis à disposition initialement : les niveaux des variables catégorielles. Ceci permet les interprétations très précises discutées dans la section B.3 qui sont mises en pratique dans la section 3.2.

B.3. Interprétations graphiques et numériques

Grâce aux liens étroits entre profils ligne et profils colonne, et aux propriétés duales de l'AFC (relations quasi barycentriques entre lignes et colonnes), toutes les modalités des deux métavariables parviendront à être représentées sur un seul et unique plan. Ce plan sera celui construit à partir des deux premières composantes principales, ou axes principaux, de l'AFC.

La projection d'un profil ligne (respectivement colonne) est alors, à une constante près, la moyenne pondérée des projections de tous les profils colonne (respectivement ligne).

Par conséquent :

- Des profils ligne (respectivement colonne) proches entre eux seront interprétables comme des profils semblables, ayant un certain nombre de caractéristiques en commun ;
- Des profils ligne (respectivement colonne) éloignés du centre de la représentation, c'est-à-dire de l'origine 0, seront interprétables comme des profils éloignés du « profil moyen ». Il est important ici de faire le lien avec l'hypothèse d'indépendance évoquée plus tôt. Il a été dit que l'AFC allait plus loin que le test d'hypothèse d'indépendance dans le sens où il devenait maintenant possible de comprendre pourquoi cette hypothèse était amenée à être rejetée, sur quelles bases, à cause de quels individus ou à cause de quelles modalités. Et c'est exactement de ça qu'il s'agit ici : si un lien existe (rejet de l'hypothèse d'indépendance), il y aura obligatoirement des profils, ligne et/ou colonne, qui apparaîtront et qui se mettront en évidence par rapport aux autres. Ils seront ainsi différents du « profil moyen » ;
- Le cas impliquant un profil ligne (par exemple un individu ou une commune dans le cas présent) et un profil colonne (variable) éloignés chacun du centre de la représentation et proches l'un de l'autre sera l'expression d'une

association forte positive. Là aussi, il est intéressant de revenir en arrière et de faire un rapprochement avec la comparaison des effectifs observés et des effectifs attendus sous hypothèse d'indépendance. Si des écarts significatifs existent entre ces effectifs, nous avons compris que cela entraîne un rejet de l'indépendance, et donc la mise à jour de l'existence d'un lien. Une association forte positive est directement reliée à un tel écart significatif. Mais, plus précisément, il s'agira d'un cas où l'observé est significativement supérieur à l'attendu. Autrement dit, d'un cas de **surreprésentation** ;

- Le cas impliquant un profil ligne et un profil colonne éloignés chacun du centre de la représentation et également éloignés entre eux sera l'expression d'une **association forte négative**. Par le même raisonnement, une telle association mettra en avant un cas où un effectif observé est significativement inférieur à l'effectif attendu correspondant. Autrement dit, un cas de **sous-représentation**.

Ces interprétations, graphiques et immédiates, sont directement visualisées, mises en avant et interprétées dans la section 3.2 dédiée à l'analyse des causes des accidents de la route en Wallonie (sur la base des données décrites à la section 2.1).

De façon davantage numérique, les principaux logiciels statistiques proposent différents *outputs* sur lesquels s'appuyer pour mieux comprendre encore ces interactions entre modalités. Il s'agit là :

- Des **valeurs propres** mentionnées plus haut, qui permettent de choisir le nombre d'axes à retenir et d'évaluer ainsi le pourcentage d'inertie, ou de variance expliquée, qui est atteint avec ce nombre précis de dimensions ;
- Des **coordonnées**, calculées par le modèle, des points dans le plan des deux premiers axes principaux ;
- Des **cosinus carrés** (cos^2) ou mesures de la qualité de représentation des lignes et des colonnes. En choisissant le nombre de dimensions du problème grâce à l'analyse des valeurs propres, on sait quelle est la part de l'inertie totale contenue dans les données. Cependant, durant la phase de décomposition spectrale, tous les points ne sont pas aussi bien représentés dans ces dimensions (et notamment les deux premières, les plus importantes, qui servent à construire les sorties graphiques). Le cosinus carré mesure le degré d'association entre les lignes/colonnes et un axe particulier. Les valeurs de cos^2 sont toujours comprises entre 0 et 1 et la somme des cos^2 pour les lignes/colonnes sur toutes les dimensions de l'AFC est égale à 1. Sachant ceci, si un point ligne/colonne est bien représenté par les deux premières dimensions, la somme des deux cos^2 correspondants sera déjà proche de 1. Cette somme des deux premiers cos^2 est ainsi un chiffre clé à toujours vérifier ;
- Les **contributions** des lignes/colonnes (en %) dans le calcul des dimensions. Les lignes/colonnes qui ont des valeurs élevées contribuent le mieux à la construction des dimensions. Autrement dit, les lignes/colonnes qui contribuent le plus dans les deux premières dimensions sont les plus importantes pour expliquer la variabilité dans le jeu de données (ou tableau de contingence). C'est donc sur ces lignes/colonnes qu'il faut insister pour interpréter les résultats. Par contre, les lignes/colonnes qui ne contribuent pas beaucoup à aucune dimension ou qui contribuent aux dernières dimensions uniquement ont moins d'intérêt.

Concernant ces contributions, il est utile de faire une distinction entre les contributions absolues et les contributions relatives. Les contributions absolues déterminent directement les variables qui « tirent » un axe, ou le déploient. Comme lors de l'interprétation des associations fortes positives ou négatives, plus les points sont loin sur l'axe plus ils contribuent de manière absolue. En régression, une analyse équivalente définirait un « effet levier ». La somme des contributions absolues de toutes les modalités sur un axe vaut 1 (100%).

Les contributions relatives quant à elles se réfèrent à la qualité de représentation sur tous les axes en

même temps. La somme sur tous les axes des valeurs absolues des contributions relatives d'une modalité vaut 1.

Pour parfaitement comprendre ce qu'il se passe axe par axe, dans un ordre croissant d'importance, une analyse des contributions absolues est sans doute préférable.

Certains tableaux portant sur l'analyse des accidents de la route en Wallonie reprenant coordonnées, cosinus carrés et contributions absolues sont reportés en annexe et interprétés dans la section 3.2.

C. Analyse factorielle des Correspondances Multiples (ACM)

L'Analyse factorielle des Correspondances Multiples (ACM, Escofier, 1979) est le prolongement naturel de l'AFC dans le cas où plus de deux variables catégorielles sont considérées. L'AFC était basée sur un tableau de contingence entre deux variables catégorielles, ou entre des individus et une variable catégorielle. Ici, avec l'ACM, nous retrouvons le cadre classique avec une base de données dans laquelle les individus retrouvent leur rôle strict d'identifiants (en lignes) et dans laquelle les variables catégorielles, sans *a priori* sur leur nombre, se retrouvent en colonnes. Comme auparavant, chacune de ces variables catégorielles comporte plusieurs modalités, ou niveaux.

C.1. Généralisation à plus de deux variables catégorielles

Le cadre théorique général de l'AFC vu à la section B.2 ne change pas radicalement en passant à l'ACM. L'objectif est de s'adapter à un plus grand nombre de variables catégorielles à traiter simultanément. Lorsqu'une ACM est mise en œuvre, la base de données initiale qui comporte *p* variables catégorielles (ou discrètes ou discrétisées) est implicitement transformée en une **table disjonctive complète** (notée *Z*). Cette table *Z* contient alors *M* variables binaires, avec :

$M = \sum_{j=1}^{p}(I_j - 1)$ où I_j dénote le nombre de niveaux de la variable *j*.

Un niveau est mis de côté (d'où le « moins un ») pour chaque variable. Ce niveau est en quelque sorte pris en référence et est de toute façon déduit par la non-occurrence de tous les autres niveaux. Un exemple concret est montré dans le tableau A2.

Dans le but de décomposer une matrice cible en valeurs et vecteurs propres comme dans le processus de l'AFC, la table disjonctive complète est récupérée et permet de construire un objet global appelé table de Burt (notée *B*). Il s'agit alors d'une sorte de table de contingence multivariée telle que :

$$B = Z^T Z = \begin{pmatrix} D_1 & Z_1^T Z_2 & \cdots & Z_1^T Z_p \\ Z_2^T Z_1 & D_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ Z_p^T Z_1 & \cdots & \cdots & D_p \end{pmatrix}$$

Dans *B*, on retrouve les éléments $Z_j^T Z_k$ qui ne sont rien d'autre que les tables de contingence (partielles) entre les variables *j* et *k* successives (*j* et *k* allant de 1 à *p*). Les éléments D_j, quant à eux, se réfèrent à la matrice diagonale des effectifs des $(I_j - 1)$ niveaux de la variable *j* (*j* allant de 1 à *p*).

