

## **Génération des émissions et attractions à l'échelle des aires de rabattement du réseau TER**

Déterminer les émissions et attractions liées à un service de transport en commun nécessite d'élaborer des aires de rabattement au tour des points d'entrée de ce réseau. Il existe énormément de facteurs pouvant, à un moment ou un autre du processus décisionnel, expliquer le choix d'un point de rabattement plutôt qu'un autre. Aux vues de la très grande quantité d'informations à exploiter, nous avons dû mettre en place, pour la modélisation du découpage de l'espace, un degré d'automatisation élevé. Mais nous avons tenté de conserver un regard critique sur le processus comme le résultat de cette modélisation afin d'éviter l'effet « boîte noire », couramment rencontré lors de la modélisation de systèmes aussi complexes que la mobilité quotidienne et son organisation spatiale.

Nous avons fait appel à des méthodes issues de disciplines non géographiques comme la biologie par le biais des algorithmes métaheuristiques. Les choix de modélisation défendus ici sont donc ceux d'une pleine utilisation des possibilités informatiques nouvelles combinées au refus d'une automatisation aveugle et souvent vide de sens.

C'est dans cet esprit que, après avoir rappelé le rôle fondamental de la partition spatiale dans la gestion des transports, nous présenterons notre protocole de modélisation lié à la partition de l'espace dans la problématique de la mobilité quotidienne. Enfin, nous situerons notre propos par la mise en œuvre des résultats fournis par le modèle.

## **1. La problématique du zonage, outil de compréhension territoriale**

Il a été montré précédemment que la mobilité quotidienne était liée à une multitude de facteurs combinés par les individus pour façonner leur processus de choix. Ce processus peut se diviser en plusieurs étapes. Quelle que soit l'étape à étudier dans ce processus, la modélisation qui en découle dépend toujours des zones que l'on a définies comme génératrices des déplacements. Le décompte des émissions et attractions, lors de la phase de génération, s'en trouve modifié, tout comme l'étape de distribution ne fournit pas les mêmes résultats selon les limites ou la taille des zones ; enfin l'affectation de trafic et le choix modal, conçus autour de fonctions d'utilité basées sur des temps de parcours, se trouvent changés en fonction du zonage effectué. Il s'agit donc ici de définir les éléments spatiaux servant de base à l'ensemble de la modélisation de trafic.

Le choix du zonage doit suivre certaines contraintes tout en répondant à certains objectifs, fixés au préalable, et en accord avec l'ensemble du processus de modélisation. Ainsi, chaque zone doit présenter une certaine cohérence interne, une certaine homogénéité, garantissant une logique lors de l'application de ces frontières. Au-delà de cette cohérence interne, une cohérence avec le réseau s'impose également : soit avec le réseau dans son ensemble, soit avec un réseau spécifique dans le cas d'étude centrée sur un mode particulier. Enfin, le zonage choisi doit garantir une certaine applicabilité : sur la disponibilité des données d'une part qui doivent fournir une certaine compatibilité spatiale, et sur des temps de calcul réalistes d'autre part.

Lors de la première partie, nous avons discuté de la définition du périmètre de l'aire d'étude en géographie des transports et plus particulièrement dans le cas de la gestion de la mobilité quotidienne. Cette définition revient à déterminer un isolat qui nous conduit à considérer ce qui est extérieur à ce périmètre comme exogène.

Le zonage est lié à cette définition de périmètre mais trouve une traduction différente en terme de modélisation. Il est mis en place afin d'agréger certaines données liées à l'individu ou au transport (temps d'accès au réseau...). On abandonne ainsi toute information désagrégée sur le plan spatial à l'intérieur de la zone. De ce fait, on a longtemps considéré que le découpage le plus fin possible fournissait la garantie d'une modélisation optimale. Le frein principal à ce désir de détail résidait alors dans le coût d'acquisition des données, le

temps de travail que représentait leur traitement, et connaissait la plus grande entrave par le temps de calcul nécessaire aux ordinateurs.

L'amélioration récente des capacités de traitement du matériel informatique nous ramène souvent à des temps de calcul réalistes pour un zonage d'un haut niveau de détail. Cette amélioration pousse souvent les concepteurs de modèles de simulation de trafic à faire le choix du zonage le plus fin qu'autorisent les temps de calcul. Ce choix, servant de base à l'ensemble des calculs opérés au cours des quatre étapes de la modélisation, n'est pourtant pas obligatoirement le plus judicieux en fonction des objectifs du modèle. L'étude du rôle, parfois multiple, de ce zonage va donc nous permettre d'avoir un œil critique sur les différentes solutions déjà appliquées dans ce domaine.

Notre travail s'articulera donc autour de l'enjeu du partitionnement spatial dans le modèle à quatre étapes et pour toute gestion d'un système de transport, ce qui nous amènera par la suite à faire un tour d'horizon des solutions existantes et des différentes possibilités de découpage déjà mises en œuvre en géographie des transports.

### **1.1. Le partitionnement spatial dans le modèle à quatre étapes**

Un découpage, fonction de ses objectifs

Le choix d'un zonage doit clairement se faire en accord avec les objectifs fixés pour la modélisation. Cela ne veut pas dire pour autant que les zones définies vont permettre tout type d'analyse mais bien qu'elles doivent permettre de répondre aux questions que l'on se pose. La modélisation liée à la gestion du trafic autoroutier, par exemple, amène à découper le territoire en fonction des zones de chalandise de chaque axe autoroutier. On peut ainsi « isoler » les territoires et les individus concernés par cette infrastructure. Dans le cas où les objectifs du modèle sont axés sur les aménagements des échangeurs, il est souhaitable de disposer de zones définissant les aires d'attraction de chacun de ces échangeurs. On peut par la suite imaginer des zonages beaucoup plus complexes, prenant en compte les problématiques de correspondance ou des études liées à des générateurs spécifiques de mobilité qu'il est bon de pouvoir isoler dans des zones spécifiques.

### Un zonage disposant d'entités cohérentes et homogènes

La cohérence des zones, dans le cas de modèles désagrégés, (à l'exception de l'étape de génération de trafic pour laquelle, comme nous l'avons vu, il est difficile d'appliquer un tel modèle), ne présente pas une nécessité flagrante. On évite, avec une telle modélisation, l'approximation due à l'illusion de l'individu moyen d'une zone qui reste plus ou moins forte pour les modèles agrégés. Il est alors utile pour ces derniers de pouvoir prétendre à des zones présentant une certaine cohérence interne par rapport aux variables explicatives de la mobilité. Il est par exemple judicieux de séparer les zones d'activités des zones d'habitat dans la mesure où les données le permettent. On peut même imaginer séparer, principalement en milieu urbain, des zones ne présentant pas un comportement de mobilité uniforme.

On fait souvent appel aux frontières naturelles ou à toute infrastructure qui se présentent comme un frein à la mobilité entre les espaces situés de chaque côté. On obtient ainsi de possibles ensembles présentant une cohérence comportementale.

### Un zonage organisé autour des réseaux

Comme nous l'avons déjà illustré avec l'exemple de la gestion du trafic autoroutier, l'organisation du zonage en fonction des réseaux est fortement conditionnée par les objectifs du modèle. Pour tout modèle lié à un réseau spécifique, l'organisation du zonage correspond aux aires de chalandise de ce réseau. Le problème est un peu différent dans le cas d'une modélisation de la mobilité pour l'ensemble des modes et concernant donc l'ensemble des réseaux. Il est globalement recommandé de limiter à la fois le nombre de déplacements émis ou attirés au regard des infrastructures d'accès à traiter, et le nombre de points d'entrée sur le réseau par zone. Certains réseaux comme le réseau routier classique fournissent toutefois une irrigation du territoire telle, qu'il paraît très difficile de disposer de données à une échelle aussi fine. On peut alors organiser le zonage en fonction du réseau représentant une part forte du trafic global et isoler au maximum les points d'accès à ce réseau tout en conservant l'échelle la plus fine possible.

Le choix de zones de taille réduite nous amène à soulever un autre aspect lié au réseau et à l'organisation du découpage qui en découle. Au-delà des aires de chalandise (des points d'entrée sur le réseau ou du réseau lui-même), le zonage doit tenir compte de la structure du réseau. Il s'agit en fait de définir des zones

pour lesquelles le temps d'accès au réseau principal est relativement uniforme. Ce problème est commun aux approches agrégées et désagrégées puisque la localisation au sein de la zone reste agrégée, la désagrégation ne portant que sur les données socio-économiques. On traite donc, à l'intérieur de chaque zone, une distance moyenne et un temps d'accès moyen au point d'entrée sur le réseau.

#### L'adéquation entre découpage et données disponibles

L'ensemble des critères vus jusqu'ici conditionnant le zonage nous pousse à préférer des zones de taille réduite en adéquation avec les sorties attendues du modèle. Les considérations évoquées maintenant nous ramènent souvent à des ambitions plus mesurées.

Concernant le modèle de génération de trafic, en premier lieu, il est fondamental de disposer de zones représentant chacune un nombre de déplacements suffisants pour permettre un calage fiable du modèle. Faire le choix de zones trop réduites pour représenter une population statistique suffisante nous amène à considérer démesurément des phénomènes locaux, voire ponctuels temporellement. Ainsi, la précision que l'on pensait gagner par un découpage fin n'est qu'une illusion en conséquence d'une estimation faussée par le modèle.

Il est également important de prendre en compte la correspondance entre le découpage et la précision spatiale des variables socio-économiques explicatives de la mobilité. Cette correspondance ne pose pas de problèmes importants concernant la modélisation du trafic interurbain. La plupart de ces données sont disponibles à l'échelle communale et permettent une analyse à ce niveau de détail. En milieu urbain, certaines données sont disponibles à l'échelle de l'îlot ; toutefois, il est rare de disposer de l'ensemble des données à un tel niveau de détail. De plus, que la modélisation porte sur de la mobilité intra ou interurbaine, travailler à un niveau très fin est totalement illusoire et en inadéquation complète avec le but prévisionnel des modèles. Plus la population traitée est faible, plus les marges d'erreur en prévision sur les données socio-économiques sont fortes. On perd ainsi, à l'horizon d'étude, le détail gagné en situation de référence.

Certains modèles font le choix d'élaborer différents découpages en fonction de l'étape à traiter dans la modélisation. On fait ainsi appel à un découpage plus grossier concernant l'étape de génération, la plus problématique en terme de

disponibilité de données comme en terme de prévision. Puis on construit un zonage plus fin pour les étapes de distribution et d'affectation, afin de réduire au maximum le biais lié à la fonction d'utilité. Cette manière de procéder ne fait que déplacer le problème puisqu'on introduit une marge d'erreur supplémentaire dans la procédure de répartition des origines et destinations en désagrégeant l'information agrégée à l'étape de génération.

### La limite technique

Le temps nécessaire à la production des données mais également à l'analyse permettant le calage des différents modèles propres à chaque étape croît énormément avec l'augmentation du nombre de zones à traiter. Et même si, comme nous l'avons vu, les progrès informatique permettent de traiter un nombre considérable d'informations, le temps alloué à l'étude conditionne fortement la précision du zonage que l'on souhaite développer.

Il faut prendre garde de ne pas se laisser aveugler par les possibilités techniques et bien considérer les avantages et inconvénients des différents niveaux de détail, plus particulièrement avec les sorties attendues de la modélisation. On rejoint ici les problématiques liées à la géographie dite « computationnelle » qui s'inscrit dans une philosophie de la technique. Elle découle de la notion de *technique originaire*, à savoir que la technique possède une identité propre dans le monde de la recherche, jouant un rôle constitutif dans l'évolution des connaissances, voire de la pensée. L'outil est donc constituant de sciences. La géographie computationnelle correspond donc à la géographie du possible constituée et portée par les nouveaux supports techniques et dynamiques. Chaque technique permet donc de constituer de nouveaux ensembles conceptuels élargissant les possibilités géographiques.

