

LA MOBILITÉ QUOTIDIENNE DANS LES GRANDES VILLES DU MONDE : APPLICATION DE LA THÉORIE DES RÉSEAUX BAYÉSIENSⁱ

DAILY MOBILITY IN MAJOR WORLD CITIES: AN APPLICATION OF BAYESIAN NETWORKS' THEORY

[Giovanni Fusco](#)

UMR 6012 ESPACE, Université de Nice-Sophia Antipolis
98 Bd Herriot, 06204 Nice cedex 3, France
Tél : ++33 4 93 37 54 64

DiAP

Politecnico di Milano
Via Bonardi 3, 20133 Milano, Italia
Tél : ++39 02 23 99 41 02

Résumé : À la frontière entre statistiques multivariées et intelligence artificielle, les réseaux bayésiens sont un puissant outil de modélisation de connaissance systémique. Une utilisation particulièrement intéressante est celle de la recherche des liens causaux entre variables dans une base de données. Dans cet article, nous analysons une base de données urbaines sur la mobilité quotidienne couvrant plus de cent grandes villes mondiales. À l'aide d'algorithmes de génération de réseaux bayésiens, nous avons produit deux réseaux reliant les variables et traduisant deux hypothèses différentes sur la hiérarchie causale entre elles. Si l'on introduit l'hypothèse que la mobilité détermine la structure urbaine, on crée alors un réseau qui identifie la motorisation de la population comme étant une variable cruciale. Ce réseau s'avère plus probable que celui résultant de l'hypothèse opposée. En revanche, les deux réseaux concordent sur l'importance de l'offre de voirie rapide urbaine et de la vitesse moyenne du réseau routier dans l'explication du partage modal et des coûts et des externalités des transports urbains.

Mots-clés: réseaux bayésiens, knowledge discovery, mobilité urbaine, villes mondiales, indicateurs urbains.

Abstract : At the frontier between multivariate statistics and artificial intelligence, Bayesian networks are a powerful tool to model systemic knowledge. A particularly interesting use of Bayesian networks is the search of causal links between variables in a database. In this paper, we analyse a database on urban daily mobility covering more than one hundred big world cities. Through the generating algorithms of Bayesian networks, we produced two networks derived from two different hypotheses on the causal hierarchy among the variables. Car ownership is identified as a crucial variable when considering the network resulting from the hypothesis that daily mobility determines urban structure. This network has a higher likelihood than the one resulting from the opposite hypothesis. On the other hand, the two networks converge on the importance of the offer of urban freeway and of the average speed of the road network. These two factors explain the modal split and the costs and externalities of urban transportation.

Key-words : Bayesian networks, knowledge discovery, urban mobility, world cities, urban indicators.

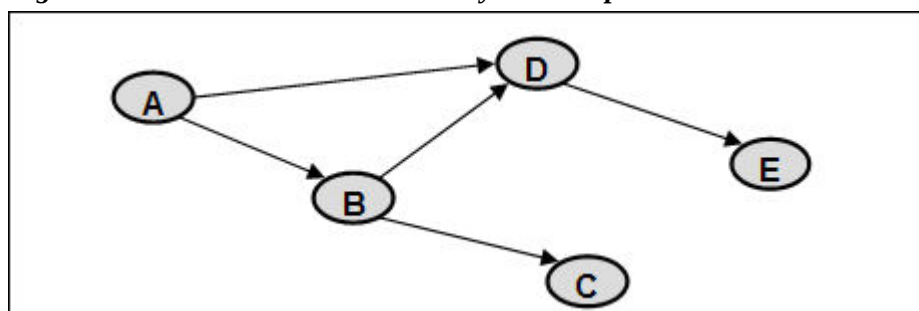
1. Les réseaux bayésiens : une formalisation de la connaissance causale

Les réseaux bayésiens s'inscrivent à la confluence entre statistiques multivariées et intelligence artificielle. Ils constituent aujourd'hui l'une des formalisations les plus complètes et les plus cohérentes pour l'acquisition, la représentation et l'utilisation des connaissances par des ordinateurs (Becker et Naïm 1999). Tout en étant connue depuis longtemps, cette théorie développe ses premières applications opérationnelles lors de la dernière décennie, à la suite de l'augmentation de la puissance de calcul des ordinateurs et surtout grâce aux importantes avancées théoriques constituées par les travaux des Judea Pearl à l'Université de Californie à Los Angeles et de Finn Jensen à l'Université d'Aalborg (Pearl 2000, Jensen 2001).

Les réseaux bayésiens sont utilisés depuis peu dans les études géographiques et, particulièrement, dans les études urbaines. Les premières applications ont concerné le traitement d'images numériques et l'analyse de l'information contenue dans les SIG. Stassopoulou et al. (1996) ont développé un modèle bayésien pour l'évaluation du risque de désertification des aires soumises à un incendie en milieu méditerranéen à partir des connaissances du modélisateur sur les relations fonctionnelles parmi les variables étudiées. En France, Cavarroc et Jeansoulin (1998), dans les travaux du GdR Cassini, ont évalué l'apport des réseaux bayésiens pour la recherche de causalités spatio-temporelles dans les données d'imagerie numérique. Des applications particulièrement intéressantes ont été développées au Département de Géographie de la Pennsylvania State University : l'équipe de Colin Flint (Flint et al. 2000) en géographie électorale, William Pike (2001) dans l'étude des dépassements des normes de qualité des eaux. En Italie, Francesco Scarlatti et Giovanni Rabino, au Département d'Architecture et Aménagement du Polytechnique de Milan, ont proposé l'utilisation des réseaux bayésiens pour l'analyse du paysage (Scarlatti 2002). Ces études pionnières montrent le grand potentiel des réseaux bayésiens dans la mise en évidence de liens systémiques entre variables territoriales. Dans le domaine de l'interaction ville-transports-environnement, l'intérêt d'avoir recours à l'approche des réseaux bayésiens a été mis en évidence par P. Waddell (2001), mais n'a pas jusqu'à présent abouti à une réalisation opérationnelle.

Un réseau bayésien est un modèle graphique (au sens de la théorie des graphes) de connaissance en probabilité. Graphiquement, il est constitué d'un ensemble d'événements (ou variables) reliés par des flèches (figure 1). L'existence d'une flèche orientée allant d'un événement A à un événement B symbolise un lien de cause-à-effet entre eux : il peut s'agir d'une causalité stricte ou d'un type de dépendance causale moins contraignante (A influence de façon générale B). Selon la terminologie des graphes, A est un nœud parent de B et B est un nœud enfant de A. Finalement, dans la figure 1, A est la seule variable indépendante du réseau.