Dans le même esprit que dans la section B.2, c'est la décomposition spectrale de la table de Burt (B) qui permet l'Analyse des Correspondances Multiples. Pour davantage de détails techniques et mathématiques, sans doute non-nécessaires dans le cadre de ce travail, voir Greenacre et Blasius, 2006 ou Le Roux et Rouanet, 2004.

Tableau A2 : Exemple simplifié d'un passage de variables catégorielles initiales à un tableau disjonctif complet

Type de route (2 niveaux)	Agglomération (2 niveaux)	Type de collision (8 niveaux)	Luminosité (4 niveaux)
Communale	Agglomération	Frontale	Plein jour
Communale	Hors agglomération	Par le côté	Nuit sans éclairage
Autoroute	Hors agglomération	En chaîne	Plein jour
Autoroute	Agglomération	Contre un obstacle	Crépuscule
Communale	Agglomération	Par le côté	Plein jour

Route	Agglo.	C.1	C.2	C.3	C.4	C.5	C.6	C.7	L.1	L.2	L.3
1	1	0	0	1	0	0	0	0	1	0	0
1	0	0	0	0	0	0	0	1	0	0	0
0	0	0	1	0	0	0	0	0	1	0	0
0	1	0	0	0	0	1	0	0	0	1	0
1	1	0	0	0	0	0	0	1	1	0	0

À nouveau, une fois la décomposition spectrale effectuée, les valeurs et vecteurs propres obtenus, la première réflexion à avoir consistera à choisir de façon adéquate le nombre d'axes principaux, ou de dimensions principales, à conserver et interpréter. Ce choix se basera toujours sur le pourcentage de variance expliquée et sera toujours un compromis entre possible perte d'information et possible prise en compte d'un bruit de fond.

C.2. Interprétations graphiques et numériques

Autant graphiquement que numériquement, les interprétations permises par une ACM sont très semblables par rapport à celles d'une AFC. Ainsi, la section B.3 reste en grande partie d'actualité ici.

Pour l'ACM, l'objectif final est d'obtenir une représentation graphique dans laquelle :

- Toutes les modalités de toutes les variables initiales et tous les individus peuvent être représentés sur un même plan ;
- La proximité (éloignement) entre une modalité et un individu peut prendre un sens intrinsèque ;

- Le centre du plan ou d'un axe, l'origine, peut aussi avoir un sens intrinsèque.

Concrètement, la projection d'une catégorie est, à une constante près, la moyenne arithmétique des projections des individus qui la composent. La projection d'un individu est quant à elle, à une constante près, la moyenne arithmétique des projections des modalités auxquelles il appartient.

Graphiquement, dans le plan des deux premiers axes principaux, cela se traduit par les possibilités d'interprétation suivantes :

- Dans le cas de modalités proches, d'une même variable ou de variables catégorielles différentes, cela signifie que ces modalités ont beaucoup d'individus en commun. Il en ressort une **association forte positive**, ou une **surreprésentation**, comme déjà rencontrée avec l'AFC. Comme auparavant, il s'agira d'un motif de rejet de l'hypothèse préalable d'indépendance entre les variables ;
- Dans le cas de modalités éloignées entre elles, et éloignées de l'origine, qui s'opposent dans le plan, on en déduira qu'elles ont peu d'individus en commun (moins que dans la moyenne). On parle alors d'**association forte négative** ou de **sous-représentation** ;
- Le centre du graphique est assimilable à un centre de gravité du nuage de points.

De façon numérique, les mêmes *outputs* que pour l'AFC sont disponibles : **valeurs propres**, **coordonnées**, **cosinus carrés** et **contributions** (autant pour les individus que pour les modalités dans les trois derniers cas). Les interprétations restent semblables (cf. section B.3).

Les résultats et graphes obtenus grâce aux ACM sur la base du jeu de données multivarié présenté à la section 2.2 sont mis à disposition en détail dans la section 3.3.

C.3. Discussion ACM versus *AFC*

Au-delà du nombre de variables catégorielles considérées, qui est la différence majeure entre l'AFC et l'ACM, une autre considération doit être prise en compte au moment de choisir l'analyse adéquate. Comme on l'a vu dans les sections 2.1 et 2.2, lorsque les données exploitées ont été présentées, pour pouvoir affirmer que deux modalités ont ou pas beaucoup d'individus en commun, il faut pouvoir compter sur une base de données dans laquelle chaque individu est référencé dans chacune des variables catégorielles. Autrement dit, il ne suffit pas d'avoir un premier tableau de contingence entre les individus et une quelconque variable *A*, puis une table de contingence entre ces mêmes individus et une autre variable *B*.

Avoir les chiffres de ces deux tables de contingence ne fournit pas d'information sur le croisement entre ces variables *A* et *B*. Par exemple, une première table pourrait nous dire qu'il y a eu 392 accidents de la route hors agglomération dans la commune de Braine-l'Alleud entre 2009 et 2018. Une seconde table pourrait nous dire qu'il y a eu 61 accidents de la route dans des conditions de luminosité d'aube ou de crépuscule dans la même commune et ce sur la même période de temps. Combien d'accidents, parmi ces 61 accidents, ont eu lieu hors agglomération ? Combien d'accidents, parmi lesdits 392, ont eu lieu dans des conditions de luminosité d'aube/crépuscule ? Impossible à dire en l'état.

Ainsi, il est important de comprendre, à titre d'indication méthodologique, que si l'on ne dispose que de tableaux de contingence séparés, seules des AFC, elles-mêmes séparées, seront pertinentes. Si l'on dispose par contre d'un profilage plus précis des individus (comme dans la seconde base de données, section 2.2), un par un, parmi les modalités de plusieurs variables catégorielles simultanément, une ACM sera pertinente et hautement valorisable. En effet, dans ce cas précis, la reconnaissance de la simultanéité de plusieurs facteurs favorisera l'interprétation d'interactions et d'associations fortes positives ou négatives entre modalités de variables différentes.

ANNEXE 2 : EXEMPLE DE SORTIE NUMÉRIQUE DÉTAILLÉE POUR UNE AFC (COORDONNÉES, COSINUS CARRÉS ET CONTRIBUTIONS DES MODALITÉS ET DES COMMUNES – VARIABLE « ÉTAT DE LA CHAUSSÉE »)

Modalités	Coordonnées			Cosinus carrés			Contributions		
	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3
CHAUSSÉE PROPRE	-0,049	-0,027	0,0183	0,695	0,207	0,099	12,947	9,361	5,622
CHAUSSÉE HUMIDE	0,0566	0,0952	-0,076	0,177	0,501	0,322	5,684	39,155	31,743
VERGLAS/NEIGE	0,5202	-0,213	-0,028	0,855	0,143	0,003	68,291	27,761	0,621
CHAUSSÉE SALE	0,3814	0,3293	0,4745	0,304	0,226	0,47	13,078	23,723	62,013

Communes	Coordonnées			Cosinus carrés			Contributions		
	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3
Aiseau-Presles	-0,010	-0,013	-0,126	0,006	0,010	0,985	0,001	0,006	0,727
Amay	0,001	-0,064	0,008	0,000	0,987	0,013	0,000	0,246	0,004
Amblève	0,308	-0,043	-0,110	0,871	0,017	0,112	1,150	0,056	0,454
Andenne	-0,048	0,026	0,020	0,677	0,203	0,120	0,117	0,085	0,063
Anderlues	-0,141	-0,050	0,007	0,885	0,113	0,002	0,353	0,110	0,002
Anhée	0,004	0,086	0,089	0,001	0,482	0,517	0,000	0,222	0,300
Ans	-0,022	-0,087	-0,038	0,053	0,795	0,152	0,036	1,290	0,310
Anthisnes	0,191	0,129	-0,204	0,386	0,176	0,438	0,310	0,343	1,075
Antoing	0,129	0,167	0,345	0,101	0,171	0,728	0,310	1,274	6,840
Arlon	0,121	0,006	-0,060	0,802	0,002	0,197	0,922	0,005	0,693
Assesse	0,199	0,000	-0,008	0,998	0,000	0,002	0,925	0,000	0,005
Ath	-0,042	-0,032	0,094	0,152	0,089	0,759	0,119	0,168	1,813
Attert	0,230	0,206	-0,083	0,518	0,414	0,068	0,318	0,619	0,127
Aubange	0,072	-0,016	-0,035	0,781	0,037	0,182	0,143	0,017	0,102
Aubel	-0,020	-0,224	0,126	0,006	0,756	0,239	0,003	0,902	0,359
Awans	-0,002	-0,209	0,036	0,000	0,972	0,028	0,000	2,097	0,077
Aywaille	0,067	0,119	-0,064	0,197	0,624	0,180	0,152	1,171	0,425
Baelen	0,445	-0,063	0,194	0,826	0,017	0,157	2,412	0,118	1,407
Bassenge	-0,182	-0,088	0,000	0,810	0,191	0,000	0,512	0,293	0,000
Bastogne	0,235	-0,095	-0,078	0,786	0,128	0,086	2,181	0,866	0,732
Beaumont	0,097	-0,101	-0,059	0,408	0,443	0,149	0,261	0,690	0,293
Beauraing	0,028	-0,033	0,026	0,314	0,424	0,262	0,022	0,073	0,057
Beauvechain	0,254	-0,052	0,031	0,946	0,040	0,014	0,439	0,045	0,020
Beloeil	-0,099	-0,058	0,042	0,652	0,228	0,120	0,339	0,288	0,191
Berloz	0,034	0,048	0,181	0,032	0,062	0,905	0,006	0,026	0,477
Bernissart	-0,048	0,040	-0,060	0,307	0,211	0,483	0,073	0,122	0,351
Bertogne	0,374	-0,252	-0,006	0,687	0,313	0,000	0,818	0,908	0,001
Bertrix	0,122	-0,046	-0,002	0,877	0,123	0,000	0,316	0,108	0,000
Beyne-Heusay	-0,115	-0,003	0,091	0,611	0,001	0,389	0,195	0,000	0,380
Bièvre	0,120	0,088	-0,045	0,594	0,323	0,084	0,122	0,162	0,053
Binche	-0,052	-0,018	0,007	0,887	0,099	0,014	0,151	0,041	0,007