Il nous paraît risqué de laisser les possibilités techniques nous guider vers des solutions plus complexes à mettre en œuvre, et ne garantissant que rarement une amélioration à terme du pouvoir prévisionnel des modèles de simulation de trafic.

## **1.2. Les solutions existantes**

Les solutions existantes concernant un découpage spatial optimal dans le cadre de la modélisation de trafic sont relativement restreintes. La plupart des modèles interurbains récents traitent l'information au niveau le plus fin que

leur permet la disponibilité des données. Ils sont donc classiquement centrés sur l'échelle communale. Certains modèles, comme le modèle MODEV (un modèle géographique multimodal de transport de marchandises), pour lequel le zonage est basé sur les zones d'emplois, fonctionnent sur des zones spécifiques en lien avec leurs objectifs.

La plupart des modèles intra urbains, quant à eux, sont la plupart du temps basés sur les îlots en raison d'une adéquation plus facile avec les découpages de l'INSEE.

Certains zonages ont été mis en place pour la gestion spécifique de transports collectifs. Dans la majorité des cas, l'application de buffers autour des points d'entrée sur le réseau est utilisée. Cette technique, bien que ne prenant pas en compte les différents éléments détaillés précédemment, fournit l'avantage de garantir une relative homogénéité pour les temps d'accès à ces connexions au réseau. Ces découpages sont donc plus adaptés à l'étape d'affectation de trafic qu'à celle de génération, pour laquelle la question du zonage n'est que rarement traitée en tant que telle.

D'autres applications ont vu le jour pour tenter de répondre mieux à ces problématiques de zonage. On peut citer le zonage mis en place sur l'agglomération lyonnaise [SEMALY 2000] en 196 zones, élaboré de manière intuitive en respectant au mieux une certaine homogénéité des zones ainsi qu'une identification des pôles de rabattement.

Il n'existe à notre connaissance pas de procédure automatisée prenant en compte la cohérence interne de chaque zone, la structure des réseaux et de leurs points d'encrage au territoire, adaptée à la gestion d'un mode particulier, attaché à une infrastructure ou un réseau qui lui est propre.

Nous présenterons, au cours du chapitre prochain, une procédure de découpage spatial adapté à la gestion d'un service de transport en commun, fondée sur la cohérence locale et globale du zonage.



## **2. Le modèle de structuration spatiale, l'optimisation du rabattement**

Avant de mettre en place un protocole de modélisation pour la structuration spatiale de notre territoire d'étude, nous avons mené une réflexion préalable sur la complexité et la difficulté d'appliquer un modèle à un tel système. Nous avons tenté, en réalité, de porter un regard critique sur la modélisation des systèmes complexes, sur les difficultés qu'elle engendre et sur les erreurs qu'elle peut introduire.

Nous présenterons, par la suite, la méthodologie mise en place pour la construction des aires de rabattement autour des points d'entrée du réseau TER. Il nous a fallu, pour mener à bien ce processus complexe de modélisation, procéder par étapes. Nous avons dans un premier temps identifié des indicateurs pertinents pour guider la construction des agrégats. Dans un second temps, nous avons fait appel à une méthode classique de partitionnement basée sur la solution optimale commune par commune. Enfin, dans un troisième temps, notre réflexion s'est portée sur un autre niveau d'échelle pour transcender le simple optimum local et tenter d'évaluer le zonage dans son ensemble.

### **2.1. Modélisation des territoires de la mobilité et complexité**

La compréhension et la modélisation du transport de personnes nécessitent de « penser la complexité ». Il faut donc réussir à admettre un mode de raisonnement –voire même un état d'esprit– complexe en acceptant ses implications, en percevant ses possibilités et en gardant en tête ses restrictions.

#### *2.1.1. La mobilité quotidienne, un phénomène complexe*

Comme nous l'avons déjà évoqué concernant les difficultés liées à sa modélisation, la mobilité quotidienne est un phénomène qui, par définition, est en mouvement permanent. Le modèle de génération de trafic mis en place auparavant a mis en évidence la multitude de variables explicatives de la mobilité et nous a montré que l'ensemble du phénomène n'était pas expliqué par le modèle. De fait, d'autres facteurs entrent en compte dans les déplacements et accroissent ainsi la difficulté de compréhension de ce phénomène.

De plus, la mobilité est soumise à une fluctuation constante, le choix d'un des individus conditionnant par ses conséquences une part du choix du reste de la population. Le système dans son entier est donc soumis à une instabilité. La notion de choix en elle-même nous pousse à considérer le phénomène de la mobilité comme complexe par son aspect flou. Le processus de choix de l'individu n'est en effet que rarement résumable à un comportement binaire.

### *2.1.2. Les difficultés liées à la modélisation de la complexité*

En tant qu'approche, la complexité est « fondée sur des nouvelles représentations de la réalité prenant en compte l'instabilité, l'ouverture, la fluctuation, le chaos, le désordre, le flou, la créativité, la contradiction, l'ambiguïté, le paradoxe » [Donnadieu et Karsky, 2002]. Ces différents concepts, souvent complémentaires, donnent une vision assez large mais réaliste de l'étendue de la complexité. Un système complexe peut donc être défini comme une entité cohérente selon certains critères mais dont les éléments, les interactions et la dynamique (de par leur très grand nombre et leur dynamique différente) engendrent des structures qui recèlent des surprises, de la nouveauté, que l'on ne peut pas définir *a priori*. Les systèmes complexes sont donc plus que la somme de leurs parties, la conséquence étant que tout modèle de leur structure ou de leur fonctionnement est nécessairement incomplet et partiel. Cependant, plus qu'une définition arrêtée du terme, c'est la philosophie et le mode de réflexion attachés au vocable qui nous intéressent ici.

Accepter la complexité signifie plus qu'admettre uniquement l'ensemble des concepts cités précédemment. « Penser la complexité, c'est affirmer à la fois que le monde échappe à nos concepts, les dépasse, tout en maintenant la possibilité d'agir, de schématiser provisoirement le réel » [Duits 1999]. Le mode de réflexion développé ici par Duits nous encourage à traiter avec circonspection toute étude d'un environnement complexe, sans toutefois réfuter toute possibilité d'analyse. Aucune grille de lecture n'est suffisante pour cerner la dimension complexe, mais une accumulation de plusieurs grilles, de multiples approches, permet de saisir de mieux en mieux les objets complexes (sociaux, comportementaux...) sans jamais les épuiser.

Ce mode de pensée, encourageant dans un processus de compréhension et de modélisation, soulève pourtant un véritable défi vis-à-vis de cette complexité. La multiplicité des approches et des critères pris en compte tend à développer une certaine inintelligibilité de la démarche d'analyse du système dans sa complexité. Ainsi, la tentative de représentation des modèles au plus près de la réalité complexe s'apparie de manière presque systématique avec le sacrifice de

la compréhension du processus de modélisation. Une conséquence directe de cette renonciation apparaît alors, avec l'aboutissement à des solutions potentiellement en décalage avec la réalité et donc à un amoindrissement du pouvoir prévisionnel du modèle.

L'enjeu de la modélisation de la complexité réside donc dans la capacité d'une procédure d'intégrer un grand nombre de données différenciées d'une part (et ainsi se rapprocher au plus près de la complexité du système) sans perdre le caractère intelligible de la démarche d'autre part, dans le but de conserver une possibilité de contrôle des prévisions proposées.

### *2.1.3. Proposition méthodologique de modélisation*

Plus qu'un simple modèle ou procédure, il est nécessaire de développer un véritable principe de raisonnement en adéquation avec un mode de pensée complexe. Le principe de complémentarité, développé dans les années 1920 par le physicien danois Niels Bohr, répond assez convenablement aux exigences présentes. Il reprend, de manière formalisée, le fameux dualisme esprit et matière de Descartes : « Les deux notions doivent être vraies en même temps, donc il doit exister deux royaumes de réalités ». Le principe de Bohr s'applique à la mécanique quantique dans la résolution de l'antinomie onde-particule<sup>18</sup>. Le principe de complémentarité, dans sa philosophie, consiste à résoudre le problème de la description des particules, en terme d'onde ou de corpuscule : pour ne pas trahir la conviction qu'il n'y a qu'une seule réalité, Bohr conclue que chaque description est vraie, mais dans des circonstances différentes. Cela ne signifie pas qu'il faille abandonner la métaphysique, mais plutôt qu'il faut reconnaître qu'un concept ne peut pas être défini par une seule phrase. Einstein, pour sa part, avait toujours refusé que la description du réel puisse dépendre de nos observations. Ceci aurait signifié qu'il fallait renoncer à parler d'un réel en soi. Or, adopter un raisonnement complexe n'autorise pas une telle renonciation.

Nous allons ici adapter à la problématique de description d'un système complexe le principe de complémentarité. Nous accepterons donc qu'il existe une vision de la réalité issue d'un modèle complexe ; assortie d'une autre vision de la réalité, découlant d'un modèle "simple" (en opposition au premier

---

<sup>18</sup> Une entité quantique peut être simultanément observée comme une particule et comme une onde, mais ces deux états sont mutuellement exclusifs à l'observation. Pour plus de détails, consulter [Rosenfeld L., article Bohr (Niels), Encyclopaedia Universalis, 1974] ou [Bohr N., Causality and Complementarity, in Philosophy of Science, 4, 1937]

modèle). Ces deux représentations de la réalité, sans s'inscrire dans des circonstances différentes, correspondent dans les faits à des angles de vue dissemblables. En fait, « en jouant des deux images, en passant de l'une à l'autre, en revenant à la première nous obtenons finalement l'impression juste sur l'étrange sorte de réalité qui se cache derrière nos expériences ». [Eisenberg]

Nous nous proposons donc ici de confronter deux visions, par deux prismes différents, d'un système complexe. Il s'agira en fait de fournir, par le moyen d'un modèle simple (et donc intelligible dans son déroulement), un contrôle des résultats obtenus par un système intégrant de multiples approches et critères.

## **2.2. Le modèle d'optimum local**

La première démarche, concernant le découpage de l'espace, répond au problème de l'affectation d'une commune sur un point d'entrée sur le réseau TER de manière purement locale. Il s'agit donc de mettre en place un protocole permettant d'évaluer la pertinence d'une commune en tant que part d'un agrégat de rabattement.

### *2.2.1. Définition de la fonction d'utilité du rabattement de chaque commune*

L'évaluation de chaque commune passe donc par la construction d'un indice mesurant la pertinence territoriale du rabattement effectué. Cet indice est construit à partir de la maximisation de l'utilité d'utilisation du transport régional sur lequel est rabattue la commune par rapport à une solution en transport personnel. Il s'agit donc en fait de mettre en évidence le point d'entrée sur le réseau régional de transport en commun le plus avantageux pour la commune concernée, en comparaison de l'utilisation d'une voiture personnelle. En clair, un rabattement est considéré comme bon lorsque la fonction d'utilité du transport en commun pour l'ensemble des individus concernés est acceptable et qu'elle reste relativement concurrentielle vis-à-vis de la fonction d'utilité de la voiture personnelle. On tente ainsi de ne pas privilégier une solution de transport en commun qui paraît acceptable au regard des autres possibilités en transport en commun, mais qui serait trop défavorable par rapport à une solution en véhicule personnel.