Figure 1 – La structure d'un réseau bayésien simple constitué de 5 variables



Le formalisme graphique du réseau bayésien s'accompagne d'un formalisme probabiliste associé de type quantitatif. Selon ce formalisme, un réseau bayésien est défini par (Becker et Naïm, 1999) :

- un graphe acyclique (c'est-à-dire sans boucles) et orienté G , constitué de nœuds (les variables V_i) et d'arcs orientés (A_{ij});
- un espace probabilisé fini (Ω, Z, p) , où Ω est l'univers des possibles, Z est une tribu des évènements de Ω , et p une application $Z \rightarrow R$ avec domaine image $[0,1]$ et pour laquelle $p(\Omega)=1$;
- un ensemble de variables aléatoires correspondant aux nœuds du graphe et définies sur (Ω, Z, p) , tel que la probabilité globale du réseau soit le produit des probabilités de chaque nœud V_i conditionnellement à l'ensemble de ses nœuds parents $C(V_i)$:

$$p(V_1, V_2, \dots, V_n) = \prod_{i=1}^n p(V_i | C(V_i))$$

Cette dernière expression nous dit que la distribution des probabilités conjointes des variables du réseau a une factorisation très particulière, directement liée à la structure du graphe. Couplant ainsi le formalisme des graphes causaux à celui des équations probabilistes, les réseaux bayésiens sont un instrument apte à représenter un modèle de dépendance causale entre variables stochastiques (Pearl 2000). On appelle « structure » du réseau bayésien, l'ensemble des relations causales qui relient ses variables (les arcs orientées A_{ij}). La théorie des probabilités permet d'associer à chaque nœud enfant du graphe des fonctions de probabilité conditionnelle à la connaissance de ses nœuds parents. L'ensemble des distributions de probabilités marginales (pour les variables indépendantes) et des distributions de probabilités conditionnelles (pour les variables dépendantes) constituent les « paramètres » du réseau. La plus grande partie des algorithmes de calcul de réseaux bayésiens opèrent sur des variables discrètes. L'information probabiliste associée au graphe se réduit dans ce cas à des vecteurs de probabilités marginales pour les nœuds indépendants et à des tables de probabilités conditionnelles pour les nœuds enfants. Une table de probabilités conditionnelles fournit, pour les valeurs du nœud enfant, un vecteur de probabilité différent pour chaque combinaison possible des valeurs des variables parents. Dans la figure 2, nous montrons la table des probabilités des valeurs du nœud enfant B par rapport aux valeurs de son nœud parent A . Cette table quantifie en probabilité la dépendance causale montrée de façon qualitative dans la figure 1. La table doit être lue en ligne : si, par exemple, la variable A est comprise entre 2 et 31, alors la variable B assumera une valeur comprise entre 16 et 961 avec probabilité 0,903, comprise entre 961 et 2116 avec probabilité 0,033 et comprise entre 2116 et 5329 avec probabilité 0,063.

Figure 2 – Table des probabilités conditionnelles de la variable B

Variable_B			
Variable_A	[16.0 961.0)	[961.0 2116.0)	[2116.0 5329.0]
[2.0 31.0)	0.903	0.033	0.063
[31.0 52.0)	0.003	0.783	0.213
[52.0 88.0]	0.091	0.178	0.731

Les derniers développements théoriques (Druzdzel et Simon 1993, Pearl 2000) caractérisent la causalité comme une propriété systémique : elle devient une notion rigoureuse dès que nous mettons chaque élément dans le contexte du système que nous voulons étudier, d'autant plus que la direction d'une relation de cause-à-effet dans un système peut être aisément renversée dans un autre système. Selon Pearl (2000), il n'est pas étonnant que la réflexion sur la notion de causalité ait été remise à l'ordre du jour après l'introduction de la notion de système

(systémique) et en conséquence des problèmes opérationnels constitués par la création de systèmes artificiels (cybernétique). Dans le contexte des analyses de système, la notion de causalité s'est débarrassée du caractère encore métaphysique qu'elle présentait dans l'approche de certains positivistes. Il existe des causes et il existe des effets en référence à une définition précise du système étudié, ce qui introduit une asymétrie fondamentale dans l'observation des phénomènes. La façon la plus pertinente de modéliser les relations causales opérant dans un système est de rendre explicite cette asymétrie et d'avoir recours à un double formalisme couplant un graphe causal avec des relations mathématiques, ce qui est rendu possible par la technique des réseaux bayésiens.

La structure des réseaux bayésiens introduit cependant des contraintes pour la modélisation. En particulier, la structure du réseau ne peut pas contenir des boucles. Cela a d'importantes conséquences sur le type de modèle qui peut être développé pour le système étudié : l'existence de boucles d'interaction entre éléments ne peut être mise en évidence par le biais d'un seul réseau bayésien. Or, en ce qui concerne notre champ d'étude, le système de la mobilité urbaine dans les grandes villes du monde, nous savons que d'importantes interrelations peuvent exister. Ainsi, il n'est pas rare, dans le système de la mobilité urbaine, qu'un élément soit en même temps cause et effet d'un autre. L'analyse conduite par le biais des réseaux bayésiens revient à privilégier à chaque fois un point de vue, et à mettre en évidence causes et effets "en dernière instance". Cette notion est utilisée dans le même sens que François Durand-Dastès, dans ses réflexions sur l'approche systémique en géographie (Durand-Dastès 1985, 1999).

2. La recherche de relations causales dans l'interaction ville-transports-environnement

L'une des utilisations les plus intéressantes des réseaux bayésiens est celle de la *causal knowledge discovery*, c'est-à-dire de la recherche automatique des liens causaux entre les variables d'une base de données. À partir d'un ensemble de variables observées, même de façon lacunaire et incertaine, sur un certain nombre d'individus indépendants, de puissants algorithmes de génération de réseaux bayésiens permettent de déduire en probabilité des liens de cause à effet entre les variables. Les algorithmes les plus complets peuvent apprendre, et ainsi reproduire, tant la structure que les paramètres du réseau.

Selon Pearl et Russel (2000), la possibilité de découvrir des structures causales à partir des données statistiques est l'une des perspectives les plus excitantes ouvertes à l'analyse scientifique au cours de ces dernières années. Il s'agit d'une tâche qui était considérée impossible auparavant sans effectuer des expérimentations contrôlées. Les mêmes auteurs soulignent également que les structures causales réelles ont normalement une stabilité de comportement qui fait qu'elles produisent des ensembles de co-occurrences statistiques qui, tout en étant dépourvues d'information temporelle, sont compatibles avec certaines asymétries dans les relations entre les données et ne le sont pas avec d'autres. Mises ensemble de façon systématique, ces asymétries peuvent être utilisées pour inférer, à partir des données, les structures causales originelles et pour garantir que toute autre structure alternative compatible avec les données soit moins stable que celle (ou celles) inférée(s). Cette tâche peut être accomplie, en probabilité, par les algorithmes d'apprentissage bayésien.

Dans cet article, nous utilisons la technique des réseaux bayésiens pour déceler les relations systémiques existantes entre 21 indicateurs du système de la mobilité urbaine. Ces indicateurs sont issus de la base de données urbaines *Millennium Cities* (UITP-ISTP 2001), mise au point par l'Union Internationale des Transports Publics en collaboration avec l'ISTP de la Murdoch University de Perth, Australie. Ils décrivent de façon synthétique les points principaux de l'interaction ville-transports-environnement : la structure urbaine (densité de population,

centralisation des emplois), l'offre de transport (offre d'infrastructures, offre de services de transports publics, motorisation de la population), la demande de mobilité (nombre et longueur des déplacements, partage modal), les vitesses moyennes des réseaux de transport, les flux produits (de trafic de véhicules et flux de passagers-Km), les coûts de la mobilité (coûts financiers mesurés comme pourcentage du PIB métropolitain et consommation énergétique), les externalités des transports urbains (émissions de polluants atmosphériques locaux et taux de mortalité par accident de la route). Nous avons également inclus deux indicateurs exogènes au système de la mobilité urbaine, mais susceptibles de l'influencer profondément : la taille démographique de la ville et sa richesse appréhendée par le PIB métropolitain par habitant. Les indicateurs et leurs définitions opérationnelles sont indiqués dans le tableau 1.