Communes	Coordonnées			Cosinus carrés			Contributions		
	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3
Blégny	-0,013	-0,051	0,058	0,026	0,424	0,550	0,004	0,170	0,278
Bouillon	0,178	0,218	0,052	0,386	0,582	0,033	0,671	2,465	0,174
Boussu	-0,191	-0,105	0,055	0,723	0,218	0,059	0,784	0,576	0,196
Braine-L'Alleud	0,003	0,036	0,059	0,002	0,273	0,725	0,001	0,250	0,837
Braine-Le-Château	-0,106	0,094	0,118	0,330	0,260	0,410	0,257	0,492	0,975
Braine-Le-Comte	0,137	-0,091	0,052	0,631	0,278	0,091	0,742	0,796	0,328
Braives	0,249	0,058	0,116	0,786	0,042	0,172	0,795	0,103	0,533
Brugelette	0,100	-0,122	0,087	0,307	0,458	0,235	0,088	0,321	0,207
Brunehaut	-0,175	0,058	0,098	0,704	0,077	0,219	0,276	0,073	0,264
Bullange	0,417	-0,184	0,111	0,790	0,154	0,056	1,378	0,653	0,301
Burdinne	0,014	0,132	0,033	0,010	0,931	0,059	0,001	0,141	0,011
Burg-Reuland	0,168	0,039	0,106	0,690	0,038	0,273	0,353	0,047	0,427
Bütgenbach	0,321	-0,310	-0,094	0,496	0,462	0,043	1,017	2,306	0,267
Celles	-0,016	0,102	0,168	0,007	0,267	0,726	0,003	0,303	1,035
Cerfontaine	0,127	-0,077	-0,101	0,498	0,184	0,318	0,292	0,262	0,570
Chapelle-Lez-Herlaimont	-0,150	-0,034	-0,049	0,865	0,044	0,091	0,566	0,070	0,183
Charleroi	-0,071	0,007	-0,006	0,984	0,010	0,006	2,543	0,062	0,048
Chastre	0,010	-0,100	-0,081	0,006	0,598	0,397	0,000	0,118	0,099
Châtelet	-0,069	0,070	-0,036	0,435	0,447	0,118	0,367	0,915	0,305
Chaudfontaine	-0,081	-0,088	-0,047	0,400	0,468	0,132	0,178	0,508	0,181
Chaumont-Gistoux	0,132	-0,057	0,030	0,809	0,150	0,042	0,309	0,139	0,049
Chièvres	0,063	0,015	0,197	0,092	0,005	0,902	0,046	0,007	1,379
Chimay	0,054	-0,036	-0,010	0,678	0,300	0,022	0,114	0,123	0,012
Chiny	0,022	0,039	0,017	0,204	0,668	0,127	0,004	0,034	0,008
Ciney	0,097	-0,028	0,029	0,855	0,069	0,077	0,293	0,057	0,081
Clavier	0,263	-0,151	0,037	0,741	0,244	0,015	0,592	0,474	0,036
Colfontaine	-0,170	-0,169	0,045	0,486	0,481	0,033	0,369	0,889	0,078
Comblain-Au-Pont	0,062	0,124	-0,026	0,193	0,774	0,033	0,045	0,444	0,024
Comines-Warneton	-0,014	-0,162	0,057	0,007	0,882	0,111	0,006	1,853	0,294
Courcelles	-0,010	-0,062	-0,028	0,022	0,816	0,162	0,007	0,606	0,151
Court-Saint-Étienne	0,122	0,140	0,074	0,373	0,489	0,138	0,326	1,040	0,370
Couvin	0,063	0,170	-0,139	0,077	0,552	0,372	0,157	2,732	2,316
Crisnée	0,075	-0,129	0,002	0,252	0,748	0,000	0,051	0,369	0,000
Dalhem	0,078	0,021	-0,019	0,882	0,065	0,052	0,094	0,017	0,017
Daverdisse	0,215	-0,394	0,007	0,230	0,770	0,000	0,100	0,812	0,000
Dinant	-0,089	0,018	0,038	0,815	0,035	0,150	0,230	0,024	0,130
Dison	0,028	-0,004	0,012	0,832	0,021	0,147	0,016	0,001	0,009
Doische	0,148	-0,054	0,229	0,283	0,038	0,679	0,142	0,046	1,040
Donceel	0,080	-0,077	-0,127	0,223	0,211	0,566	0,026	0,060	0,201
Dour	-0,031	-0,008	0,048	0,287	0,019	0,694	0,020	0,003	0,147
Durbuy	0,067	0,156	0,077	0,129	0,702	0,169	0,159	2,103	0,637
Écaussinnes	-0,079	-0,090	0,067	0,329	0,433	0,238	0,094	0,302	0,209

Communes	Coordonnées			Cosinus carrés			Contributions		
	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3
Eghezée	0,040	0,031	0,002	0,615	0,384	0,002	0,061	0,092	0,001
Ellezelles	0,183	-0,082	0,171	0,481	0,096	0,423	0,337	0,163	0,908
Enghien	0,033	-0,084	0,041	0,109	0,720	0,171	0,026	0,421	0,126
Engis	0,173	-0,034	0,044	0,907	0,034	0,059	0,428	0,039	0,085
Erezée	0,250	-0,036	0,028	0,968	0,020	0,012	0,628	0,032	0,025
Erquelinnes	-0,141	-0,053	0,024	0,856	0,119	0,025	0,526	0,178	0,047
Esneux	0,036	0,152	-0,157	0,027	0,472	0,501	0,037	1,584	2,112
Estaimpuis	-0,185	-0,063	0,067	0,800	0,093	0,107	0,721	0,204	0,294
Estinnes	0,066	0,013	0,135	0,191	0,007	0,802	0,072	0,007	0,928
Etalle	-0,036	-0,013	-0,044	0,377	0,051	0,572	0,016	0,005	0,075
Eupen	0,033	-0,103	0,029	0,086	0,846	0,068	0,047	1,114	0,113
Faimes	0,020	-0,044	0,036	0,108	0,542	0,350	0,002	0,027	0,022
Farciennes	0,029	0,053	0,005	0,233	0,759	0,008	0,016	0,123	0,002
Fauvillers	0,107	-0,135	-0,104	0,283	0,450	0,268	0,040	0,154	0,115
Fernelmont	-0,052	-0,065	0,025	0,362	0,555	0,084	0,071	0,266	0,051
Ferrières	0,172	-0,013	-0,032	0,962	0,005	0,033	0,434	0,006	0,045
Fexhe-Le-Haut-Clocher	-0,121	-0,132	-0,001	0,455	0,545	0,000	0,078	0,227	0,000
Flémalle	-0,048	0,036	-0,038	0,452	0,264	0,284	0,117	0,166	0,224
Fléron	-0,079	0,035	0,046	0,655	0,127	0,218	0,159	0,075	0,163
Fleurus	0,002	-0,079	-0,038	0,001	0,811	0,189	0,000	0,758	0,222
Flobecq	-0,163	-0,176	0,045	0,446	0,520	0,034	0,078	0,223	0,018
Floreffe	-0,039	-0,019	0,057	0,293	0,069	0,638	0,023	0,013	0,155
Florennes	0,002	0,009	-0,008	0,037	0,540	0,423	0,000	0,006	0,006
Florenville	0,072	-0,029	-0,031	0,745	0,120	0,135	0,073	0,029	0,041
Fontaine l'Évêque	-0,050	-0,079	-0,011	0,286	0,701	0,014	0,078	0,466	0,012
Fosses-La-Ville	-0,078	-0,018	-0,051	0,671	0,037	0,292	0,130	0,017	0,174
Frameries	-0,082	-0,199	0,022	0,145	0,844	0,011	0,102	1,438	0,023
Frasnes-Lez-Anvaing	0,026	0,088	0,041	0,065	0,768	0,167	0,014	0,400	0,109
Froidchapelle	0,209	0,086	0,039	0,830	0,141	0,029	0,577	0,239	0,061
Gedinne	0,091	-0,060	0,011	0,692	0,299	0,009	0,094	0,099	0,004
Geer	-0,055	0,103	0,130	0,099	0,349	0,552	0,017	0,147	0,292
Gembloux	0,074	0,087	0,080	0,281	0,389	0,331	0,275	0,928	0,994
Genappe	0,096	0,077	0,026	0,581	0,376	0,043	0,341	0,539	0,078
Gerpinnes	0,023	-0,010	-0,047	0,188	0,035	0,778	0,014	0,006	0,182
Gesves	0,078	0,162	-0,062	0,167	0,726	0,107	0,074	0,781	0,144
Gouvy	0,366	-0,013	0,088	0,944	0,001	0,055	1,800	0,005	0,319
Grâce-Hollogne	0,041	0,029	-0,014	0,625	0,307	0,069	0,110	0,131	0,037
Grez-Doiceau	0,024	0,018	0,292	0,007	0,004	0,990	0,011	0,015	4,992
Habay	0,169	0,011	-0,065	0,867	0,004	0,129	0,469	0,005	0,213
Hamoir	0,080	0,279	0,199	0,051	0,628	0,321	0,060	1,797	1,157
Hamois	-0,014	0,020	-0,144	0,009	0,019	0,973	0,003	0,016	1,056
Ham-Sur-Heure-Nalinnes	0,189	0,045	0,209	0,440	0,025	0,535	0,538	0,075	1,999