La fonction d'utilité du rabattement est basée sur le coût généralisé de la liaison. Ce coût généralisé, concernant le transport en commun, entre une commune origine  $i$  et une commune destination  $j$  est fonction de différents facteurs (de coût et de temps) indépendants les uns des autres. Nous sommes donc ici, dans la conception du modèle, dans une logique additive. Ces facteurs sont définis par :

- La valeur du temps de rabattement effectué en voiture personnelle :  $\lambda_{VP_1}$
- Le temps de rabattement en voiture personnelle entre la commune  $i$  et la gare  $G$  de rabattement :  $t_{iG}^{Rab}$
- La valeur du temps du trajet en transport en commun entre la gare  $G$  et la commune  $j$  :  $\lambda_{TC}$
- Le coefficient de pénibilité de l'attente au départ du transport en commun :  $p^{Att}$
- Le temps d'attente au départ du transport en commun :  $t^{Att}$
- Le temps trajet en transport en commun entre la gare  $G$  et la commune  $j$  :  $t_{Gj}^{TC}$
- Le coefficient de pénibilité de l'attente des correspondances :  $p_c^{Cor}$
- Le temps d'attente de la correspondance :  $t_c^{Cor}$
- Le coût du déplacement en transport en commun entre la gare  $G$  et la commune  $j$  :  $c_{Gj}^{TC}$
- Le coût du déplacement en voiture personnelle entre la commune  $i$  et la gare  $G$  :  $C_{iG}^{VP}$

La formule de coût généralisé liée au transport en commun s'écrit donc :

$$C_{ij}^{TC} = \lambda_{VP_1} \cdot t_{iG}^{Rab} + \lambda_{TC} (p^{Att} t^{Att} + t_{Gj}^{TC} + \sum_{c=1}^q p_c^{Cor} t_c^{Cor}) + c_{Gj}^{TC} + C_{iG}^{VP}$$

Dans le cas du choix modal de la voiture personnelle, la fonction de coût généralisé est fonction de :

- La valeur du temps du trajet en voiture personnelle entre la commune  $i$  et la commune  $j$  :  $\lambda_{VP_j}$

- Le temps de trajet en voiture personnelle entre la commune  $i$  et la commune  $j$  :  $t_{ij}^{VP}$
- Le coût du déplacement en voiture personnelle entre la commune  $i$  et la commune  $j$  :  $c_{ij}^{VP}$

La formule de coût généralisé liée à la voiture personnelle s'écrit donc :

$$C_{ij}^{VP} = \lambda_{VP_{ij}} \cdot t_{ij}^{VP} + c_{ij}^{VP}$$

Se pose alors le problème de l'évaluation des différents coefficients de pénibilité et du calcul des indices de valeur de temps. Cette étape se présente comme particulièrement importante ; les gains de temps représentent l'avantage principal que retirent les usagers au cours du déplacement [Hammadou Jayet, 2002]. Il va de soi que le temps n'est pas perçu de la même manière selon les conditions du transport, l'environnement dans lequel il est effectué, et le motif justifiant ce déplacement. Il existe différentes évaluations de la valeur du temps. On peut faire le choix de considérer le temps comme une ressource à utiliser ou comme une variable non affectée spécialement à une activité et dont la valeur se modifie en fonction d'un environnement spécifique [SMALL, 1982]

Le temps comme ressource est alloué à différentes activités et déplacements qui permettent à l'individu de produire des biens ou services. Il en résulte une valeur de ce temps, fonction de la ressource temporelle attachée à l'individu et de l'utilité retirée pour l'activité concernée [Becker, 1965]. On peut, à partir de cette acceptation du temps comme ressource, identifier plusieurs types ou valeurs de temps : la première se rapportant à la définition de Becker, reconnue comme ressource et attachée à une activité ; la seconde comme gain de temps dans une activité donnée. Il est alors nécessaire de définir une contrainte liée aux durées nécessaires à certaines activités ou à certains temps de transport [De Serpa, 1971].

Nous nous sommes appuyé sur les travaux de A. Causse, [Causse, 1999] afin de définir les différents indices de valeur de temps. Ces travaux se basent sur la théorie de l'allocation des ressources temporelles de Becker et définissent une valeur de temps fonction du mode de transport utilisé, du motif du déplacement et de la durée du déplacement. Ces motifs sont répartis en deux grands ensembles : les motifs professionnels et les motifs privés. Même si cette répartition peut paraître assez grossière, les motifs privés pouvant subir parfois une contrainte temporelle aussi importante que les motifs professionnels, elle

reste relativement adaptée à la perception générale du temps par les individus au cours de leurs déplacements.

	Voiture personnelle	Transport en commun
Motif privé	1,0	1,3
Motif professionnel	2,0	4,4

Tableau 23 : Valeur du temps en fonction du mode et du motif  
Source : d'après [Causse 99]

Le tableau 23 nous présente les coefficients relatifs à la valeur du temps en fonction du mode et du motif du déplacement. On constate que la valeur du temps est plus importante lorsque le déplacement est à but privé ou quand il s'effectue en transport en commun.

Les différents indices de valeur de temps évoqués précédemment sont donc définis en fonction du temps par :

$$\lambda_{TC} = \omega_m^{TC} (1,4 \times 10^{-3} \times (t_{Gj}^{TC} + \sum_{c=1}^q t_c^{Cor}) + 2,27$$

$$\lambda_{VP} = \omega_m^{VP} (2,3 \times 10^{-3} \times (t_{ij}^{VP}) + 1,67 \quad \text{d'après [Causse, 1999]}$$

avec  $\omega_m^{TC/VP}$  définit par le tableau 23.

### 2.2.2. Mise en place du modèle

Comme nous l'avons vu, le modèle d'optimum local est construit sur la mise en concurrence des deux modes de transport principaux que sont le transport en commun et la voiture personnelle. On calcule ainsi un indice affecté à chaque commune pour toutes les gares de rabattement de la région. De plus, la fonction d'utilité d'une origine – destination est pondérée par le poids qu'elle représente dans la fréquentation quotidienne du trajet concerné. On multiplie donc l'indice par  $U_{ij}$  afin de ne prendre en compte que les individus liés au trajet. Le tout est normalisé par le nombre d'individus se déplaçant par jour dans la commune origine afin de ne pas attribuer un indice trop élevé aux communes représentant un grand nombre d'usagers. Le modèle fonctionnant sur une logique additive des indices attribués à chaque trajet, un grand nombre d'émissions liées à une commune garantirait, sans cette normalisation, une valeur de  $I_{iG}$  élevée, même dans le cas d'une qualité de rabattement médiocre.

Il se traduit formellement par :

$$I_{iG} = \sum_{j=1}^n \frac{\lambda_{VP_1} \cdot t_{iG}^{Rab} + \lambda_{TC} (p^{Att} t^{Att} + t_{Gj}^{TC} + \sum_{c=1}^q p_c^{Cor} t_c^{Cor}) + C_{Gj}^{TC} + C_{iG}^{VP}}{\lambda_{VP_{ij}} \cdot t_{ij}^{VP} + C_{ij}^{VP}} \cdot \frac{U_{ij}}{U_i}$$

$$I_{iG} = \sum_{j=1}^n \frac{C_{ij}^{TC}}{C_{ij}^{VP}} \cdot \frac{U_{ij}}{U_i}$$

Avec  $U_{ij}$  = Nombre d'individus effectuant un déplacement depuis la commune  $i$  vers la commune  $j$

Et  $U_i$  = Nombre d'individus effectuant un déplacement ayant pour origine la commune  $i$

On affecte ensuite la commune à la gare présentant l'indice le plus élevé en respectant une contrainte de contiguïté de niveau 1. En clair, une commune ne peut être affectée à une gare que si elle est contiguë avec la commune sur le territoire de laquelle la gare est située ou avec des communes déjà agrégées à cette gare. On raisonne donc de manière itérative pour agréger dans un premier temps les communes pour lesquelles l'indice est maximum et présentant un degré de contiguïté avec la gare concernée, puis de proche en proche en redéfinissant à chaque étape un nouvel agrégat autour de chaque gare. La contrainte de contiguïté nous pousse naturellement à isoler certaines communes de leur agrégat. Nous avons alors mis en place un protocole de réaffectation de ces communes vers la gare située en deuxième, puis en troisième position pour l'indice  $I_{iG}$  avec une valeur seuil minimum d'exclusion de la réaffectation.

Le résultat final présenté par la figure 40 laisse donc apparaître encore des communes non affectées.



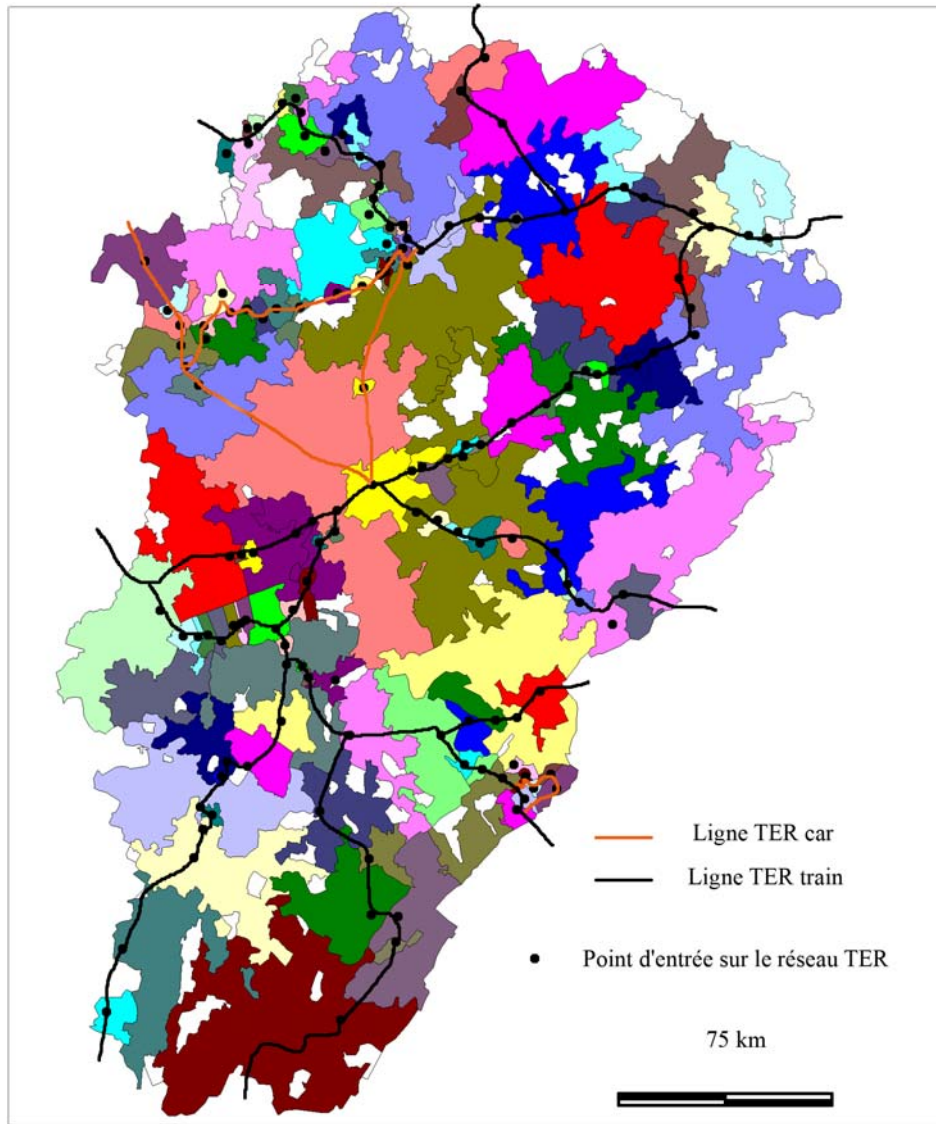


Figure 40 : Zonage établi sur le principe de l'optimum local

Avec le système de multi rabattement mis en place, peu de communes ne sont pas affectées à une gare. La majorité de ces communes exclues se situe aux marges de notre territoire d'étude, cet isolement peut s'expliquer par un effet de frontière. Ces communes trouvent en effet souvent un rabattement intéressant sur les territoires frontaliers de la Franche-Comté.