Les unités couvertes constituent 102 villes mondiales situées dans les cinq continents et appréhendées en tant qu'agglomérations. Il s'agit des 100 villes contenues dans *Millenium Cities*, auxquelles nous avons ajouté les deux conurbations de Nice-Côte d'Azur et de Gênes. Les villes de notre étude sont indiquées dans le tableau 2. Pour la plupart, il s'agit de villes grandes (entre 500 000 et 2 millions d'habitants) ou très grandes (plus de 2 millions d'habitants). Les villes européennes sont au nombre de 43. 11 cités sont issues de l'Afrique et du Moyen Orient, 18 sont des villes asiatiques, 5 des villes d'Océanie, 15 des villes nord-américaines et 10 des villes latino-américaines. La base de données n'est pas complète, les lacunes représentant 6,3% du total des valeurs.

Le logiciel que nous utilisons pour la génération et l'analyse des réseaux bayésiens est *Discoverer* de *Bayesware* (2000), qui permet de rechercher automatiquement les réseaux bayésiens les plus probables même à partir d'une base de données incomplète. *Discoverer* utilise uniquement des variables discrètes. Les variables continues de notre base de données sont par conséquent discrétisées selon une double méthode. Soit elles sont divisées en intervalles de même amplitude (*range discretization*), soit elles sont divisées en intervalles contenant approximativement le même nombre d'individus (*frequency discretization*). Pour nos générations de réseaux, compte tenu du faible nombre d'individus dans la base de données, nous avons toujours utilisé une discrétisation par fréquence en trois classes seulement.

Tableau 1 – Les 21 indicateurs retenus dans l'étude du système de la mobilité urbaine

Libellé	Indicateur	Unité de mesure
Population	Population résidant dans la ville	Habitants
PIB_hab	PIB métropolitain par habitant	Dollars Américains 1995
Densité	Densité nette de population	Habitants par hectare de surface urbanisée
Centralisation	Centralisation des emplois	% emplois dans le centre-ville
Voirie_Rapide	Offre de voirie rapide par habitant	m / habitant
Infrastructures_TP	Offre d'infrastructures réservées aux transports publics par habitant	m / habitant
Service_TP	Offre de service de transports publics par habitant	Véhicules Km / habitant / an
Motorisation	Motorisation de la population	Véhicules motorisés privés / 1000 habitants
Vitesse_Route	Vitesse moyenne du réseau routier	Km / h
Vitesse_TP	Vitesse moyenne du réseau des transports publics	Km / h
Longueur_Depl	Longueur moyenne des déplacements	Km
Mobilite_Tot	Mobilité totale de la population	Déplacements / habitant / jour
Part_Doux	Part de la marche et du vélo dans les déplacements totaux	%
Part_MP	Part des modes motorisés privés (automobile, deux-roues motorisés) dans les déplacements totaux	%
Part_TP	Part des transports publics dans les déplacements totaux	%
Passkm_Motor	Distances totales parcourues en modes motorisés	Passagers Km motorisés / habitant / an
Flux_Trafic	Trafic de véhicules motorisés par habitant	Véhicules Km / habitant / an

Energie	Consommation énergétique des transports des personnes par habitant	MJ / habitant / an
Cout_Tot	Part du PIB métropolitain dépensé pour le transport des personnes	%
Emiss_Poll	Emissions de polluants atmosphériques locaux (CO, NO _x , SO ₂ , COV) par habitant	Kg / habitant / an
Morts	Taux de mortalité par accident de la route	Morts / 100 000 habitants / an

Tableau 2 – Les 102 grandes villes mondiales de l'étude

Europe	Vienne	Varsovie	Asie	Océanie	Phoenix
Lisbonne	Graz	Cracovie	Manille	Brisbane	San Diego
Madrid	Berne	Prague	Bangkok	Melbourne	San Francisco
Barcelone	Genève	Budapest	Bombay	Perth	Washington
Nantes	Zurich	Istanbul	Madras	Sydney	Amérique Latine
Marseille	Bologne	Athènes	New Delhi	Wellington	
Nice	Gênes	Afrique et Moyen Orient	Kuala Lumpur	Amérique du Nord	Buenos Aires
Lyon	Milan		Jakarta		Rio de Janeiro
Paris	Rome	Abidjan	Séoul	Calgary	Curitiba
Lille	Turin	Casablanca	Ho Chi Minh City	Montréal	Sao Paulo
Bruxelles	Londres	Dakar	Pékin	Ottawa	Brasilia
Amsterdam	Manchester	Cape Town	Shanghai	Toronto	Salvador
Berlin	Newcastle	Johannesburg	Canton	Vancouver	Santiago
Francfort	Glasgow	Harare	Osaka	Atlanta	Bagota
Hambourg	Oslo	Tel Aviv	Sapporo	Chicago	Mexico
Düsseldorf	Copenhague	Téhéran	Tokyo	Denver	Caracas
Munich	Stockholm	Riyad	Hong Kong	Houston	
Ruhr	Helsinki	Le Caire	Singapour	Los Angeles	
Stuttgart	Moscou	Tunis	Taipei	New York	

Discoverer utilise l'algorithme *K2* pour la génération des réseaux bayésiens et cela par différentes stratégies de recherche qui mélangent connaissance acquise des données et connaissance fournie par le modélisateur. Mis au point par G. Cooper au début des années 90 (Cooper et Hersovits 1992), l'algorithme *K2* sélectionne le réseau bayésien le plus probable a posteriori (c'est-à-dire sachant les données) parmi un sous-ensemble de réseaux possibles. Ce sous-ensemble est défini par l'utilisateur, qui doit identifier l'ordre dans lequel les variables de la base de données doivent être testées. L'ordre d'évaluation est une hiérarchie causale entre variables proposée par le modélisateur et limite le nombre de réseaux possibles. La position de chaque variable dans cette hiérarchie détermine le nombre de variables qui seront testées comme parents possibles. Soit *X* la dernière variable dans la hiérarchie causale. L'implémentation de l'algorithme commence avec la prise en compte de *X*, et calcule en premier lieu la probabilité (plus précisément la log-vraisemblance) du modèle sans aucun lien causal allant des autres variables vers *X*. L'étape suivante est le calcul de la log-vraisemblance de tous les modèles avec un seul lien causal vers *X*. Si aucune de ces dépendances ne possède une log-vraisemblance supérieure à celle du modèle sans lien vers *X*, ce dernier est retenu comme modèle le plus probable et *Discoverer* passe à l'évaluation de la seconde variable dans la liste. Si au moins l'un de ces modèles possède une log-vraisemblance supérieure à celle du modèle sans lien vers *X*, le lien correspondant de dépendance causale est retenu et la recherche continue en essayant d'ajouter un deuxième lien vers *X*, et cela jusqu'à que la log-vraisemblance cesse d'augmenter. Une fois l'évaluation d'une variable terminée, l'algorithme passe, par la même méthode, à la recherche des liens les plus probables vers la deuxième variable de la liste.