Communes	Coordonnées			Cosinus carrés			Contributions		
	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3
Hannut	-0,069	-0,123	-0,008	0,240	0,757	0,003	0,184	1,413	0,007
Hastière	0,079	-0,037	0,011	0,807	0,179	0,014	0,067	0,036	0,004
Havelange	0,156	-0,002	0,072	0,827	0,000	0,173	0,226	0,000	0,145
Hélécine	-0,184	0,045	0,087	0,778	0,047	0,175	0,173	0,025	0,119
Hensies	-0,057	-0,024	0,027	0,710	0,129	0,161	0,046	0,020	0,032
Herbeumont	0,267	0,279	-0,149	0,415	0,455	0,130	0,217	0,578	0,207
Héron	0,077	0,089	-0,164	0,145	0,195	0,660	0,080	0,260	1,111
Herstal	-0,136	-0,022	-0,016	0,962	0,025	0,013	1,725	0,108	0,073
Herve	0,241	0,011	-0,041	0,970	0,002	0,028	2,807	0,015	0,247
Honnelles	0,151	-0,147	0,046	0,488	0,467	0,045	0,121	0,282	0,034
Hotton	0,000	-0,110	0,079	0,000	0,659	0,341	0,000	0,508	0,331
Houffalize	0,251	0,043	0,022	0,964	0,029	0,007	1,266	0,092	0,029
Houyet	0,001	-0,121	0,105	0,000	0,569	0,431	0,000	0,544	0,519
Huy	-0,057	0,083	-0,025	0,300	0,641	0,059	0,181	0,939	0,109
Incourt	0,058	-0,026	-0,149	0,128	0,027	0,845	0,037	0,019	0,751
Ittre	0,097	0,006	-0,020	0,958	0,003	0,039	0,146	0,001	0,018
Jalhay	0,497	-0,121	-0,090	0,916	0,054	0,030	6,354	0,911	0,640
Jemeppe-Sur-Sambre	-0,045	0,040	-0,064	0,260	0,208	0,532	0,087	0,169	0,545
Jodoigne	-0,017	-0,034	0,052	0,069	0,278	0,653	0,008	0,075	0,222
Juprelle	0,145	-0,058	0,008	0,858	0,139	0,003	0,412	0,162	0,004
Jurbise	-0,017	-0,068	-0,040	0,043	0,709	0,247	0,005	0,196	0,086
La Bruyère	0,053	0,135	0,071	0,107	0,700	0,193	0,061	0,973	0,338
La Calamine	-0,144	-0,061	-0,034	0,810	0,145	0,045	0,260	0,113	0,045
La Hulpe	-0,098	-0,015	-0,006	0,974	0,023	0,004	0,107	0,006	0,001
La Louvière	-0,071	0,007	-0,024	0,890	0,010	0,101	0,805	0,021	0,279
La Roche-en-Ardenne	0,260	-0,117	0,082	0,769	0,155	0,076	1,096	0,538	0,334
Lasne	0,043	-0,069	-0,037	0,232	0,599	0,169	0,041	0,259	0,092
Le Roeulx	-0,044	0,115	0,108	0,072	0,493	0,435	0,050	0,838	0,931
Léglise	0,453	-0,029	-0,181	0,859	0,004	0,138	2,418	0,024	1,189
Lens	-0,084	-0,053	-0,067	0,492	0,196	0,312	0,078	0,076	0,152
Les Bons Villers	0,175	0,045	-0,045	0,884	0,057	0,059	0,528	0,083	0,107
Lessines	0,010	0,011	0,084	0,014	0,015	0,971	0,003	0,008	0,646
Leuze-en-Hainaut	0,001	-0,078	-0,006	0,000	0,995	0,005	0,000	0,461	0,003
Libin	0,312	0,003	-0,094	0,917	0,000	0,083	1,775	0,000	0,492
Libramont-Chevigny	0,248	-0,015	-0,051	0,956	0,003	0,041	1,941	0,017	0,255
Liège	-0,111	0,006	-0,068	0,726	0,002	0,272	4,548	0,031	5,216
Lierneux	0,060	-0,055	0,050	0,398	0,327	0,275	0,034	0,069	0,073
Limbourg	0,172	0,097	0,078	0,656	0,209	0,135	0,361	0,280	0,227
Lincent	0,088	-0,034	-0,012	0,859	0,125	0,016	0,046	0,016	0,003
Lobbes	-0,071	-0,080	0,003	0,443	0,556	0,001	0,075	0,230	0,000
Lontzen	0,124	-0,052	0,003	0,853	0,147	0,000	0,241	0,101	0,000
Malmédy	0,093	-0,053	-0,055	0,595	0,192	0,213	0,239	0,188	0,263