Nous voyons également se dégager des aires de rabattement importantes aux entrées des grandes villes. Ces aires, recadrées dans le contexte de la maximisation de la fonction de coût généralisé, traduisent spatialement les difficultés de circulation des entrées des agglomérations. Ainsi, un individu aura intérêt à garer sa voiture à la gare la plus proche de l'agglomération constituant sa destination et effectuer un report modal sur un transport en

commun afin d'éviter la circulation d'entrée de ville. Ce résultat, dans le contexte actuel de mise en place de parkings relais aux entrées des agglomérations, paraît tout à fait intéressant pour la localisation d'espaces privilégiés de report modal.

Toutefois, l'ensemble des résultats fournis par ce modèle ne paraît pas satisfaisant. En effet, les aires d'attraction des points d'arrêt TER aux marges des agglomérations ne paraissent pas crédibles en terme de zones d'émissions ou d'attractions de déplacements. De plus, elles ne sont pas cohérentes d'un point de vue purement pratique. Imaginer un report modal sur ces points d'arrêt nous pousse à considérer les possibilités de stationnement à chacun de ces points et nous pousse à relativiser les capacités d'accueil de certains arrêts.

Il manque donc certains critères d'évaluation à ce modèle afin de le rendre opérationnel. Toutefois, il n'est pas possible d'intégrer ces facteurs aux vues du processus d'affectation des communes. De plus, le principe de fonctionnement de l'optimum local, qui nous permet d'affecter une à une les communes aux gares présentant pour chacune un rabattement optimal, ne garantit pas obligatoirement la solution optimale pour l'ensemble de la Région. En effet, le déplacement d'une ou plusieurs communes d'un agrégat à un autre peut localement correspondre à une solution moins adéquate mais fournir pour l'ensemble du territoire une réponse plus satisfaisante par rapport aux attentes et aux buts de la modélisation.

En conséquence, nous avons fait le choix d'appliquer une seconde méthode, nous permettant d'inclure d'autres facteurs aux protocoles de définition de zones et de dépasser la simple échelle de raisonnement locale.

### **2.3. Le modèle d'optimum régional**

La seconde méthode employée pour l'optimisation du zonage de l'espace pose comme différence principale avec la méthode précédente le niveau d'échelle de la réflexion portée sur l'évaluation du découpage. Ici, l'évaluation ne porte pas sur chaque commune mais sur la Région dans son ensemble. Nous avons donc fait appel à une méthode nous permettant de modifier à volonté un découpage par l'introduction d'un aléa et d'évaluer les différentes solutions ainsi proposées.

### *2.3.1. Les algorithmes génétiques appliqués à la partition spatiale*

Nous avons fait appel à une seconde méthode de partition de l'espace, nous permettant à la fois de modifier le niveau d'échelle de notre raisonnement, de faire appel à de nouvelles variables explicatives et d'introduire un aléa dans la construction des zones. Dans ce but, nous avons opté pour l'utilisation des algorithmes génétiques.

#### Définition et principes de fonctionnement

« Un algorithme génétique est un algorithme stochastique itératif qui opère à partir d'une population initiale et qui est bâti à l'aide de trois opérateurs : croisement, mutation, sélection. » [Goldberg, 1994]. Cette définition résume parfaitement le principe et la méthode.

Le principe est donc basé sur une construction stochastique, les solutions fournies résultant du hasard, et sur la recherche de l'optimisation de ces solutions. La méthode employée pour cette optimisation, quant à elle, consiste en un dispositif appelé « méthode d'essai et erreur » [Birchenhall, 1995]. Il s'agit de tester un certain nombre de solutions potentielles jusqu'à l'obtention d'une solution adéquate. Cette méthode pourrait tout à fait se traduire par le principe Shaddock présenté à la figure 41.

Les algorithmes génétiques s'inspirent de la théorie de l'évolution de Darwin ; ils consistent à faire évoluer une population virtuelle vers un individu moyen. Cette méthode se présente donc comme une procédure qui s'inspire des mécanismes de sélection naturelle et des phénomènes génétiques. L'algorithme simule donc un processus d'évolution dans un environnement spécifique.

Nous utiliserons donc ici un vocabulaire propre à la méthodologie employée et se référant à celui de la génétique. Nous parlerons ainsi d'individus, inclus dans une population. L'individu est défini par ses gènes qui contiennent le caractère héréditaire de l'individu. Ces individus sont soumis à différents processus de sélection, de croisement ou de mutation que nous définirons plus en détail par la suite.

Ici, un individu correspond à un zonage de la Région de Franche-Comté, dont les gènes sont définis par la combinaison des différentes zones de rabattement. La population est donc composée d'un ensemble de zonages différents les uns des autres.

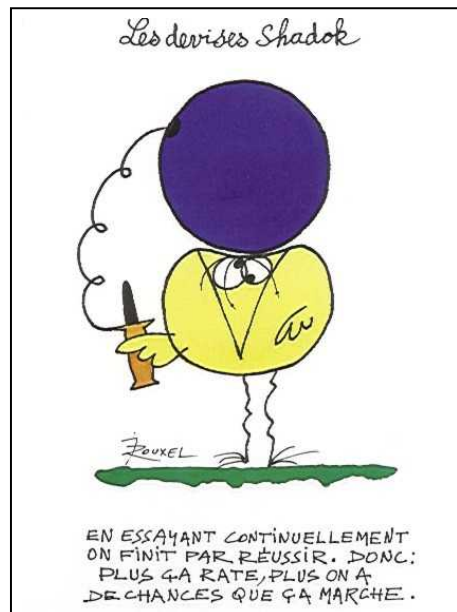


Figure 41 : Devise Shadok n°2  
Source : Jacques Rouxel

Darwin montre que l'apparition d'espèces distinctes est liée à la sélection naturelle de certaines variations individuelles [Darwin, 1876]. Cette sélection est basée sur la lutte pour la survie et est fonction de l'adaptation au milieu. Les individus les mieux adaptés tendent donc à survivre et à se reproduire de manière plus importante.

Les variations individuelles constituent donc le point fondamental du fonctionnement de l'évolution. Elles interviennent par hasard et permettent d'expliquer le phénomène d'évolution sans avoir recours à une modification directe de l'individu ni à la création d'une population nouvelle. L'espèce évolue donc mais ne tend pas pour autant vers un but prédéterminé.

Il existe différents types d'algorithmes génétiques, basés sur les mêmes principes, mais différant par leurs méthodes d'application. Nous utiliserons ici les algorithmes dits évolutionnaires basés sur les opérateurs génétiques déjà cités de sélection, de croisement et de mutation. La figure 42 présente l'organigramme de l'algorithme évolutionnaire.

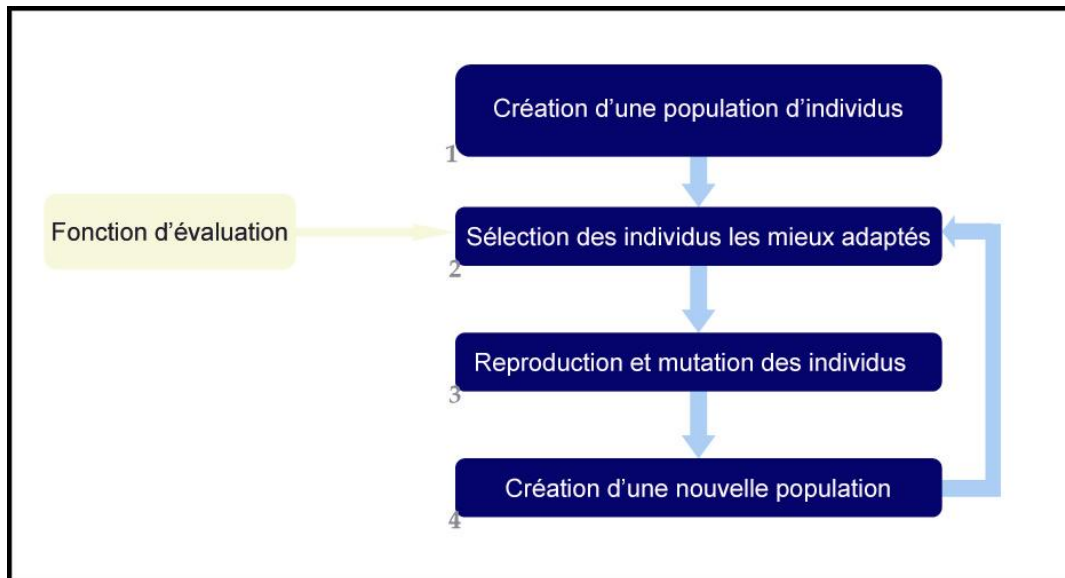


Figure 42 : Le principe itératif d'un algorithme évolutionnaire

Il s'agit donc de simuler l'évolution d'une population d'individus de départ fondée sur l'aléa, à laquelle on applique les principes de croisement et de reproduction de la génétique. Ce principe de reproduction suit le même principe que l'évolution des espèces édictée par Darwin à savoir la sélection des individus les plus adaptés au milieu. Cette sélection s'opère donc à partir d'une fonction d'évaluation qui pousse la population à s'améliorer à chaque génération. Cette méthode ne nécessite pas la compréhension du problème en lui-même, ici le profil génétique des individus, mais requiert la connaissance de l'optima désiré.

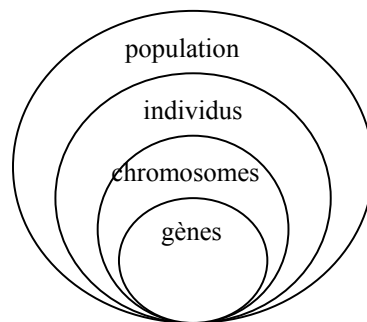
Les éleveurs de bétail par exemple ont ainsi procédé pendant plusieurs siècles en sélectionnant selon leur désir les bêtes les plus adaptées à une certaine demande. Ils ont réussi à modifier le patrimoine génétique de certaines espèces sans connaissance en génétique ou biologie moléculaire.

Il ne s'agit donc pas de fournir une solution analytique exacte ou une approximation de qualité, mais d'identifier les solutions correspondant au mieux à des critères pré-établis.

Il est fondamental, dans la mise en place d'un algorithme évolutionnaire, de respecter certaines étapes, qui garantissent le bon fonctionnement et le respect des principes de base que nous venons de voir.

### Le codage des variables

La première étape consiste en la définition de codage du problème. Il s'agit, en clair, de définir ce qui compose les gènes des individus à évaluer. Ces gènes composent des chromosomes qui assemblés constituent le génotype de l'individu (figure 43). Il faut donc identifier avec précision l'élément d'évaluation qui permet de créer des individus dissemblables.



*Figure 43 : Les quatre niveaux d'organisation d'un algorithme évolutionnaire  
[Magnin 2006]*

Dans le cas présent, les individus, identifiés comme nous l'avons vu par les zonages dans leur ensemble de la Région de Franche-Comté, sont caractérisés par des gènes composés des communes situées sur le territoire régional. Elles se combinent en agrégats, sous forme d'aires de rabattement autour des points d'arrêt du réseau TER, qui constituent les chromosomes. La population est définie, quant à elle, par un nombre  $N$  d'individus, pouvant varier selon la simulation désirée.

### La genèse, ou création d'une population initiale

La seconde étape consiste à générer une population initiale. Car si l'algorithme permet l'amélioration de la vie, il ne la crée pas. Il paraît peu important que cette population initiale soit performante ou non, l'algorithme se chargeant de faire converger la population vers les caractéristiques souhaitées. La population initiale peut donc être définie aléatoirement, en respectant les contraintes liées aux chromosomes.

Dans notre cas, les aires de rabattement ont été définies aléatoirement, en respectant une contrainte de contiguïté de niveau un.