L'algorithme *K2* peut être implémenté selon deux stratégies : *Greedy* et *ArcInversion*. Dans *Greedy*, les variables testées comme parents possibles d'une variable donnée sont seulement celles qui la précèdent dans la hiérarchie causale. Dans *ArcInversion*, au contraire, les variables plus en bas dans l'échelle peuvent elles-aussi être testées comme parents, à condition que cela ne produise pas une boucle (les réseaux bayésiens ne peuvent pas contenir de boucles). *Greedy* est la stratégie la plus simple qui donne au modélisateur le plus grand

contrôle du processus de génération du modèle. Elle est la stratégie utilisée dans la génération des réseaux présentés dans cet article.

Avant de procéder à la génération des réseaux bayésiens, nous avançons deux hypothèses sur la hiérarchie causale entre les variables étudiées. Chaque hypothèse relève d'une approche différente dans l'étude de l'interaction ville-transport-environnement. Selon la première approche, la ville produit sa mobilité, selon la seconde la mobilité produit sa ville.

La figure 3 présente les schémas conceptuels correspondants aux deux approches. Ces schémas peuvent être traduits dans les hiérarchies causales indiquées dans la figure 4, qui serviront donc de base pour produire deux réseaux bayésiens différents. La hiérarchie causale proposée agit comme une contrainte sur la génération du réseau, qui ne peut pas contenir des liens causaux remontant la hiérarchie causale. En revanche, la structure causale du réseau généré peut ne pas contenir le simple enchaînement causal représenté dans le schéma. Des parcours de causalité indirecte peuvent s'avérer statistiquement plus significatifs, ou bien la hiérarchie causale supposée peut être très éloignée des données réelles, produisant un réseau très mal connecté.

Les deux approches possèdent quelques points communs et des différences majeures. Un premier point commun est de considérer les facteurs économiques et démographiques de la ville, variables exogènes au système de la mobilité urbaine, en haut de la hiérarchie causale. Un second point commun est de considérer en bas de la hiérarchie causale d'abord les flux de trafic produits par le système de la mobilité urbaine (les passagers Km-produits par les modes motorisés et les flux de véhicules), ensuite les ressources consommées (ressources financières, énergie) et enfin les externalités produites (émission de polluants, morts des accidents de la route).

Les deux approches diffèrent par contre dans l'enchaînement des phénomènes intermédiaires. Pour la première approche, la ville physique (densité urbaine, centralisation des emplois) est susceptible de déterminer l'offre de transport qui peut se mettre en place (l'offre d'infrastructures, l'offre de services de transports publics et, en conséquence de ces éléments, la motorisation). Ces éléments de l'offre déterminent les vitesses des différents réseaux de transport, qui sont à leur tour à la base de la demande de mobilité (longueur des déplacements, nombre des déplacements et leur partage modal entre modes doux, modes motorisés privés et transports publics). Il s'agit en général d'une approche très fonctionnaliste, basée sur les effets produits par le monde physique sur les comportements humains.

La seconde approche privilégie l'action des acteurs sur le monde physique. Le premier élément dépendant des facteurs sociodémographiques est la motorisation de la population. Les autres éléments de l'offre de transport sont le résultat du jeu d'acteurs, et en particulier des pressions politiques des citoyens motorisés et non motorisés pour avoir des infrastructures et des services répondant à leurs situations. Les éléments de l'offre sont clairement susceptibles de déterminer les vitesses des réseaux de transport, qui sont à leur tour à la base de plusieurs aspects de la demande de mobilité : mobilité totale et partage modal. Vitesses et habitudes modales peuvent à leur tour engendrer la structure physique de la ville (densité, centralisation) et, en conséquence de cette dernière, la longueur des déplacements. La structure urbaine serait alors le produit des comportements humains, permis par les vitesses, possibles grâce aux infrastructures que les pressions politiques des citoyens même ont fait construire.

Figure 3 – Deux approches dans l'étude de l'interaction ville-transports-environnement