Communes	Coordonnées			Cosinus carrés			Contributions		
	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3
Manage	-0,023	0,018	-0,045	0,184	0,107	0,710	0,025	0,035	0,293
Manhay	0,425	-0,209	-0,098	0,773	0,187	0,041	2,929	1,722	0,474
Marche-en-Famenne	0,053	0,002	0,001	0,998	0,002	0,000	0,149	0,001	0,000
Marchin	0,133	0,114	-0,140	0,353	0,257	0,390	0,119	0,210	0,402
Martelange	0,338	-0,203	-0,168	0,622	0,224	0,154	0,558	0,489	0,423
Meix-Devant-Virton	0,216	0,096	-0,072	0,765	0,149	0,085	0,180	0,086	0,062
Merbes-Le-Château	0,065	0,011	0,172	0,125	0,003	0,872	0,032	0,002	0,678
Messancy	0,183	-0,026	-0,065	0,871	0,018	0,111	0,616	0,031	0,241
Mettet	-0,076	-0,058	0,046	0,506	0,303	0,191	0,174	0,253	0,200
Modave	-0,051	0,052	0,029	0,425	0,440	0,135	0,039	0,098	0,038
Momignies	0,029	-0,090	0,038	0,080	0,784	0,136	0,011	0,256	0,056
Mons	-0,098	0,058	-0,049	0,623	0,218	0,159	1,968	1,678	1,536
Mont-de-l'Enclus	0,048	-0,190	0,069	0,053	0,836	0,111	0,018	0,674	0,113
Montigny-Le-Tilleul	0,010	0,045	-0,036	0,028	0,601	0,371	0,002	0,096	0,074
Mont-Saint-Guibert	0,025	0,015	-0,064	0,121	0,046	0,834	0,011	0,010	0,232
Morlanwelz	-0,152	-0,019	0,073	0,805	0,012	0,183	0,621	0,023	0,432
Mouscron	-0,151	-0,033	0,029	0,921	0,045	0,035	2,803	0,333	0,322
Musson	0,038	0,083	0,031	0,157	0,738	0,105	0,011	0,121	0,022
Namur	-0,076	0,005	0,032	0,843	0,004	0,154	1,293	0,015	0,722
Nandrin	0,017	-0,037	-0,064	0,052	0,239	0,709	0,004	0,050	0,186
Nassogne	0,200	0,098	0,054	0,762	0,183	0,056	0,503	0,293	0,113
Neufchâteau	0,097	0,047	-0,168	0,238	0,056	0,706	0,184	0,105	1,671
Neupré	0,091	-0,012	-0,096	0,471	0,008	0,521	0,187	0,007	0,634
Nivelles	0,029	0,138	-0,035	0,040	0,902	0,058	0,058	3,185	0,259
Ohey	0,241	0,169	0,145	0,540	0,265	0,196	0,504	0,602	0,560
Olne	0,087	-0,066	0,064	0,473	0,272	0,256	0,032	0,045	0,054
Onhaye	-0,085	0,121	0,024	0,326	0,649	0,026	0,064	0,309	0,016
Oreye	0,228	-0,069	-0,090	0,801	0,073	0,126	0,407	0,091	0,196
Orp-Jauche	0,154	0,197	0,162	0,267	0,438	0,295	0,250	0,998	0,845
Ottignies-Louvain-La-Neuve	0,087	0,119	-0,027	0,338	0,629	0,033	0,353	1,598	0,104
Ouffet	0,160	0,168	0,142	0,346	0,381	0,273	0,194	0,518	0,467
Oupeye	-0,078	-0,059	0,052	0,494	0,282	0,224	0,221	0,307	0,307
Paliseul	0,187	0,056	-0,038	0,884	0,079	0,037	0,411	0,089	0,053
Pecq	-0,118	0,009	0,069	0,740	0,005	0,255	0,226	0,003	0,239
Pepinster	0,020	-0,015	-0,039	0,188	0,105	0,706	0,007	0,010	0,081
Péruwelz	-0,059	-0,028	0,004	0,815	0,181	0,004	0,151	0,081	0,002
Perwez	-0,099	-0,032	0,063	0,662	0,068	0,270	0,249	0,062	0,311
Philippeville	0,078	0,075	-0,106	0,264	0,246	0,489	0,228	0,517	1,294
Plombières	0,056	0,003	0,070	0,386	0,001	0,614	0,046	0,000	0,222
Pont-à-Celles	0,176	0,003	0,010	0,997	0,000	0,003	1,044	0,001	0,010
Profondeville	-0,042	-0,025	0,031	0,531	0,183	0,286	0,046	0,039	0,076

Communes	Coordonnées			Cosinus carrés			Contributions		
	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3
Quaregnon	-0,204	-0,088	0,049	0,804	0,149	0,047	0,904	0,406	0,163
Quévy	-0,008	0,000	0,029	0,062	0,000	0,938	0,001	0,000	0,056
Quiévrain	-0,109	-0,211	0,043	0,204	0,765	0,031	0,121	1,104	0,056
Raeren	0,147	-0,136	0,054	0,504	0,429	0,067	0,426	0,882	0,173
Ramillies	-0,083	0,115	0,043	0,313	0,603	0,084	0,055	0,259	0,045
Rebecq	0,099	0,014	0,123	0,389	0,008	0,604	0,161	0,008	0,764
Remicourt	0,033	-0,216	-0,020	0,023	0,969	0,008	0,010	0,995	0,011
Rendeux	-0,006	0,103	0,054	0,002	0,782	0,216	0,000	0,177	0,061
Rixensart	-0,005	0,014	0,066	0,006	0,045	0,949	0,001	0,012	0,330
Rochefort	0,038	0,097	0,029	0,126	0,802	0,072	0,052	0,803	0,091
Rouvroy	-0,089	0,209	-0,014	0,154	0,842	0,004	0,032	0,432	0,003
Rumes	-0,108	-0,109	0,128	0,291	0,297	0,412	0,067	0,167	0,291
Sainte-Ode	0,200	0,086	0,093	0,713	0,133	0,154	0,276	0,125	0,182
Saint-Georges-Sur-Meuse	0,067	-0,081	0,136	0,152	0,222	0,626	0,061	0,216	0,769
Saint-Ghislain	-0,071	0,119	-0,042	0,242	0,673	0,085	0,193	1,304	0,208
Saint-Hubert	0,359	0,068	-0,072	0,929	0,033	0,038	1,633	0,142	0,202
Saint-Léger	0,380	0,200	0,167	0,681	0,188	0,131	0,622	0,419	0,367
Saint-Nicolas	-0,105	-0,153	0,005	0,321	0,678	0,001	0,231	1,188	0,001
Saint-Vith	0,118	0,002	0,095	0,608	0,000	0,392	0,292	0,000	0,577
Sambreville	-0,106	-0,049	-0,002	0,824	0,176	0,000	0,688	0,357	0,001
Seneffe	0,089	0,147	0,013	0,267	0,727	0,006	0,292	1,930	0,019
Seraing	-0,054	0,073	-0,023	0,334	0,607	0,059	0,447	1,976	0,240
Silly	0,058	-0,011	0,033	0,739	0,025	0,236	0,075	0,006	0,073
Sivry-Rance	0,086	-0,126	0,050	0,288	0,617	0,096	0,086	0,449	0,088
Soignies	-0,028	0,027	0,071	0,120	0,112	0,767	0,042	0,096	0,823
Sombreffe	-0,063	-0,046	-0,081	0,310	0,166	0,525	0,079	0,103	0,408
Somme-Leuze	0,121	0,124	0,225	0,182	0,190	0,629	0,178	0,455	1,894
Soumagne	0,069	-0,008	-0,045	0,691	0,010	0,300	0,172	0,006	0,228
Spa	0,039	0,028	-0,044	0,355	0,189	0,456	0,030	0,039	0,119
Sprimont	0,001	0,012	0,014	0,004	0,398	0,598	0,000	0,011	0,021
Stavelot	0,198	-0,078	-0,024	0,856	0,132	0,012	0,862	0,324	0,037
Stoumont	0,184	0,179	0,153	0,381	0,358	0,262	0,423	0,967	0,891
Tellin	0,051	-0,220	-0,026	0,051	0,936	0,013	0,028	1,257	0,022
Tenneville	0,291	0,002	0,163	0,762	0,000	0,238	0,841	0,000	0,804
Theux	0,007	-0,124	-0,014	0,004	0,985	0,012	0,002	1,246	0,019
Thimister-Clermont	0,168	-0,089	-0,103	0,604	0,168	0,228	0,428	0,290	0,496
Thuin	-0,067	-0,070	0,022	0,458	0,495	0,048	0,135	0,354	0,043
Tinlot	0,163	-0,039	-0,097	0,708	0,040	0,252	0,251	0,035	0,273
Tintigny	0,289	0,104	0,010	0,884	0,115	0,001	0,836	0,265	0,003
Tournai	-0,050	0,046	0,028	0,464	0,394	0,142	0,481	0,993	0,451
Trois-Ponts	0,191	0,037	0,041	0,923	0,035	0,042	0,326	0,030	0,046
Trooz	-0,074	-0,069	0,098	0,274	0,241	0,486	0,059	0,125	0,319

Communes	Coordonnées			Cosinus carrés			Contributions		
	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3
Tubize	-0,104	0,018	0,120	0,421	0,012	0,567	0,335	0,023	1,381
Vaux-Sur-Sûre	0,045	-0,013	0,009	0,887	0,076	0,038	0,021	0,004	0,003
Verlaine	0,276	0,061	-0,001	0,954	0,046	0,000	0,653	0,077	0,000
Verviers	-0,043	0,040	-0,092	0,155	0,136	0,709	0,187	0,400	2,622
Vielsalm	0,309	-0,047	-0,108	0,873	0,020	0,107	2,311	0,127	0,869
Villers-La-Ville	0,081	-0,049	0,084	0,412	0,151	0,437	0,100	0,089	0,324
Villers-Le-Bouillet	-0,019	-0,135	-0,004	0,019	0,980	0,001	0,007	0,912	0,001
Viroinval	0,088	0,006	0,113	0,375	0,002	0,624	0,090	0,001	0,460
Virton	0,056	0,049	-0,067	0,309	0,237	0,454	0,085	0,158	0,381
Visé	-0,058	-0,068	0,032	0,372	0,512	0,116	0,169	0,568	0,161
Vresse-Sur-Semois	0,049	-0,211	0,300	0,018	0,325	0,657	0,019	0,849	2,163
Waimes	0,281	-0,158	-0,036	0,750	0,238	0,012	1,028	0,792	0,052
Walcourt	-0,014	0,002	-0,025	0,238	0,003	0,759	0,008	0,000	0,078
Walhain	-0,038	-0,133	-0,039	0,069	0,859	0,072	0,023	0,704	0,074
Wanze	0,043	-0,083	0,021	0,199	0,752	0,049	0,037	0,344	0,028
Waremme	-0,131	-0,037	-0,026	0,893	0,072	0,036	0,507	0,099	0,062
Wasseiges	-0,062	-0,001	-0,109	0,243	0,000	0,757	0,011	0,000	0,105
Waterloo	-0,122	-0,053	-0,027	0,808	0,152	0,040	1,136	0,521	0,170
Wavre	-0,078	0,004	-0,004	0,995	0,002	0,003	0,570	0,003	0,004
Welkenraedt	0,059	0,031	-0,002	0,781	0,218	0,001	0,066	0,045	0,000
Wellin	-0,001	0,045	0,052	0,000	0,432	0,568	0,000	0,040	0,066
Yvoir	0,054	-0,111	-0,033	0,180	0,753	0,067	0,042	0,424	0,047