## La sélection des individus

Cette étape consiste à identifier les individus les plus adaptés à l'environnement, ou, comme ici, aux attentes du modélisateur. « L'évaluation, définie par les contraintes liées aux critères de convergence, a pour fonction d'attribuer une note à chaque individu » [Bolot, 1999].

Le principe de sélection repose donc sur la qualité de chaque individu. Elle est calculée à partir de la fonction d'évaluation, qui fournit un score à chaque membre de la population. Ce score permet par la suite d'opérer une sélection sous contrainte de probabilité des individus considérés comme bon. Ce principe de sélection est appelé « roue de la fortune » et consiste donc à associer à chaque individu une probabilité de tirage proportionnelle à sa part de score dans la population. En clair, les individus bénéficiant d'un score élevé ont plus de chance d'être sélectionnés pour transmettre tout ou partie de leur génotype à la génération suivante. Ces individus constituent la génération dite intermédiaire.

Il existe d'autres modes de sélection qui fournissent, selon le contexte, des convergences plus ou moins rapides. Dans le cas de population peu importante, on fait parfois appel à la sélection dite par tournoi. Deux individus sont choisis au hasard et combattent pour accéder à la part de la population pouvant subir les opérateurs de croisement ou de mutation. On conserve ainsi un certain aléa dans la sélection. On fait également appel à l'ordonnancement comme méthode de sélection. Cette méthode est employée dans les systèmes de forte convergence, c'est-à-dire lorsque l'évaluation permet de dégager de grandes différences entre les individus. On attribue à chaque individu un classement par ordre d'évaluation. On sélectionne par la suite une population intermédiaire en utilisant les probabilités indexées sur les rangs des individus. Dans le cas d'individus peu différenciés par la fonction d'évaluation, cette méthode présente le désavantage principal d'exagérer le rôle des meilleurs individus. Le second individu aura par exemple une probabilité d'être sélectionné beaucoup plus faible que le premier, bien qu'il puisse représenter une part de la réponse finale non négligeable.

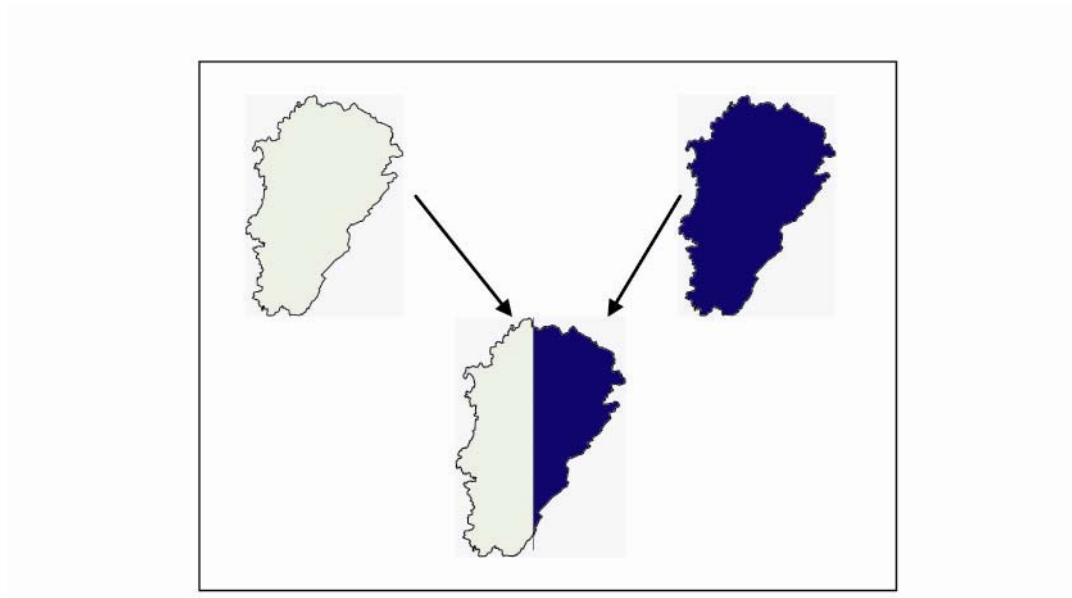
On trouve parfois d'autres méthodes, souvent hybrides, telles que la roue de la fortune modifiée [Farley et Jones, 1994], méthode plus déterministe sur laquelle on réaffecte un aléa de manière artificielle.

### Croisement et mutation des individus

Une fois la génération intermédiaire définie, on utilise différents opérateurs afin de transmettre à la génération suivante le patrimoine génétique sélectionné, tout en conservant une part d'aléa propre aux algorithmes génétiques.

Le premier de ces opérateurs est le croisement, relativement similaire à la reproduction connue dans le monde biologique. Il a pour but d'enrichir la diversité de la population en hybridant les chromosomes des individus, voire dans certains cas directement les gènes. Classiquement, cette phase de l'algorithme se conçoit avec deux parents qui génèrent deux enfants.

Le croisement consiste à scinder en deux moitiés le génotype coupé à l'identique de deux individus différents et d'affecter à l'un une des moitiés de l'autre et réciproquement (figure 44). Habituellement, la structure des chromosomes est la même d'un individu à l'autre ; ce sont les gènes qui les constituent qui diffèrent. Dans notre cas, les chromosomes, c'est-à-dire les aires de rabattement, n'ont pas obligatoirement les mêmes limites d'un zonage à un autre. On ne peut donc pas appliquer ici le croisement classique opéré à partir du découpage de chromosomes. Pratiquement, il n'y a que peu de chance que les frontières des zones d'un individu correspondent parfaitement aux frontières d'un autre individu. Il est alors nécessaire de réaffecter les gènes séparés, ici les communes isolées, à un chromosome existant.



*Figure 44 : Le principe du croisement*



Nous avons pris le parti, dans le processus de réaffectation des communes isolées par l'étape de croisement, de réagrèger ces communes à une des aires de rabattement qui lui est contiguë sans chercher à optimiser cette affectation.

Vient ensuite le processus de mutation. Cette étape a pour fonction de fournir à l'algorithme l'aléa nécessaire à une exploration efficace des solutions. Cet opérateur nous garantit que l'algorithme évolutionnaire sera susceptible d'atteindre toutes les solutions potentielles sans pour autant les parcourir toutes dans le processus de résolution. [Goldberg, 1994]. Dans la pratique, la mutation permet d'explorer des possibilités génétiques que le hasard peut avoir ignorées dans la population initiale.

Dans le cas présent, la mutation s'opère directement sur les gènes. Des communes changent donc aléatoirement d'agrégat, tout en respectant la contrainte de contiguïté. La probabilité de mutation est dépendante de deux indices. Un premier fixant la probabilité qu'a ou non un individu de subir une mutation ; un second établissant la probabilité qu'a ou non une commune de muter.

Il existe plusieurs processus d'amélioration ou d'accélération des algorithmes évolutionnaires. Le but ici n'étant pas l'optimisation du temps de calcul, nous nous sommes restreint à une utilisation « classique » de cette méthode. Le lecteur pourra consulter des ouvrages plus complets à ce sujet, détaillant les méthodes de scaling qui amplifient artificiellement les écarts entre les individus, obtenus par la fonction d'évaluation<sup>19</sup>.

### *2.3.2. Mise en place de la modélisation*

La mise en place de l'algorithme, et donc principalement la conception de la fonction d'évaluation, pose plusieurs interrogations.

La première concerne les communes non affectées. Le modèle précédent a mis en évidence certaines communes pour lesquelles l'affectation obtenait une note trop basse pour être affectée à une aire de rabattement. La question se pose donc de savoir si le fait d'isoler certaines communes doit à terme améliorer l'évaluation des individus ou constituer un handicap quel que soit l'espace concerné.

---

<sup>19</sup> L. Goffe, G. D. Ferrier, J. Roger, 1994, *Global optimization of statistical functions with simulated annealing*, Journal of Econometrics, Vol 60, pp 65-99

Il faut revenir aux buts initiaux du modèle afin de répondre le plus justement possible à cette question. Le zonage, dans le cadre du modèle à quatre étapes et plus particulièrement concernant l'étape de génération de trafic, a pour but de dimensionner au mieux les aires de rabattement autour des points d'arrêt du service TER. Le but final est de pouvoir quantifier au mieux et de la manière la plus pertinente possible les usagers concernés, ou potentiellement concernés par un service particulier. Dans ce contexte, ne pas agréger certaines communes à une aire de rabattement revient à ignorer une partie de l'information afin de permettre une plus grande cohérence de l'information restante. Il paraît donc important de pouvoir isoler certaines communes, tout en s'assurant de ne pas privilégier ce processus vis-à-vis du rabattement. De plus, les données de l'enquête TER révèlent un certain nombre de communes de l'espace franc-comtois ne déclarant aucun migrant usager d'un quelconque service TER. Cette constatation renforce donc l'importance de configurer le modèle pour une acceptation de communes non rabattues.

En clair, si le fait d'isoler une commune peut fournir une solution plus satisfaisante pour le zonage dans son entier, il est nécessaire de ne pas agréger cette commune. Toutefois, il faut garder une certaine vigilance afin de ne pas écarter le modèle de son but initial qui reste d'agréger les communes de notre espace autour de points de rabattement.

Il s'agit donc de rejeter la possibilité de non agrégation d'une commune tant que l'évaluation locale reste suffisamment satisfaisante ; puis d'infléchir la fonction d'évaluation proportionnellement à la baisse de l'évaluation locale. Nous avons donc adopté ici une logique de commutation (« *switching* ») basée sur une valeur seuil, mais prenant en compte le caractère flou de la qualité du rabattement (ou du non rabattement).

Cette étape a deux fonctions. Elle permet dans un premier temps l'évaluation des communes non rabattues, et opère dans un second une forme de *scaling* qui vise à ramener les valeurs de l'indice dans un intervalle compris entre 0 et 1. Cette « remise à l'échelle » de l'indice  $I_{iG}^{géné}$  a pour but de pouvoir lui ajouter par la suite d'autres indices, d'évaluation globale du zonage ou de potentialité de parking, tout en conservant la même échelle d'évaluation. C'est ainsi qu'ont été calculées les valeurs de la fonction  $I_{iG}^{géné}$  lorsque les communes sont rabattues. Nous nous sommes basés sur des valeurs repères du premier modèle (valeur minimale, valeur maximale, moyenne) afin de traduire en une fonction retranscrivant au mieux l'évaluation attendue en sortie du modèle. Nous avons donc paramétré une fonction retranscrivant au mieux le comportement de l'indice  $I_{iG}^{géné}$  attendu, à savoir augmentation très faible de la valeur de l'indice jusqu'à la valeur seuil du non rabattement ; puis augmentation rapide pour

favoriser son rabatement ; enfin relative stagnation une fois la valeur de l'indice considérée comme satisfaisante atteinte. Nous nous sommes basé sur une fonction logistique :

$$I_{iG}^{géné} = \frac{\beta}{1 + \gamma \cdot e^{-\alpha}} \quad \text{avec}$$

$$\alpha = -\ln \frac{\sum_{t=2}^n \frac{1}{y_t y_{t-1}} - \frac{1}{n-1} \sum_{t=2}^n \frac{1}{y_{t-1}} \sum_{t=2}^n \frac{1}{y_t}}{\sum_{t=2}^n \left(\frac{1}{y_{t-1}}\right)^2 - \frac{1}{n-1} \left(\sum_{t=2}^n \frac{1}{y_{t-1}}\right)^2}$$

$$\gamma = \frac{(n-1)(1 - e^{-\alpha})}{\sum_{t=2}^n \frac{1}{y_t} - e^{-\alpha} \sum_{t=2}^n \frac{1}{y_{t-1}}}$$

$$\beta = e^{(n+1)\alpha/2 + (\sum_{t=1}^n \ln(\gamma / y_{t-1})) / n}$$

Concernant la fonction mise en place pour les communes non rabattues, nous nous sommes basé sur la valeur seuil du premier modèle correspondant à la valeur de  $I_{iG}$  pour le nombre  $n$  de communes non rabattues.