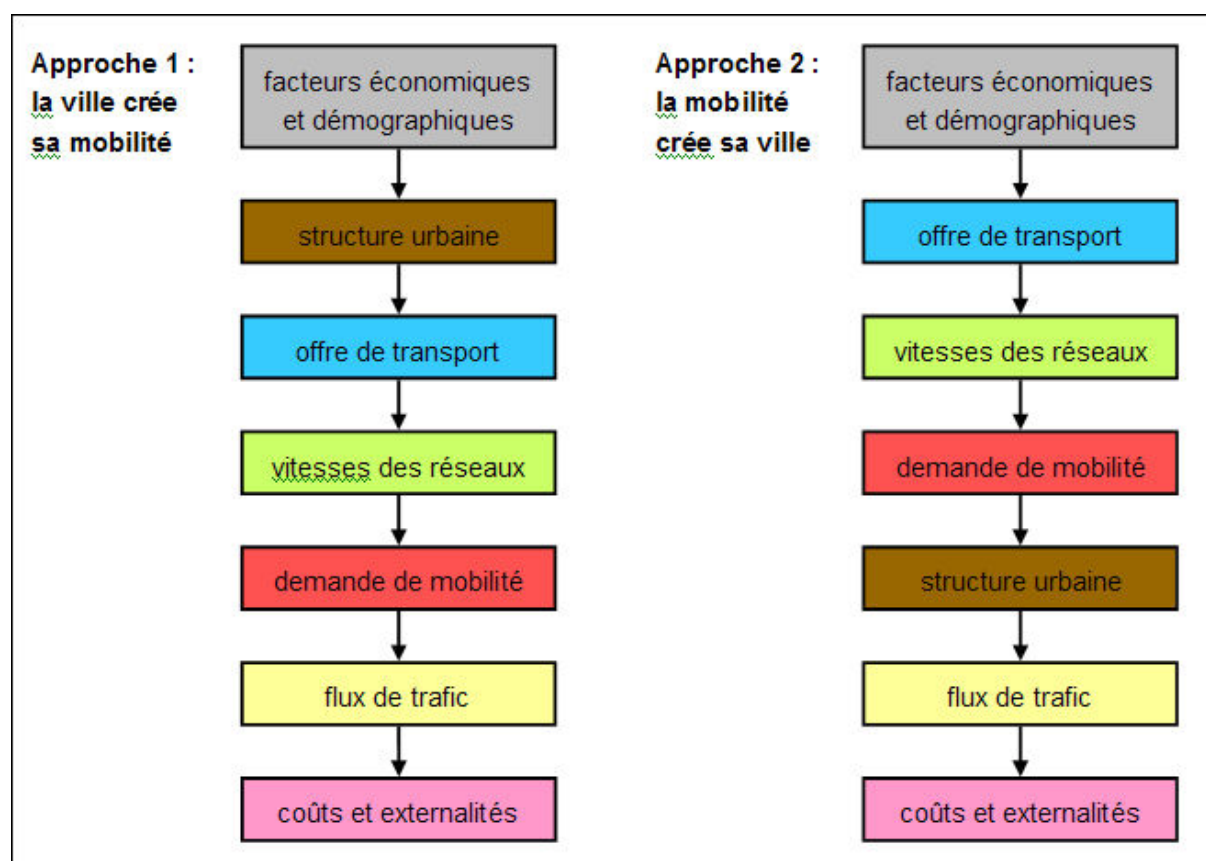
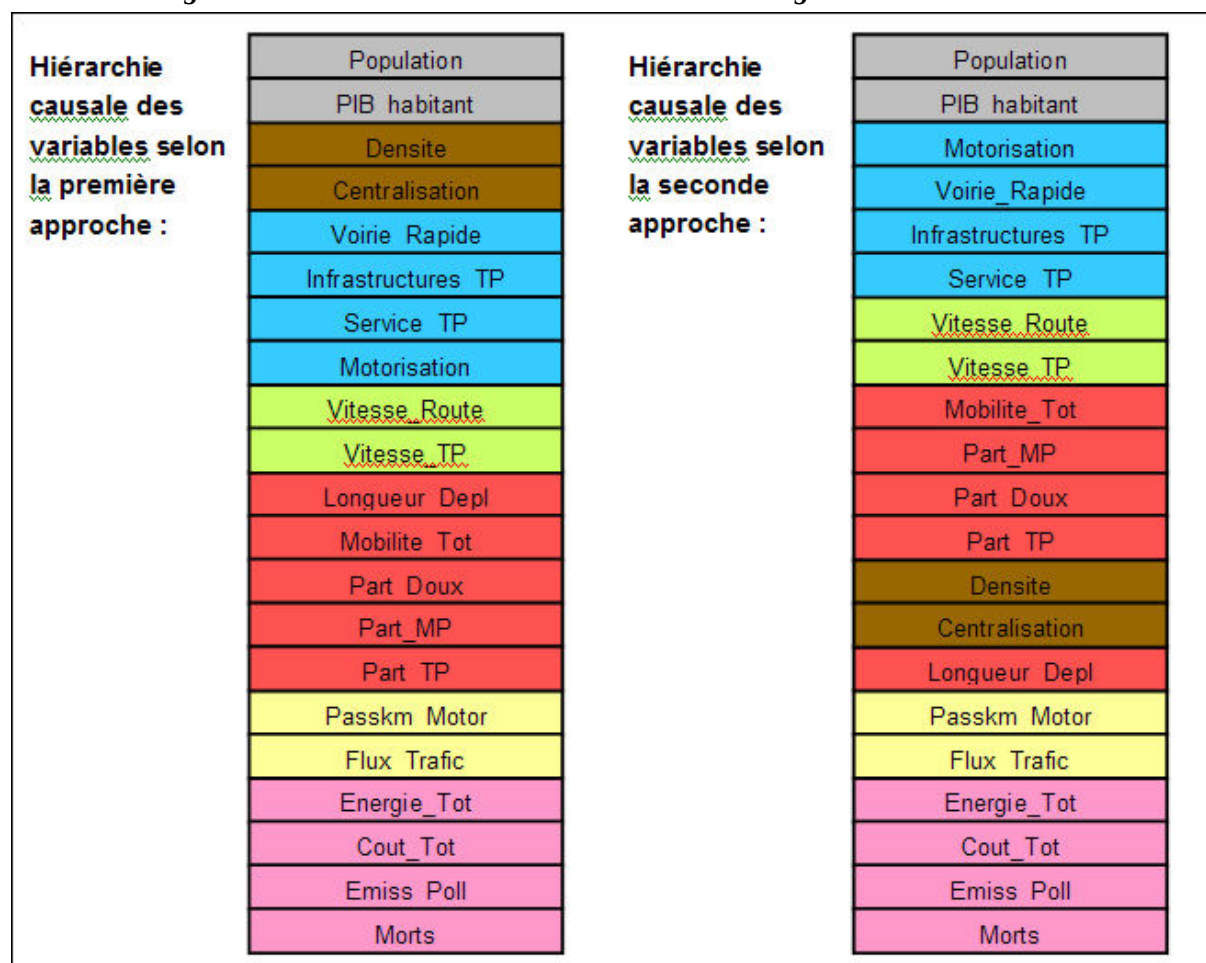