ANNEXE 3 : EXEMPLE DE SORTIE NUMÉRIQUE DÉTAILLÉE POUR UNE ACM (COMMUNE DE LIÈGE – ACCIDENTS MORTELS)

Modalités	Coordonnées			Cosinus carrés			Contributions		
	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3
En agglomération	-0,412	0,088	0,028	0,824	0,037	0,004	6,691	0,502	0,059
Hors agglomération	1,998	-0,424	-0,138	0,824	0,037	0,004	32,415	2,433	0,284
Avec un piéton	-0,673	-0,500	-0,714	0,114	0,063	0,129	4,335	3,980	8,959
En chaîne	-0,409	3,731	-0,250	0,001	0,053	0,000	0,030	4,180	0,021
Frontale	-0,356	0,753	1,725	0,013	0,057	0,299	0,549	4,088	23,681
Obstacle hors de la chaussée	0,718	0,459	-0,072	0,287	0,117	0,003	8,745	5,934	0,161
Obstacle sur la chaussée	-0,519	2,581	-0,358	0,005	0,129	0,002	0,243	9,999	0,212
Par l'arrière	0,339	-1,767	1,214	0,010	0,285	0,134	0,456	20,620	10,748
Par le côté	-0,497	-0,294	-0,202	0,059	0,021	0,010	2,276	1,327	0,692
Un seul usager sans obstacle	-0,181	0,198	-0,857	0,002	0,002	0,038	0,077	0,153	3,169
Aube ou crépuscule	0,293	-0,262	4,068	0,003	0,002	0,586	0,139	0,185	49,391
Nuit éclairage public allumé	0,319	0,842	-0,101	0,080	0,559	0,008	2,125	24,690	0,394
Plein jour	-0,287	-0,691	-0,180	0,091	0,527	0,036	2,050	19,766	1,486
Autoroute	2,339	-0,420	-0,236	0,840	0,027	0,009	34,563	1,858	0,645
Route régionale ou communale	-0,359	0,065	0,036	0,840	0,027	0,009	5,306	0,285	0,099

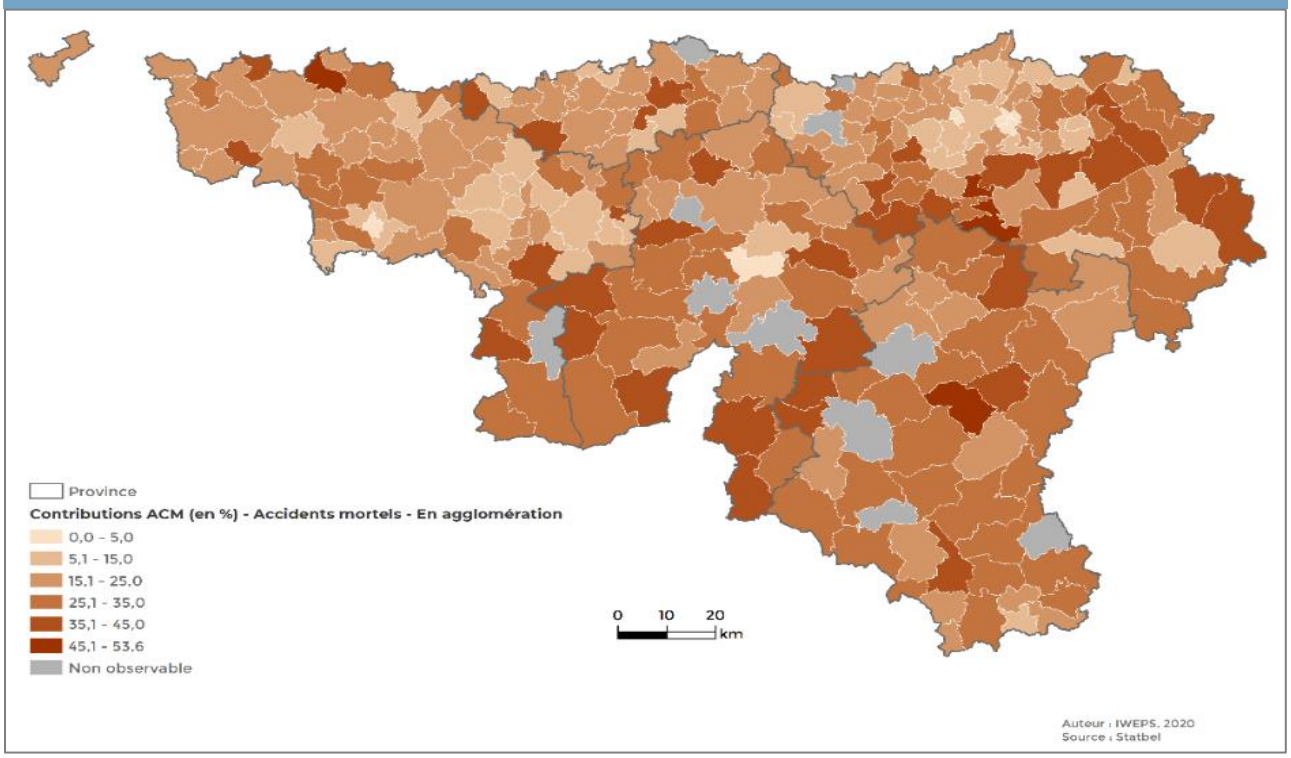
ANNEXE 4 : EXEMPLE DE SORTIE NUMÉRIQUE DÉTAILLÉE POUR UNE ACM (COMMUNE DE HERBEUMONT – ACCIDENTS MORTELS)

Modalités	Coordonnées			Cosinus carrés			Contributions		
	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3	Dim 1	Dim 2	Dim 3
Hors agglomération	0,000	0,000	0,000	NA	NA	NA	0,000	0,000	0,000
Frontale	-0,743	-1,323	0,000	0,221	0,700	0,000	10,714	50,000	0,000
Obstacle hors de la chaussée	0,990	0,000	0,000	0,736	0,000	0,000	28,571	0,000	0,000
Par l'arrière	-0,743	1,323	1,871	0,092	0,292	0,583	5,357	25,000	50,000
Par le côté	-0,743	1,323	-1,871	0,092	0,292	0,583	5,357	25,000	50,000
Nuit sans éclairage public	2,101	0,000	0,000	0,736	0,000	0,000	42,857	0,000	0,000
Plein jour	-0,350	0,000	0,000	0,736	0,000	0,000	7,143	0,000	0,000
Route régionale ou communale	0,000	0,000	0,000	NA	NA	NA	0,000	0,000	0,000