Formellement :

$$I_{iG}^{géné} \begin{cases} \frac{-0,92}{1 + 1392 \cdot e^{-26,52 \cdot I_{iG}}} & \text{si la commune est rabattue} \\ -3,8 I_{iG} + 0,5 & \text{si } I_{iG} \leq 0,12 \text{ et si la commune n'est pas rabattue} \\ 0 & \text{si } I_{iG} > 0,12 \text{ et si la commune n'est pas rabattue} \end{cases}$$

Nous avons par la suite, construit deux indices complémentaires basés sur les capacités de parking de chaque point d'arrêt et sur l'équilibre général du zonage régional.

Concernant la contrainte liée au parking, l'évaluation doit être bonne tant que la capacité d'accueil n'est pas saturée, puis doit décroître en fixant une valeur maximale du nombre d'individus acceptés à chaque point d'arrêt.

Formellement :

$$f_{park(v)} \left| \begin{array}{l} 1 \text{ si } \sum U_i \leq v \\ 1 - \frac{U_i - v_G}{\delta v_G} \text{ si } \sum U_i > v_G \text{ et } \sum U_i < \delta v_G \\ 0 \text{ si } \sum U_i > \delta v_G \end{array} \right.$$

Avec  $v_G$  = nombre de places de parking de la gare G

et  $\delta$  = coefficient de tolérance d'accueil du nombre de voitures supplémentaires lorsque celui-ci est supérieur au nombre de places de parking

Pour ce qui est de l'indice d'équilibre général du zonage, il est basé sur la relative bonne répartition de la population et du nombre de communes dans chaque agrégat. Cette répartition homogène ne se fait pas indifféremment sur l'ensemble du territoire régional mais par type de commune. Nous reprenons pour cela la typologie de communes mise en place lors de la deuxième partie. Afin de conserver la même échelle de valeur que pour les indices  $I_{iG}^{géné}$  et  $f_{park(v)}$ , nous avons adopté une mesure de dispersion comprise entre 0 et 1. Nous employons pour cela l'écart-type relatif définit par :

$$\sigma_{x'} = \frac{\sigma_x}{\sum_{i=1}^n x_i'}$$

[Groupe Chadule, 2001]

Où  $x'$  est définit par :

$$x_i' = \frac{x_i}{\sum_{i=1}^n x_i}$$

Cet indice « vaut 0 si toutes les valeurs sont constantes, et  $((n-1)/n^2)^{1/2}$  si la dispersion est minimale, c'est-à-dire si tous les  $x_i$  sauf un valent 0 » [Groupe Chadule, 2001]. Donc la relativisation de cet indice entre 0 et 1 correspond à :

$$\sigma_{x'} = \frac{\sigma_x / \sum_{i=1}^n x_i'}{((n-1)/n^2)^{1/2}}$$

Formellement, concernant le zonage :

$$E_T = \frac{\sigma_{(C)}^t / \sum C'}{2 \max \left| \frac{\sigma_{(C)}^t}{\sum C'} \right|} + \frac{\sigma_{(P)}^t / \sum P'}{2 \max \left| \frac{\sigma_{(P)}^t}{\sum P'} \right|}$$

Où C représente le nombre de communes et P la population par aire de rabattement.  $\sigma_{(C)}^t$  représente l'écart type lié à ces valeurs pour chaque type  $t$  de commune.

L'étape suivante consiste à optimiser les différents paramètres liés au déroulement de l'algorithme évolutionnaire. Nous avons testé pour cela la rapidité de convergence de l'algorithme en faisant varier la taille de la population, la probabilité de faire muter un individu, la probabilité de mutation d'un gène pour les individus enclins à muter, la probabilité de conserver des individus sans croisement ni mutation, et enfin la probabilité de sélectionner ou non les meilleurs individus en fonction leur part de score.

Le paramétrage de ces différents facteurs n'est pas crucial dans le processus même d'optimisation, l'algorithme améliorant par principe les générations au fur et à mesure de son déroulement. Il recouvre cependant une certaine importance dans la rapidité de cette optimisation. La figure 45 présente la convergence sur les 1000 premières itérations de l'algorithme pour les meilleurs paramètres.

Les tests de paramétrage concernant la taille de la population montrent que l'algorithme fournit l'optimisation la plus rapide pour une population d'environ 600 individus par génération.

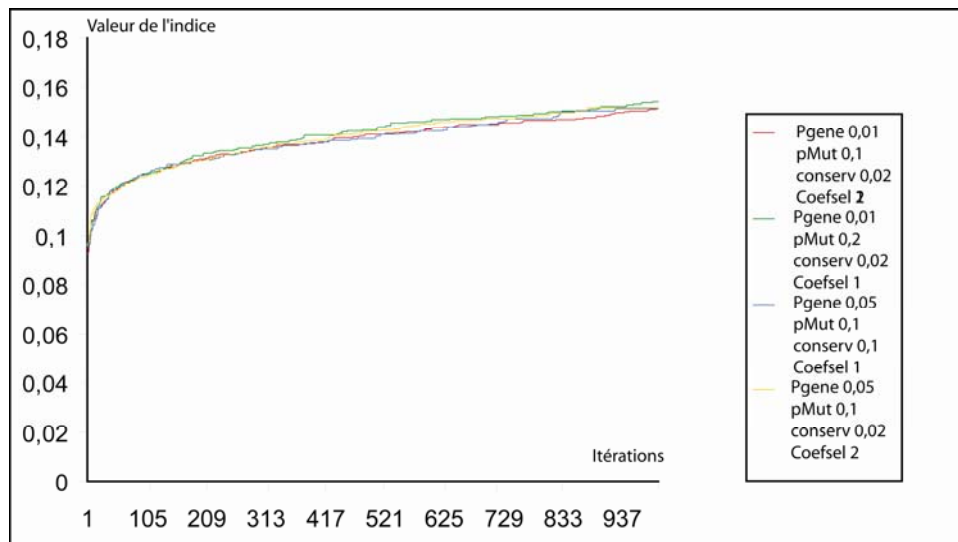


Figure 45 : Test des paramètres, la convergence de l'algorithme génétique

Nous avons par la suite « testé » l'algorithme évolutionnaire en combinant les différents indices de rabattement, de parking et d'évaluation globale mis en place.

La figure 46 présente le zonage construit à partir de l'indice  $I_{iG}^{gén\acute{e}}$  seul. On retrouve très rationnellement la même logique organisationnelle que pour le modèle d'optimum local, à quelques exceptions près de communes non rabattues et quelques dissemblances au niveau des agrégats. Cependant, les entrées d'agglomérations constituent toujours les pôles principaux et on retrouve dans l'ensemble l'effet de frontière explicité précédemment.

Le premier modèle joue donc son rôle de contrôle et nous permet de ne pas nous fier aveuglément à une modélisation dont le déroulement, plus obscur, peut conduire à des solutions erronées.

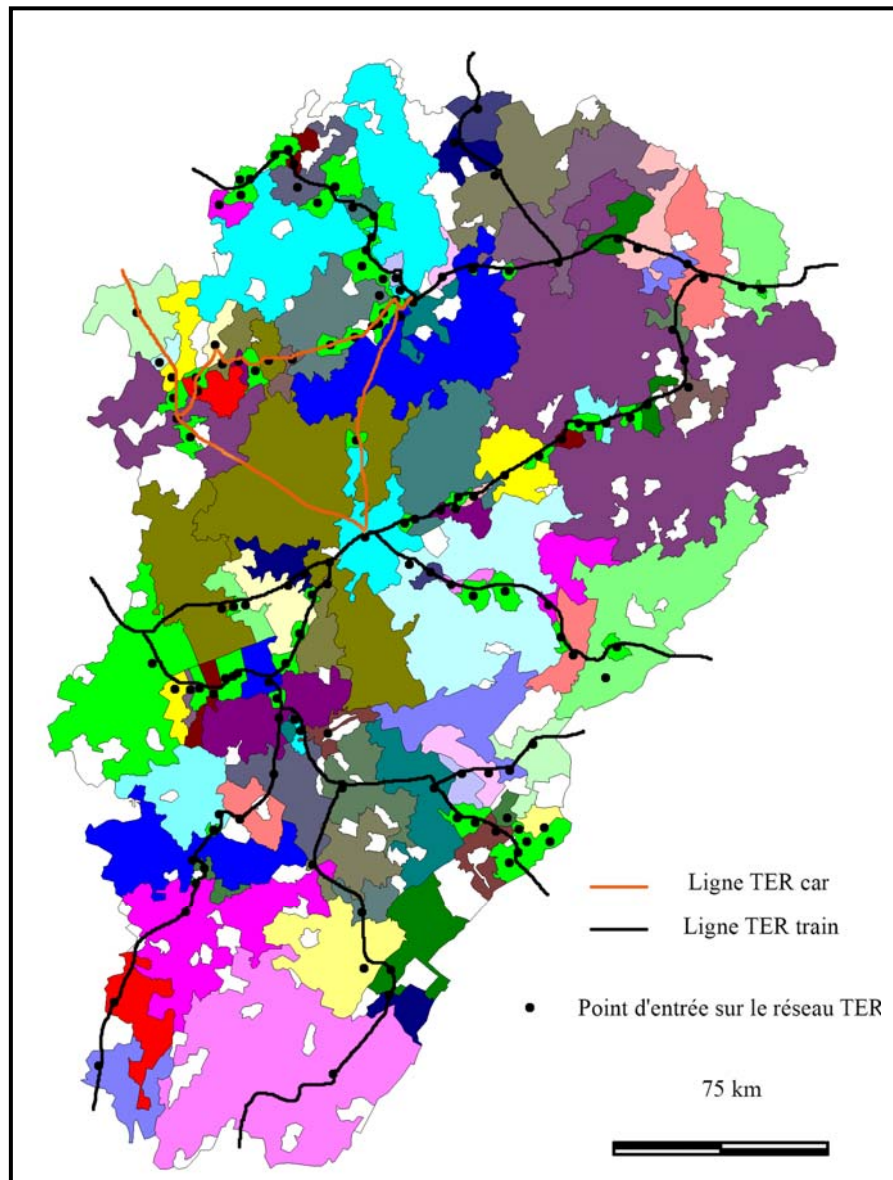


Figure 46 : Aires de rabattement autour des points TER pour l'indice de rabattement seul

Lorsque l'on ajoute l'évaluation globale pour l'élaboration du zonage, on obtient naturellement une meilleure répartition de l'espace entre les agrégats. On constate, en premier lieu, la disparition des zones de grande taille aux entrées des agglomérations au profit d'un autre type de communes. Rappelons ici que la typologie des communes a été élaborée à partir des comportements de mobilité des individus. Les points d'entrée sur le réseau présentant les aires de rabattement les plus vastes correspondent globalement aux pôles secondaires et villes de moyennes importance. Ainsi, Dole à l'ouest, Saint-Claude au sud, Pontarlier à l'est, et dans une moindre mesure Baume-les-Dames et L'Isle-sur-le-Doubs voient leur aire composer une part importante du zonage régional.

Ces points d'entrée sur le réseau TER, déjà pourvus d'aires de rabattement conséquentes dans le zonage précédent, ne subissent que peu l'effet de l'équilibrage général du zonage et correspondent vraisemblablement à des points de rabattement privilégiés du réseau franc-comtois.