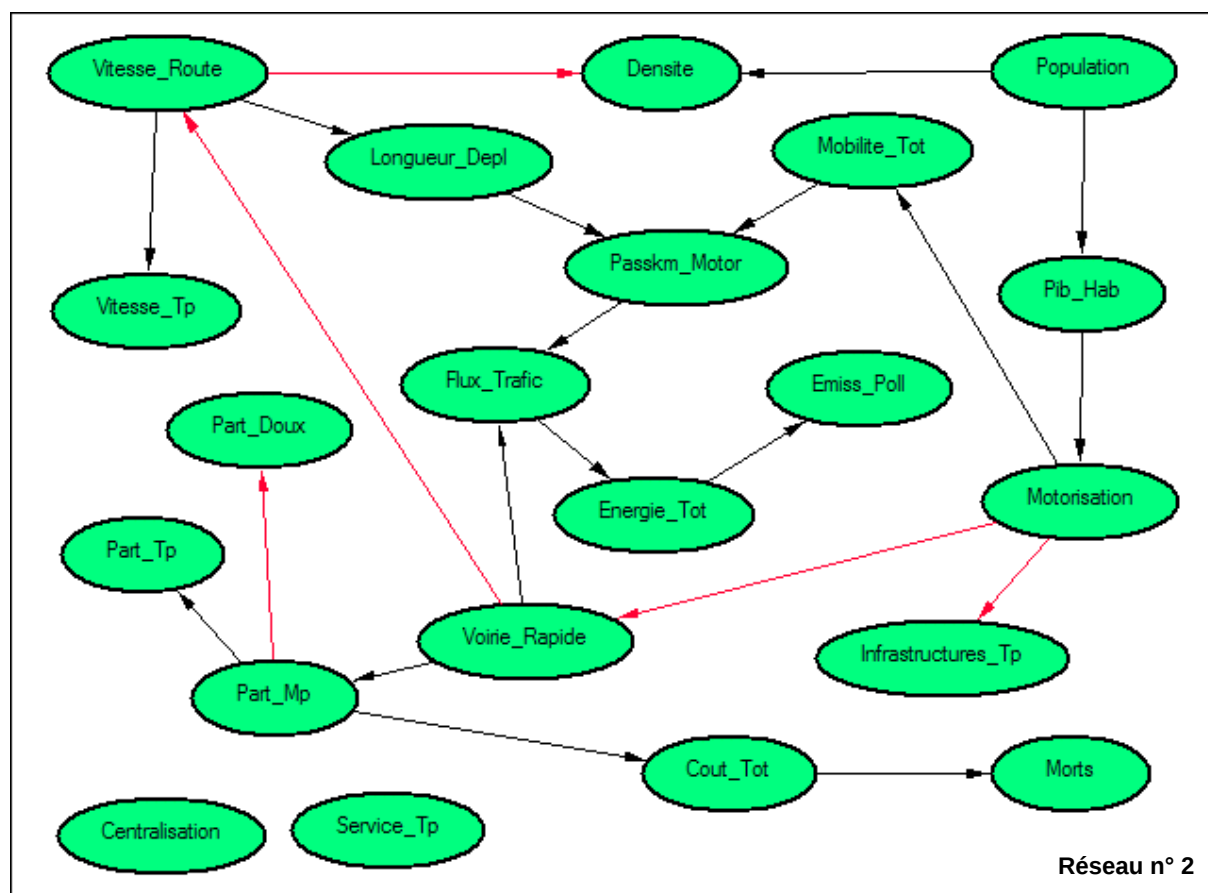
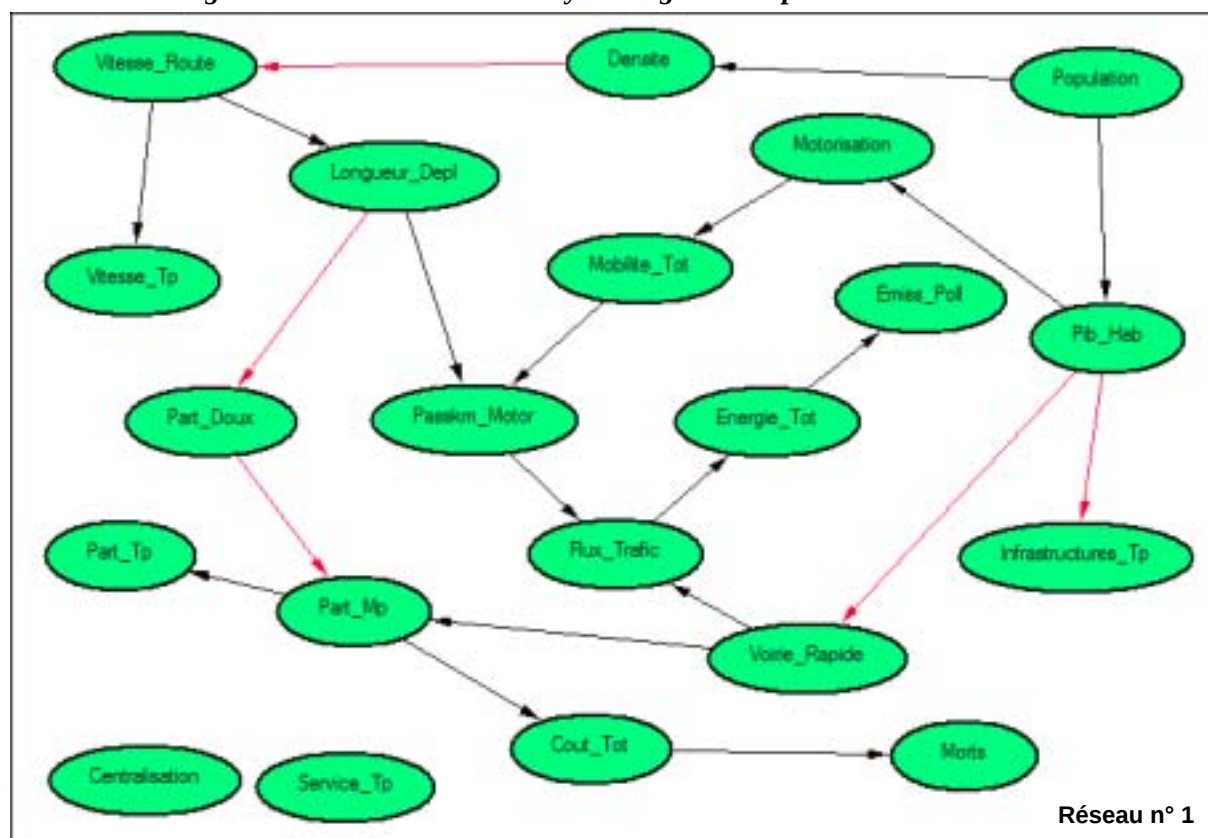
Figure 4 – Les hiérarchies causales utilisées dans la génération des réseaux



Combinées aux données réelles des 102 grandes agglomérations mondiales, les deux hiérarchies causales produisent les réseaux bayésiens présentés dans la figure 5. Les deux réseaux sont globalement bien connectés et donc capables d'expliquer la plupart des variables étudiées. Ils contiennent chacun 21 liens causaux reliant 19 des 21 variables étudiées. En effet, la centralisation des emplois et l'offre de services de transport public restent déconnectées du reste du réseau, quelle que soit la hiérarchie assumée : tels qu'ils sont définis, ces indicateurs ne semblent jouer aucun rôle à l'intérieur du système.

Le premier réseau montre clairement que les variables démographiques et économiques jouent un rôle important dans le fonctionnement du système de la mobilité urbaine. En réalité, l'importance de la taille démographique doit être mise en perspective : les tables de probabilités conditionnelles montrent que les villes plus petites ont tendance à être moins denses et surtout plus riches. Cela s'explique par le fait que les villes de plus petite taille dans la base de données (celle de moins de 1 500 000 habitants) sont principalement des villes des pays industrialisés. Il s'agit donc d'un biais de la base de données. Le rôle du PIB métropolitain par habitant est en revanche plus déterminant et moins contestable : il explique à la fois la motorisation de la population et l'offre d'infrastructures, qui sont normalement supérieures dans les villes les plus riches. En ce qui concerne la motorisation, par exemple, sa table des probabilités conditionnelles (figure 6) indique qu'une ville au faible PIB par habitant (moins de 9145 \$) a presque 80% de chance d'avoir une motorisation très faible (entre 13 et 284 véhicules par 1000 habitants). Pour des valeurs supérieures du PIB, la motorisation est soit moyenne (284-479 véhicules par 1000 habitants), soit élevée (479-756 véhicules par 1000 habitants).

Figure 5 – Les deux réseaux bayésiens générés à partir des 21 variables



→ lien causal existant dans les deux réseaux → lien causal existant dans un réseau seulement

Figure 6 – Table des probabilités conditionnelles de la variable Motorisation dans le réseau n°1

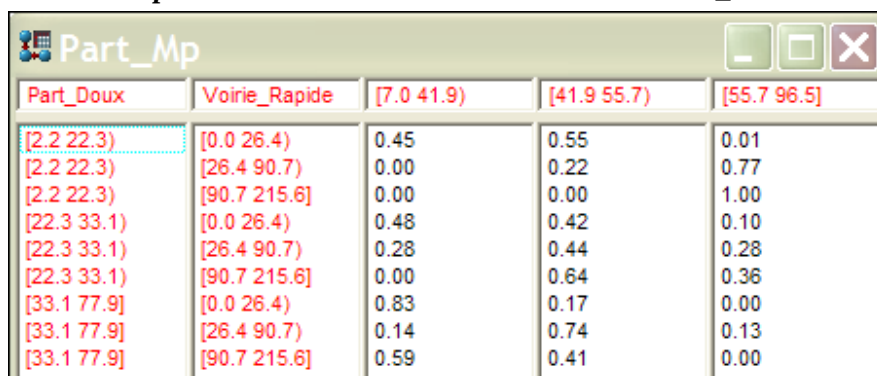


Pib_Hab	[13.0 284.0]	[284.0 479.0]	[479.0 756.0]
[396.0 9145.0]	0.79	0.21	0.00
[9145.0 28009.0]	0.13	0.36	0.51
[28009.0 54692.0]	0.04	0.42	0.54

En poursuivant la description du réseau, la densité explique la vitesse du système routier : en considérant que la longueur des infrastructures routières par hectare de surface urbanisée ne varie pas sensiblement d'une ville à l'autre (Kenworthy et Laube 1999), les villes les plus denses voient leurs réseaux empruntés par un nombre supérieur d'utilisateurs d'où leur moindre vitesse moyenne. La vitesse (fondamentalement celle du réseau routier) détermine à son tour la longueur des déplacements, en accord avec la théorie du budget-temps constant de Zahavi (1974) : les plus grandes vitesses sont utilisées pour allonger les distances parcourues et non pas pour diminuer le temps de déplacement. Cela a des conséquences directes sur le choix modal, la part de la marche et du vélo diminuant au fur et à mesure que les distances moyennes parcourues augmentent.