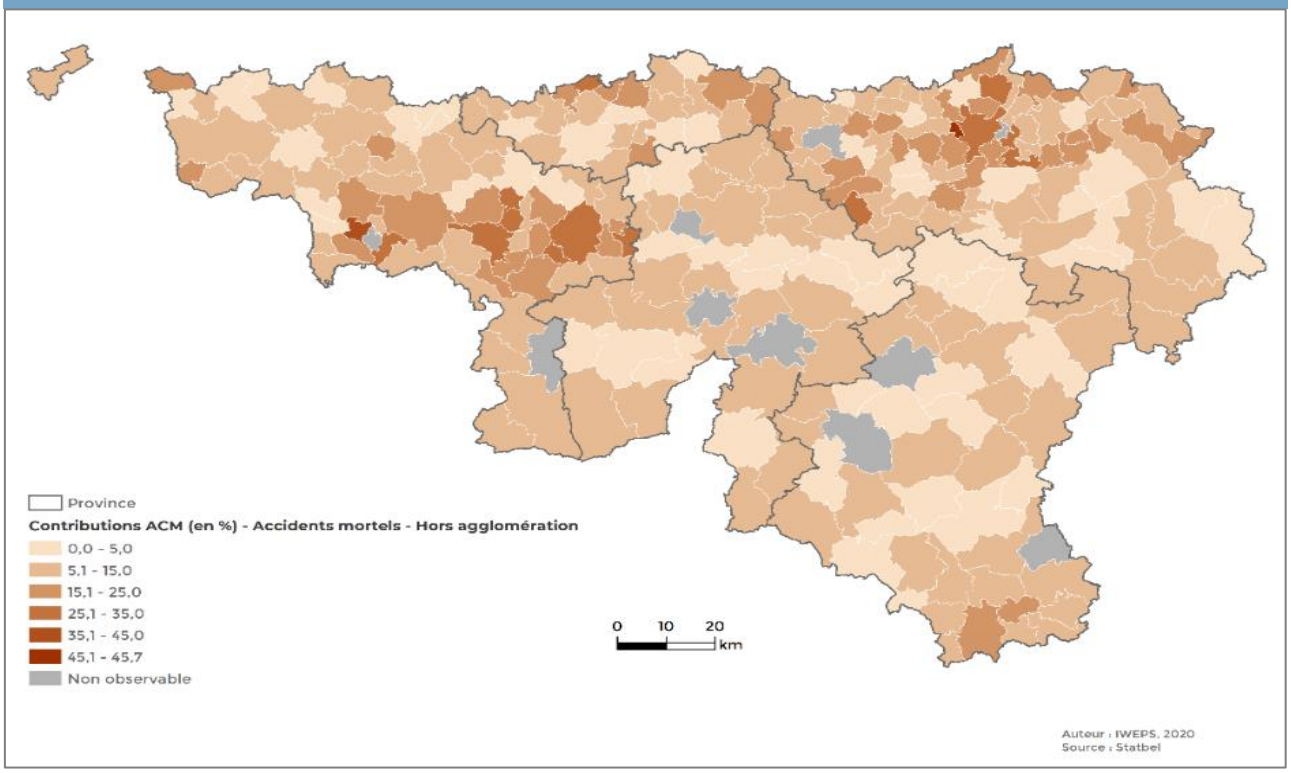
Les modalités jamais rencontrées dans cette commune (ex. : autoroute) n'apparaissent pas. Lorsque la variable catégorielle n'a que deux niveaux renseignés, et si l'un de ces niveaux n'apparaît pas, l'autre niveau ne va jamais jouer un rôle actif (ex. : route régionale ou communale). En effet, puisque l'ensemble des accidents qui ont eu lieu à Herbeumont ont eu lieu sur une route régionale ou communale, ce niveau ne peut pas être discriminant ou informatif. Il s'agit là d'un évènement certain.

ANNEXE 5 : CARTES SUPPLÉMENTAIRES BASÉES SUR LES ACCIDENTS MORTELS

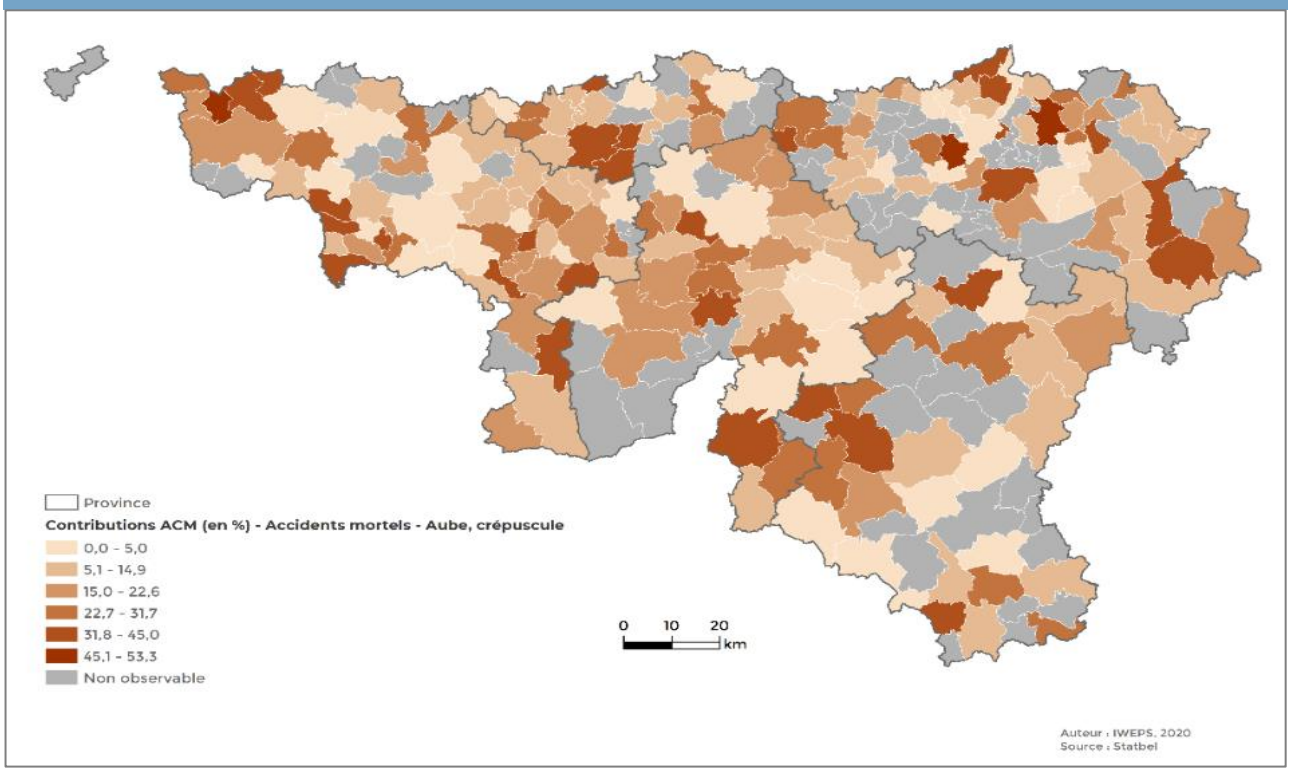
Carte A1 : Représentation des contributions ACM cumulées – Accidents mortels – Modalité « en agglomération »



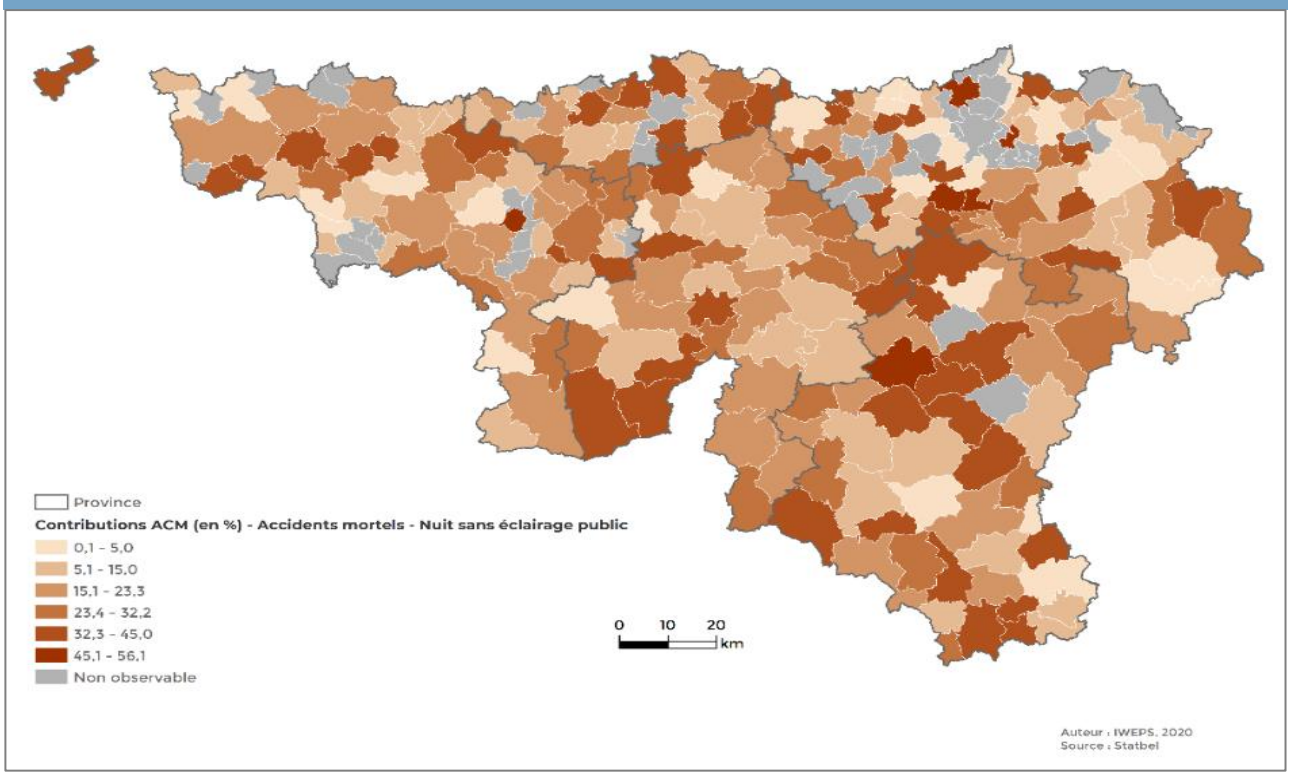
Carte A2 : Représentation des contributions ACM cumulées – Accidents mortels – Modalité « hors agglomération »