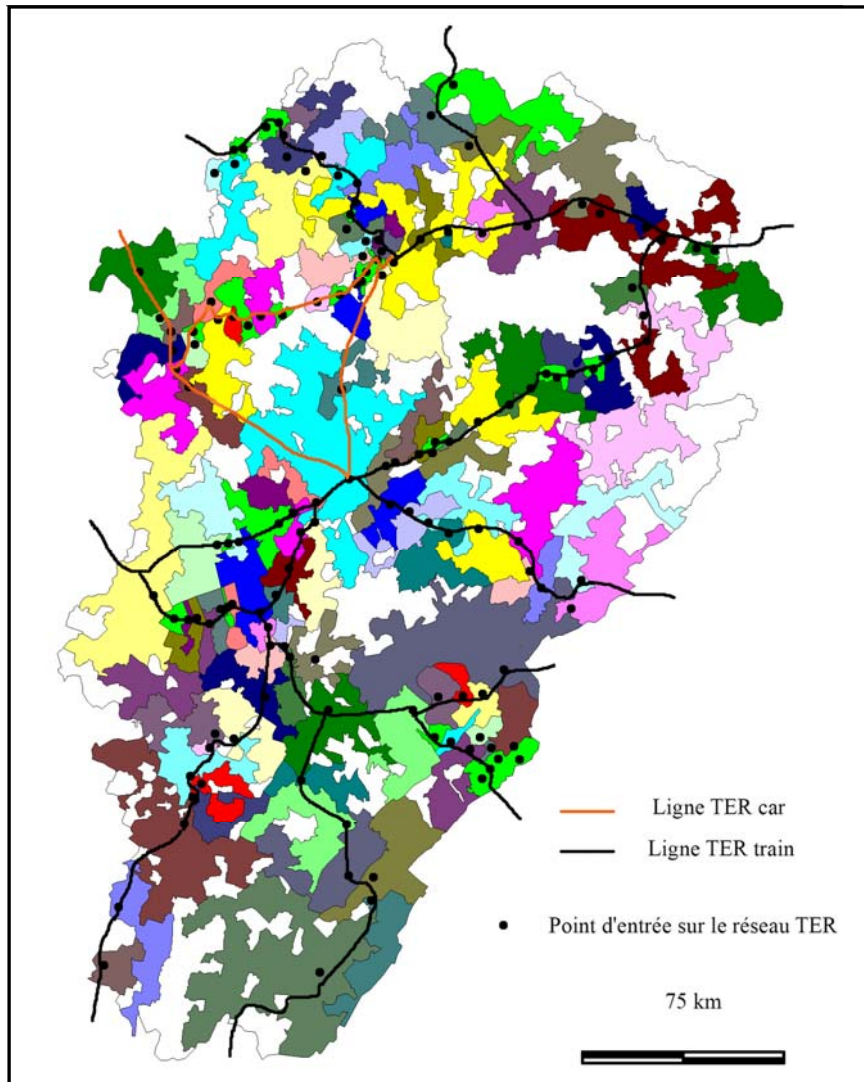


Figure 47 : Aires de rabattement autour des points TER pour l'indice de rabattement et l'indice global

L'aire d'influence de Besançon s'étend vers le nord, là où la desserte TER est moins présente. Le sud de la capitale régionale, quant à lui, présente un espace non rabattu par l'algorithme. L'isolation de ces communes, en terme d'organisation spatiale, signifie globalement pour les individus qu'il est préférable, aux vues de la fonction d'évaluation, de faire usage d'un véhicule personnel plutôt que de transports en commun, pour la majorité de la



population concernée. Le modèle met donc ici en évidence une portion du territoire régional sur laquelle le service TER n'offre certainement pas une alternative suffisante aux solutions de transport individuelles.

On retrouve ce même phénomène sur un espace relativement vaste situé entre la ligne Besançon – Strasbourg et la ligne Paris – (Vesoul – Belfort) – Bâle. L'analyse de cet espace nous permet d'affiner notre propos. L'isolation de certaines communes par le modèle ne signifie par forcément l'absence de point de rabattement proche de ces communes, mais l'inadaptation des services proposés dans ces points d'entrée par rapport aux attentes de mobilité des individus résidant dans ces communes.

On retrouve enfin l'effet de frontière, principalement dans le sud jura, dont l'organisation est guidée par des points de rabattement extérieurs à notre territoire, effet constaté également le long de la frontière suisse.

Ajouter une contrainte de parking à l'évaluation revient à traiter le problème de manière beaucoup plus appliquée. On ne travaille ainsi que sur des rabattements réalisables en l'état. Les étapes précédentes de la modélisation trouvent toutefois leur utilité dans la compréhension du potentiel d'utilisation de certains points d'entrée sur le réseau TER pouvant paraître pour l'instant sous exploités.

Comme lors de l'application précédente, les entrées d'agglomération conservent des aires de rabattement tout à fait modestes. On constate également que les pôles secondaires évoqués précédemment gardent des aires de rabattement étendues. Il ne s'agit donc pas d'un simple effet d'équilibre dicté par une appartenance au même type de commune. On constate à cette occasion que la gare de Lure, située sur la ligne Paris – Bâle, étend son aire d'influence sur une grande partie des communes non rabattues précédemment. C'est certainement l'appartenance au type de commune le plus présent sur notre territoire qui peut expliquer la restriction de son aire de rabattement lors de l'application précédente. On peut sans doute formuler les mêmes hypothèses concernant l'extension de l'aire d'influence de Vesoul.

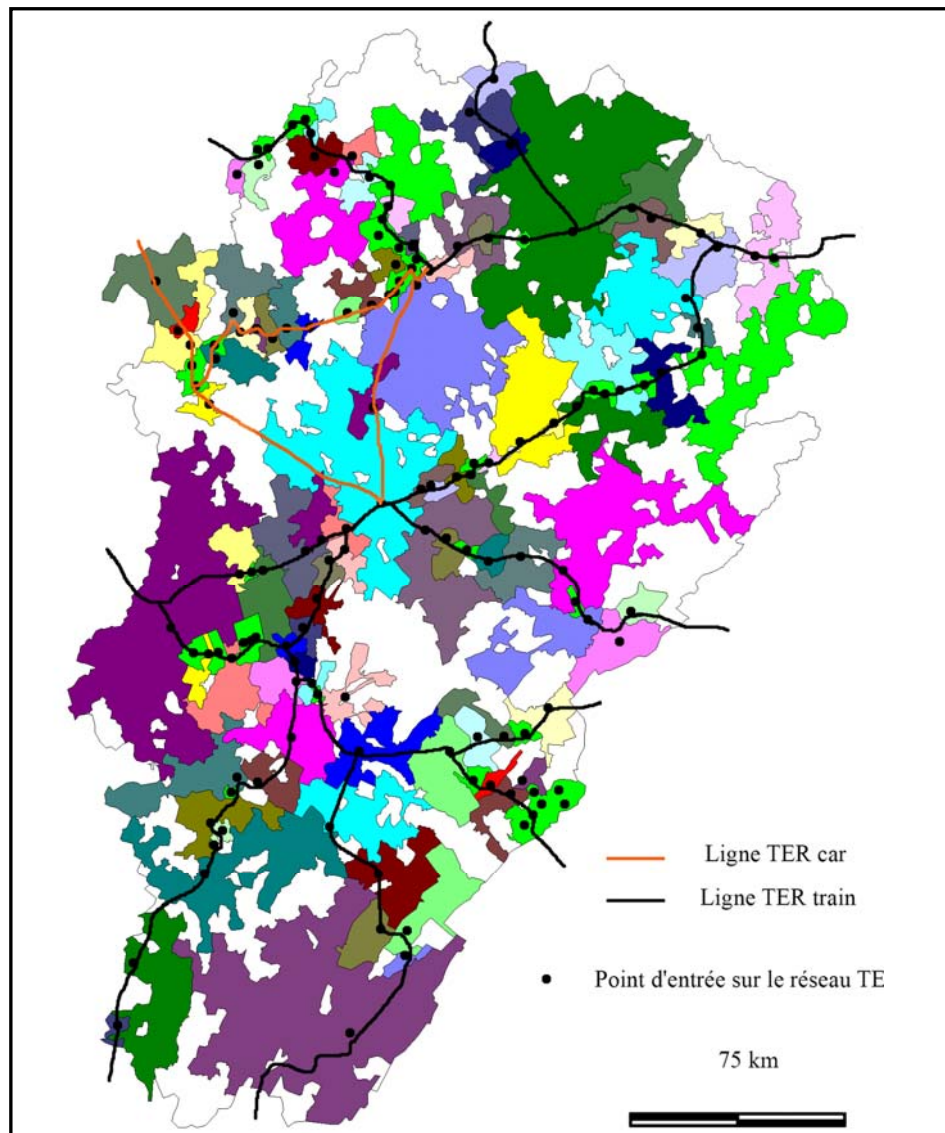


Figure 48 : Aires de rabattement autour des points TER pour l'indice de rabattement et l'indice de parking

L'espace non rabattu au sud de Besançon reste quant à lui isolé et ne trouve pas plus dans cette application de solution d'optimisation de rabattement. Les solutions en terme de service TER apparaissent donc ici comme inadaptées à la demande.

Enfin, afin de croiser l'approche très pragmatique que nous venons de voir et l'approche plus théorique précédente, nous avons configuré l'algorithme évolutionnaire pour une prise en compte de l'ensemble des indices dont nous disposons. La figure 49 présente le résultat de cette simulation.

On retrouve globalement les mêmes phénomènes déjà observés au cours des deux applications précédentes. Une certaine nuance est néanmoins apportée concernant les aires d'influences des pôles secondaires ainsi que sur les communes non rabattees. Ainsi, on voit réapparaître une partie de l'espace non rabattu entre les deux lignes ferroviaire du nord Franche-Comté.

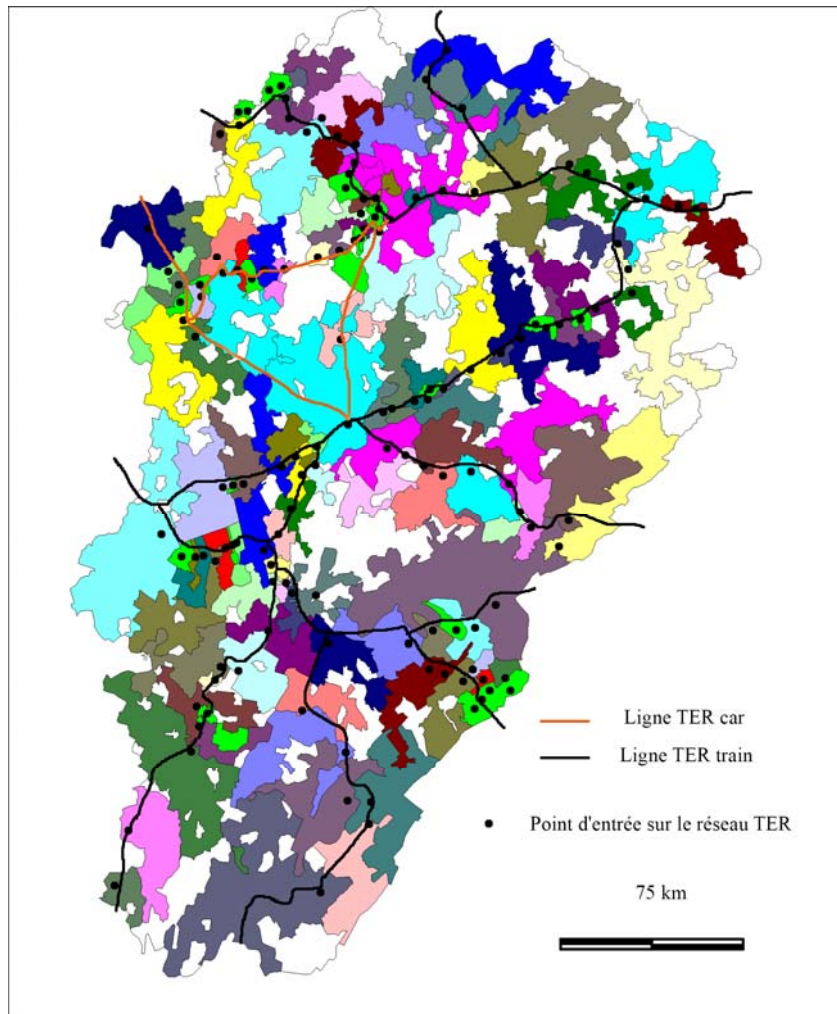


Figure 49 : Aires de rabattement autour des points TER pour l'indice de rabattement, l'indice de parking et l'indice global

On peut imaginer augmenter encore le nombre d'indices ou facteurs pouvant infléchir plus ou moins la composition et la forme des aires de rabattement. Les critères de confort, au-delà des simples possibilités de stationnement, mériteraient sans doute un approfondissement pour une meilleure prise en compte dans la fonction de coût généralisé employée ici. De même la perception de la sécurité joue un rôle dans le processus décisionnel du mode de transport et donc du rabattement. Mais le manque de données sur ces critères

explique le caractère incomplet dans cette prise en considération, les données détenues par la SNCF dans ce domaine ne nous étant pas accessibles. De même, faute de données satisfaisantes dans ce domaine, nous n'avons pas pu prendre en compte la perception et la manière dont se combinent les différents critères de coût, de temps et de confort dans le processus de choix de l'individu. Il va de soi que la même importance n'est pas accordée à la rapidité, au coût ou à tout autre critère dans la décision d'un point de rabattement. Il s'agit là de pistes encore peu traitées en géographie des transports, autant pour l'importance de ces critères en terme de perception que concernant le processus de recombinaison mentale de ces facteurs lors d'une prise de décision finale.