La motorisation de la population détermine la mobilité totale moyenne des habitants. Il s'agit d'un phénomène connu (Zahavi 1974, Merlin 1984) : dans les villes à forte motorisation, la plus grande mobilité motorisée compense largement une mobilité de proximité (à pieds, en vélo) moindre. Mobilité totale et longueur des déplacements expliquent tout logiquement les passagers-Km parcourus dans les modes motorisés (les déplacements les plus longs se font toujours dans les modes motorisés).

Figure 7 – Table des probabilités conditionnelles de la variable *Part_MP* dans le réseau n°1



Part_Doux	Voirie_Rapide	[7.0 41.9]	[41.9 55.7]	[55.7 96.5]
[2.2 22.3]	[0.0 26.4]	0.45	0.55	0.01
[2.2 22.3]	[26.4 90.7]	0.00	0.22	0.77
[2.2 22.3]	[90.7 215.6]	0.00	0.00	1.00
[22.3 33.1]	[0.0 26.4]	0.48	0.42	0.10
[22.3 33.1]	[26.4 90.7]	0.28	0.44	0.28
[22.3 33.1]	[90.7 215.6]	0.00	0.64	0.36
[33.1 77.9]	[0.0 26.4]	0.83	0.17	0.00
[33.1 77.9]	[26.4 90.7]	0.14	0.74	0.13
[33.1 77.9]	[90.7 215.6]	0.59	0.41	0.00

La part des modes motorisés privés dans les déplacements totaux est expliquée à la fois par la part des modes doux et par l'existence de voirie rapide urbaine pour l'automobile. La table des probabilités conditionnelles décrivant cette double dépendance causale est montrée dans la figure 7. Quand la part des modes doux est faible, une offre de voirie rapide moyenne suffit déjà à reporter sur les modes motorisés privés presque quatre déplacements sur cinq. En présence d'une utilisation plus importante des modes doux, l'offre de voirie rapide urbaine continue à expliquer le rôle grandissant des modes motorisés individuels, mais la part de ces derniers reste presque toujours inférieure à 55%. Tout se passe comme si, après avoir éliminé les déplacements de proximité, l'offre de voirie rapide était capable d'influencer le choix modal entre la voiture et les transports en commun. La part de ces derniers dépend de façon résiduelle de celle des modes motorisés privés. Elle semble donc être davantage influencée par l'offre d'infrastructures pour la voiture que par l'offre d'infrastructures propres aux transports en commun. Ce résultat est confirmé par l'explication des flux de trafic de véhicules. A parité de passagers-Km dans les modes motorisés, la présence de voirie rapide fait augmenter la part des passagers-Km dans les modes motorisés privés : il en résulte une

augmentation du nombre de véhicules en circulation et donc des flux exprimés en véhicules-Km. Ces derniers sont directement responsables de la consommation énergétique et par là-même des émissions des polluants atmosphériques. Finalement, ce premier réseau bayésien montre que les coûts et les externalités du système de la mobilité urbaine sont strictement liés à l'utilisation des modes motorisés privés et à l'offre de voirie rapide.

En tenant compte de l'enchaînement causal global, ce premier réseau bayésien assigne un rôle déterminant tant à la vitesse de la route qu'à l'offre de voirie rapide. La densité urbaine reste en tout cas un élément crucial en tant que variable explicative de la vitesse du réseau routier, conformément à l'analyse de Peter Newman et Jeffrey Kenworthy (1998).

Le second réseau bayésien garde les trois quarts des liens causaux du premier. Rien ne change, par exemple, dans l'explication des flux de trafic, des coûts et des externalités de la mobilité urbaine. Dans la partie plus en amont du processus causal, les différences sont néanmoins importantes. Deux liens changent de direction : désormais, c'est la vitesse du réseau routier qui détermine la densité de la ville et non pas le contraire. De même, l'ordre causal entre la part des modes motorisés privés et la part des modes doux est inversé. Enfin, trois liens disparaissent et sont remplacés par d'autres relations causales : la motorisation devient une variable cruciale, déterminant tant la mobilité totale que l'offre d'infrastructures, tandis que la vitesse du réseau routier est justement influencée par l'offre de voirie rapide.

En général, le second réseau montre que la motorisation de la population, si elle est placée en haut de la hiérarchie causale, devient un facteur déterminant dans l'explication du fonctionnement du système de la mobilité urbaine. Elle détermine directement le niveau de la mobilité totale de la population. Indirectement, étant à la base des décisions politiques en matière d'infrastructures de transport, elle définit également le partage modal de la mobilité urbaine. En outre, l'offre de voirie rapide détermine la vitesse moyenne du réseau routier et cette dernière devient une variable fondamentale dans l'explication de la densité urbaine. En dernière instance, la motorisation (donc l'automobile) semble déterminer tant les habitudes modales de la population que la structure de la ville.

Or, quel réseau explique le mieux les données sur la mobilité dans les grandes villes du monde ?

Discoverer fournit pour chaque réseau la valeur de sa log-vraisemblance, c'est-à-dire le logarithme de sa probabilité a posteriori sachant les données. Compte-tenu du nombre restreint de cas d'étude par rapport au nombre de variables, les deux réseaux ont des log-vraisemblances très faibles : -2018,6 pour le premier, -2001,3 pour le second. La différence de 17,3 est néanmoins considérable : elle signifie que le second réseau est 485 millions de fois plus probable sachant les données de notre base d'exemples. Le modèle de la ville produite par sa mobilité et, en premier lieu, par la motorisation de ses habitants, semble être donc beaucoup plus probable que celui de la ville déterminant sa mobilité.

On remarquera que la log-vraisemblance des deux réseaux est particulièrement modeste du fait des deux variables non expliquées par le modèle causal. L'algorithme de génération appliqué seulement aux 19 variables connectées produit deux réseaux bayésiens identiques à ceux de la figure 5, mais avec une log-vraisemblance de -1801,9 et -1784,6, respectivement. Le second réseau est toujours 485 millions de fois plus probable que le premier, mais chaque réseau est $1,25 \times 10^{94}$ fois plus probable sachant les 19 variables au lieu des 21.