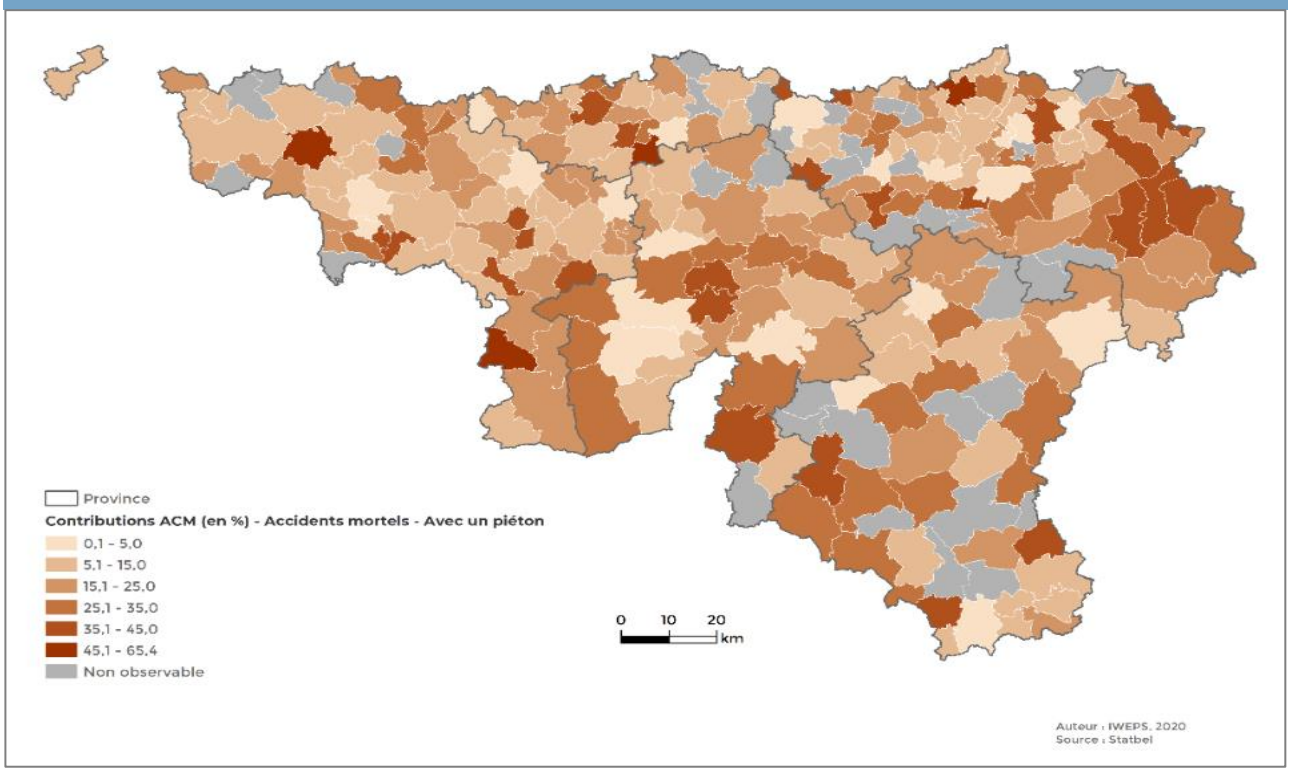
Carte A3 : Représentation des contributions ACM cumulées – Accidents mortels – Modalité « aube, crépuscule »



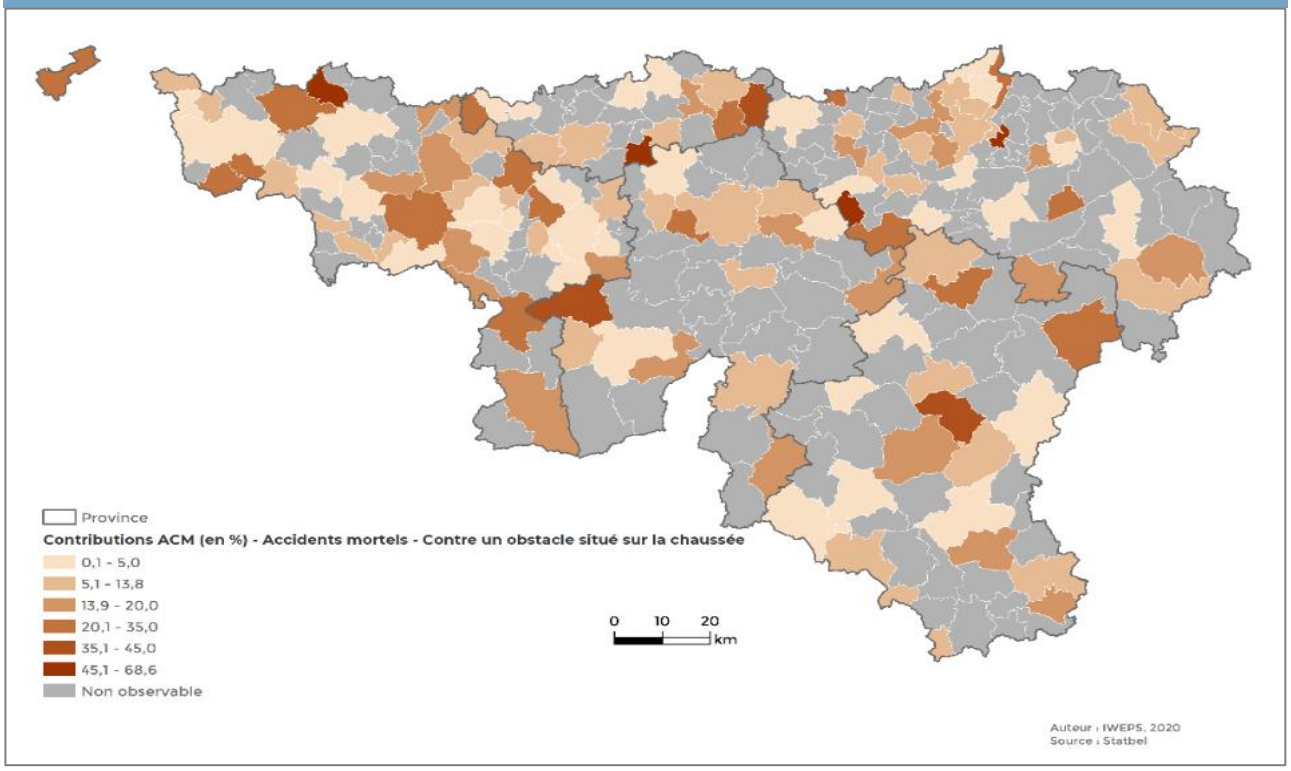
Carte A4 : Représentation des contributions ACM cumulées – Accidents mortels – Modalité « nuit, sans éclairage public »



Carte A5 : Représentation des contributions ACM cumulées – Accidents mortels – Modalité « avec un piéton »



Carte A6 : Représentation des contributions ACM cumulées – Accidents mortels – Modalité « contre un obstacle situé sur la chaussée »



L'Institut wallon de l'évaluation, de la prospective et de la statistique (IWEPS) est un institut scientifique public d'aide à la prise de décision à destination des pouvoirs publics. Autorité statistique de la Région wallonne, il fait partie, à ce titre, de l'Institut Interfédéral de Statistique (IIS) et de l'Institut des Comptes Nationaux (ICN). Par sa mission scientifique transversale, il met à la disposition des décideurs wallons, des partenaires de la Wallonie et des citoyens, des informations diverses qui vont des indicateurs statistiques aux études en sciences économiques, sociales, politiques et de l'environnement. Par sa mission de conseil stratégique, il participe activement à la promotion et la mise en œuvre d'une culture de l'évaluation et de la prospective en Wallonie.
Plus d'infos : https://www.iweps.be

2020

Institut wallon de l'évaluation, de la prospective et de la statistique
Vous pouvez consulter ou télécharger cette publication sur notre site : www.iweps.be