Une analyse temporelle plus fine pourrait également être souhaitable. Il serait alors nécessaire de disposer de données nous permettant une telle analyse. L'enquête TER nous a fourni un tel niveau de détail ; mais nous ne disposons pas de données comparables concernant les autres modes de transport. Comme nous l'avons déjà vu, l'analyse par mode de transport, concernant la génération de trafic, ne peut pas s'appliquer. Ce serait faire un postulat de départ erroné en considérant que le mode constitue la variable explicative principale dans le processus du choix de déplacement.

On peut raisonnablement penser que le zonage même se trouverait modifié selon les différents horaires de la journée. Il nous fournirait alors sans doute une meilleure compréhension du territoire de la mobilité.

On peut s'attendre à rencontrer des espaces non rabattus, et donc pour lesquels les transports régionaux ne sont pas adaptés, beaucoup plus vastes en dehors des heures d'affluences.

L'analyse spécifique du dimanche ou des jours fériés pourrait également constituer un champ d'investigation tout à fait intéressant. Les études et simulations concernant la mobilité quotidienne, souvent coûteuses, poussent les modélisateurs à centraliser leurs efforts sur les périodes regroupant la plus forte part de la population. Ces périodes où l'affluence est moindre ont donc jusqu'ici été laissées de côté et nous offrent un champ presque vierge de toute analyse. Mais, comme précédemment, il est nécessaire, afin de pouvoir réaliser une véritable étude sur ces espaces temporels, de disposer de données détaillées issues d'enquêtes encore non réalisées à ce jour.

## 2.4. Agrégation des résultats obtenus à l'échelle communale pour chaque aire de rabattement

Dénombrer les émissions par agrégat

La logique de construction des aires de rabattement nous pousse à considérer les émissions et attractions de chaque zone. En effet, l'élaboration des agrégats étant basée sur le rabattement potentiel des migrants journaliers, le calcul des émissions et attractions, dans ce cadre, ne s'opère pas par une simple addition des émissions et attractions des communes composant l'agrégat. Nous avons conçu les aires de rabattement autour des points d'entrée du réseau TER afin d'identifier au mieux les individus concernés par ce point d'entrée. Or, comme nous l'avons vu, tous les migrants d'un agrégat ne sont pas concernés par le point d'arrêt de l'agrégat auquel ils appartiennent.

Il est alors nécessaire d'indicer les émissions de chaque commune par la valeur de l'indice  $I_{iG}^{géné}$  de la commune  $i$ . Formellement,  $I_{iG}^{géné}$  étant compris entre 0 et 1 (1 correspondant au rabattement de l'ensemble des individus de la commune  $i$  sur la gare  $G$ ), les émissions liées au point de rabattement  $G$  et attachées à l'agrégat  $A_G$  sont exprimées par :

$$E_G = \sum_{i=a_1}^{a_n} I_{iG}^{géné} \cdot E_i$$

Toutefois, ce mode de calcul ne prend en compte que les émissions liées au réseau TER, de manière potentielle ou réelle, à partir des données prises en compte au départ, c'est-à-dire de la population concernée par les communes de l'agrégat.

On peut imaginer utiliser l'indice  $I_{iG}^{géné}$  en dehors du contexte du zonage dans l'optique d'une simulation présentant le désavantage certain de temps de calculs beaucoup plus conséquent. Ainsi, on peut évaluer plus précisément les émissions liées à un point de rabattement particulier, sans considérer de zonage ni de découpage communal, mais en prenant comme point d'émission la gare concernée. On traite ainsi l'ensemble du territoire régional et non plus juste la zone définie lors de la partition spatiale.

Les émissions liées au point d'arrêt  $G$  sont alors exprimées par :

$$E_G = \sum_{i=1}^n I_{iG}^{géné} \cdot E_i$$

Cette manière de procéder, bien que plus longue en terme de temps de calcul, a l'avantage certain de n'ignorer aucun individu concerné par le point d'arrêt  $G$ .

La prise en compte du zonage laisse en effet à l'écart l'ensemble des individus potentiellement concernés par le point de rabattement  $G$  mais en dehors de l'agrégat  $A_G$ .

### 2.5. Dénombrer les attractions par agrégat

Concernant le dénombrement des attractions, le problème est un peu différent. Il est nécessaire alors de prendre en compte la distribution des déplacements c'est-à-dire considérer les trajets dans leur ensemble pour pouvoir calculer les attractions à partir des émissions prises en comptes dans l'étape précédente. On fait alors l'hypothèse que les individus pris en compte lors du calcul des émissions correspondent aux individus faisant le plus augmenter l'indice  $I_{iG}^{géné}$ . On définit alors un ensemble  $K$  correspondant à l'ensemble des  $V_G$  plus forts indices  $I_{ij}^p$  pour la gare  $G$ , où

$$V_G = I_{iG}^{géné} \cdot U_i \quad \text{et,}$$

$$I_{ij}^p = \frac{\lambda_{VP_1} \cdot t_{iG}^{Rab} + \lambda_{TC} (p^{Att} t^{Att} + t_{Gj}^{TC} + \sum_{c=1}^q p_c^{Cor} t_c^{Cor}) + c_{Gj}^{TC} + C_{iG}^{VP}}{\lambda_{VP_j} \cdot t_{ij}^{VP} + C_{ij}^{VP}}$$

Les attractions sont alors définis, par aire de rabattement par :

$$A_G' = \sum_{j=a_1}^{a_n} \sum_{i=1}^n \sum_{G=1}^m U_{ij} \quad \text{sur } K$$

## CONCLUSION

Nous avons finalement modélisé dans son ensemble l'étape de génération de trafic, de la définition d'un zonage pertinent à la quantification des émissions et attractions de déplacements pour chacune de ces zones.

Il est difficile d'évaluer la pertinence du zonage défini au cours de ce chapitre. Dans la définition de partition spatiale liée à des flux, on évalue la pertinence des aires obtenues à partir d'une comparaison des flux internes et du flux entrants et sortants de la zone [Josselin et al, 2000]. Dans notre cas, comparer sur une zone ces flux n'aurait pas réellement de sens.

La comparaison des émissions et attractions théoriques par point d'arrêt avec les montées-descentes constatées peut fournir un élément de contrôle partiel de la qualité du zonage. La figure 50 qui met en relation les valeurs théoriques et réelles pour chaque point d'arrêt de ces montées-descentes (une échelle logarithmique a été appliquée pour que Besançon n'écrase pas le graphique par une valeur trop forte (figure 51) ne fait que confirmer le bon fonctionnement du modèle mais ne valide pas la pertinence de la délimitation spatiale des aires de rabattement. Le modèle de structuration spatiale étant basé, au moins en partie, sur le décompte des émissions et attractions optimales à chaque point d'arrêt, il est naturel de retrouver des valeurs estimées proches des valeurs réelles. On constate tout de même que les valeurs théoriques sont légèrement sous-estimées, certainement en raison des individus non pris en compte car en dehors de l'agrégat de rabattement. Besançon principalement, dans la table des valeurs non logarithmées, voit son nombre de montées-descentes presque divisé par deux, ce qui n'est pas le cas pour les autres gares importantes de notre aire d'étude.

La logique de rabattement concernant le pôle central relève donc d'un fonctionnement spécifique, qu'un zonage quel qu'il soit, ne peut que difficilement retranscrire.

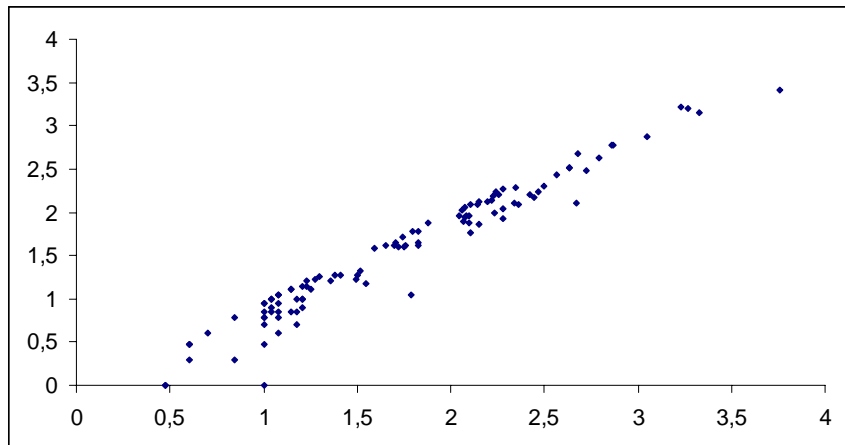


Figure 50 : Croisement des montées-descentes prédites et réelles, échelle logarithmique

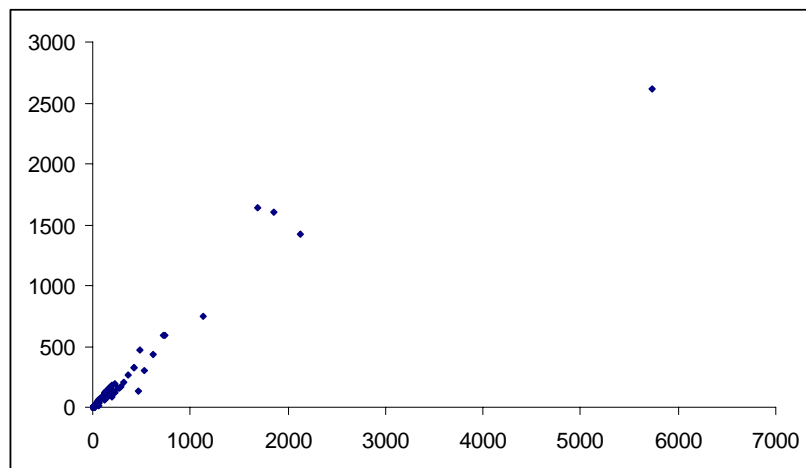


Figure 51 : Croisement des montées-descentes prédites et réelles

La mise en place d'un protocole de validation du zonage dans l'étape de génération de trafic reste donc à accomplir pour parfaire cette première étape du modèle à quatre étapes.