Par le test de la *cross validation*, *Discoverer* permet également de tester la force prédictive des modèles bayésiens trouvés à l'aide d'une seule base de données. Le test divise d'abord la base des données en k parts et, pour chaque part, prédit les valeurs d'un sous-ensemble de variables en utilisant les tables des probabilités conditionnelles estimées sur le reste de la base de données. Pour chaque réseau, nous avons effectué deux de ces tests : une fois pour prédire les valeurs des trois parts modales (Part_MP, Part_Doux et Part_TP) et une fois pour prédire les valeurs des coûts et des externalités du système de la mobilité (Energie_Tot, Emiss_Poll,

Cout_Tot et Morts). La base de données a été divisée chaque fois en cinq parts, et chaque test a été conduit dix fois pour obtenir des résultats statistiquement plus robustes. Les résultats sont affichés dans le tableau 3. La force prédictive des deux réseaux est fondamentalement équivalente : bonne pour la prédiction du partage modal des déplacements (environ 87% des valeurs prédites sont correctes), un peu moins bonne pour la prédiction des coûts et des externalités (environ 77% de valeurs correctes).

Tableau 3 – Résultats des tests de "cross validation" pour les deux réseaux après 10 itérations

Variables à prédire	Réseau n°1	Réseau n°2
Part_MP, Part_Doux, Part_TP	84,38%	85,39%
Energie_Tot, Emiss_Poll, Cout_Tot, Morts	77,36%	76,81%

3. Conclusion

La technique des réseaux bayésiens nous apparaît très prometteuse pour les analyses territoriales. Elle permet de représenter la connaissance systémique des interrelations entre variables en couplant un formalisme graphique de type qualitatif avec un formalisme mathématique de type probabiliste. En particulier, la possibilité de débusquer « automatiquement » des relations causales à partir d'une base de données fait des réseaux bayésiens un puissant outil de *knowledge discovery*.

Dans notre étude, nous avons montré que la génération de réseaux bayésiens à partir d'une base de données, pour aboutir à des structures pertinentes, doit normalement intégrer des hypothèses du modélisateur sur les hiérarchies causales des phénomènes étudiés. Ces hypothèses ont d'importantes conséquences sur les réseaux générés.

Dans le cas de l'étude de la mobilité urbaine dans les grandes villes du monde, deux hypothèses ont été testées. Selon la première, la ville détermine sa mobilité, selon la seconde, la mobilité détermine sa ville. Les deux réseaux bayésiens correspondant à ces hypothèses sont bien connectés et semblent prendre en compte correctement les données de la base d'exemples. Le réseau correspondant à la seconde hypothèse s'avère en revanche beaucoup plus probable. Dans ce réseau, la motorisation de la population joue un rôle particulièrement crucial dans l'explication à la fois des caractéristiques de la mobilité quotidienne et de la structure urbaine. La densité urbaine est, par contre, une variable explicative importante seulement dans le premier réseau. Finalement, les deux réseaux concordent pour assigner à l'offre de voirie rapide urbaine un rôle fondamental dans l'explication du partage modal et par là-même des coûts et des externalités des transports.

L'utilisation que nous avons faite des réseaux bayésiens dans la recherche de liens causaux entre indicateurs de mobilité urbaine sert ici d'évaluation de la démarche. Plutôt que proposer des résultats définitifs, elle fournit des pistes d'explication qui méritent d'être approfondies ultérieurement. En particulier, le choix des indicateurs retenus a été fortement conditionné par les données à notre disposition. En outre, les villes incluses dans la base de données utilisée appartiennent à des espaces géographiques très divers et se trouvent à différentes étapes du développement socio-économique et infrastructurel. Cela rend difficile la recherche d'un modèle unique de fonctionnement du système de la mobilité urbaine. Des modèles bayésiens pourraient ainsi être développés pour différents types de ville, au moins pour chaque macro-région géographique (Europe, Amérique du Nord et Océanie, Amérique Latine, etc.). La pertinence des hypothèses sur les hiérarchies causales parmi les indicateurs pourrait alors être évaluée par rapport aux caractéristiques du développement urbain de chaque type de ville. L'élaboration de modèles bayésiens pour des types de villes différents nécessite cependant des bases de données urbaines bien plus vastes que celle qui était à notre disposition.

BIBLIOGRAPHIE

- BAYESWARE (2000) *Bayesware Discoverer - User Manual*, London, Bayesware Limited.
- BECKER A., NAÏM P. (1999) *Les réseaux bayésiens – Modèles graphiques de connaissance*, Paris, Eyrolles.
- CAVARROC M.-A., JEANSOULIN R. (1998) *L'apport des réseaux bayésiens pour la recherche de causalités spatio-temporelles*, Assises Cassini'98, Marne la Vallée.
- COOPER G., HERSOVITS E. (1992) *A Bayesian Method for the Induction of Probabilistic Networks from Data*, *Machine Learning*, n° 9, 309-347.
- DRUZDZEL M.J., SIMON H.A. (1993) *Causality in Bayesian Belief Networks*, in "Proceedings of the Ninth Annual Conference on Uncertainty in Artificial Intelligence (UAI-93)", San Francisco, Morgan Kaufmann, 3-11.
- DURAND-DASTES F. (1985) *A propos du déterminisme : la complexité de l'ensemble des options possibles*, *L'Espace Géographique*, tome XIV, n° 2, Montpellier, RECLUS, 105-108.
- DURAND-DASTES F. (1999) *Les systèmes en géographie, remarques introductives à la séance thématique organisée par J. Dresch le 2 décembre 1979*, BAGF, n°2, Paris, 224-226.
- FLINT C., HARROWER M., EDSALL R. (2000) *But How Does Place Matter ? Using Bayesian Networks to Explore a Structural Definition of Place*, paper presented at *New Methodologies for the Social Sciences : The Development and Application of Spatial Analysis for Political Methodology*, University of Colorado at Boulder, March 10th-12th 2000.
- JENSEN F.V. (2001), *Bayesian Networks and Decision Graphs*, New York, Springer.
- KENWORTHY J., LAUBE F. (dir) (1999) *An International Sourcebook of Automobile Dependence in Cities 1960-1990*, Boulder, University Press of Colorado.
- MERLIN P. (1984), *La planification des transports urbains*, Paris, Masson.
- NEWMAN P., KENWORTHY J. (1999) *Sustainability and Cities – Overcoming Automobile Dependence*, Washington, Island Press.
- PEARL J. (2000), *Causality – Models, Reasoning and Inference*, Cambridge, Cambridge University Press.
- PEARL J., RUSSELL S. (2000) *Bayesian Networks*, working paper, Los Angeles, University of California.
- PIKE W.A. (2001) *Modelling water quality violations with Bayesian networks*, MS Thesis in Geography, Pittsburgh, Pennsylvania State University.
- SCARLATTI F. (2002) *Statistiche testuali, mappe concettuali, reti bayesiane e valutazione del paesaggio*, XXIII Conferenza Italiana di Scienze Regionali, Reggio Calabria, 10-12 ottobre 2002.
- STASSOPOULOU A., PETROU M., KITTLER J. (1996) *Application of a Bayesian Network in a GIS based Decision Making System*, working paper, Guildford, University of Surrey.
- UITP, ISTP (2001) *The Millennium Cities Database for Sustainable Transport*, Bruxelles, Union Internationale des Transports Publics.
- WADDEL P. (2001) *Towards a Behavioral Integration of Land Use and Transportation Modelling*, 9th International Association for Travel Behaviour Research Conference, Queensland, Australia, January 2001.
- ZAHAVI Y. (1974) *Travel characteristics in cities of developing and developed countries*, working paper, Washington, The World Bank.

ⁱ Article présenté aux 6^{ème} Rencontres de ThéoQuant, Besançon, 20-21 février 2003.