



HAL
open science

**Simulation et optimisation spatiales : Vers la
constitution d'un outil d'aide à la décision pour la
fermeture / ouverture de classes. Une application pour
l'enseignement agricole initial**

Marie-Benoît Magrini, Florent Bonneu, Christine Thomas-Agnan, Sandrine
Coelho

► **To cite this version:**

Marie-Benoît Magrini, Florent Bonneu, Christine Thomas-Agnan, Sandrine Coelho. Simulation et optimisation spatiales : Vers la constitution d'un outil d'aide à la décision pour la fermeture / ouverture de classes. Une application pour l'enseignement agricole initial. 2011. hal-00723220

HAL Id: hal-00723220

<https://hal.science/hal-00723220>

Submitted on 23 Aug 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

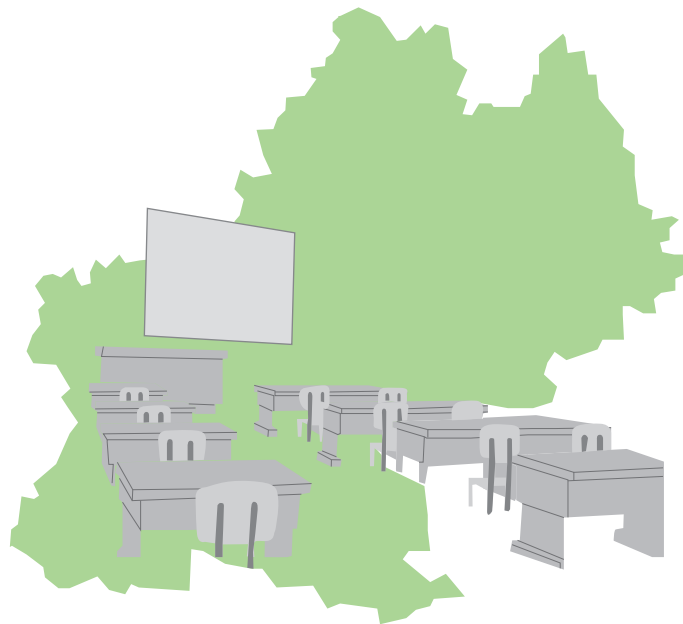
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

$$\sum_{\{\mu\}} f(b) = \int_{\Omega}$$

SIMULATION ET OPTIMISATION SPATIALES :

Vers la constitution d'un outil d'aide à la décision
pour la fermeture / ouverture de classes.

Une application pour l'enseignement agricole initial



Pour citer ce document

Magrini Marie-Benoît, Bonneau Florent, Thomas-Agnan Christine, Coelho Sandrine, (2011), *Simulation et optimisation spatiales : Vers la constitution d'un outil d'aide à la décision pour la fermeture / ouverture de classes. Une application pour l'enseignement agricole initial*, Document d'analyse PSDR3 Midi-Pyrénées-Projet COMPTEER, 28p + annexes.

© INRA - Toulouse.

Conception-réalisation : Yaël Kouzmine (INRA-Toulouse)

Sommaire

Les auteurs	5
Remerciements	5
Précision	5
Introduction	7
1. Offre scolaire et arbitrage formation-distance	8
1.1. Choix de formation et choix d'établissement : l'importance de la distance dans le processus de décision	9
1.1.1. Choix d'établissement, puis choix de formation	9
1.1.2. Choix de formation, puis choix d'établissement.....	9
2. Données : la distance dans la distribution des élèves aux sites d'enseignement agricole	11
2.1. L'enseignement agricole : une structure distincte de l'éducation nationale.....	11
2.1.1. L'enseignement agricole en France	11
2.1.2. Évolution des effectifs et offre de formation dans un monde agricole en déclin démographique	11
2.2. La répartition des effectifs de l'enseignement agricole dans la région Midi-Pyrénées	12
2.2.1. L'importance de la distance dans le choix des élèves de Midi-Pyrénées	13
2.2.2. Les temps de trajet : le choix d'un critère de distance-temps.....	15
2.3. Présentation des deux problèmes de fermeture et d'ouverture de classes étudiés.....	16
2.3.1. Un problème de fermeture pour la formation « BTSA ACSE ».....	16
2.3.2. Le problème d'ouverture pour la formation « Bac Pro Service en Milieu Rural »	17
3. Modélisation et simulation de la répartition spatiale des élèves sur le territoire	17
4. Méthodologie du choix d'ouverture et de fermeture de classes à partir des simulations	20
4.1. Estimation de la distance-temps d'une localisation simulée d'un élève	20
4.2. Minimisation de la somme des distances et fermeture	20
4.3. Minimisation de la somme des distances et ouverture.....	21
4.4. Significativité des solutions et classement des choix d'ouverture/fermeture	21
4.5. Les résultats pour les deux problèmes étudiés	22
Conclusion	25
Bibliographie	27
Annexes	29



© MAGRINI Marie-Benoit

Les auteurs

Marie-Benoit Magrini est Ingénieur de recherche à l'INRA Toulouse et membre de l'UMR 1248 AGIR (Agrosystèmes et agricultures, Gestion des ressources, Innovations et Ruralités).

@ : mbmagrini@toulouse.inra.fr

Florent Bonneau est Maître de conférences en Mathématiques à l'Université d'Avignon et des Pays de Vaucluse et membre du Laboratoire d'Analyse Non Linéaire et Géométrie.

@ : Florent.Bonneu@univ-avignon.fr

Christine Thomas-Agnan est Professeur des Universités à l'Université Toulouse 1 Capitole et membre du GREMAQ (Groupement de Recherche en Economie Mathématique et Quantitative), Toulouse School of Economics.

@ : cthomas@cict.fr

Sandrine Coelho appartient à la Société BVA de Balma (31).

@ : sandrine.coelho@bva.fr

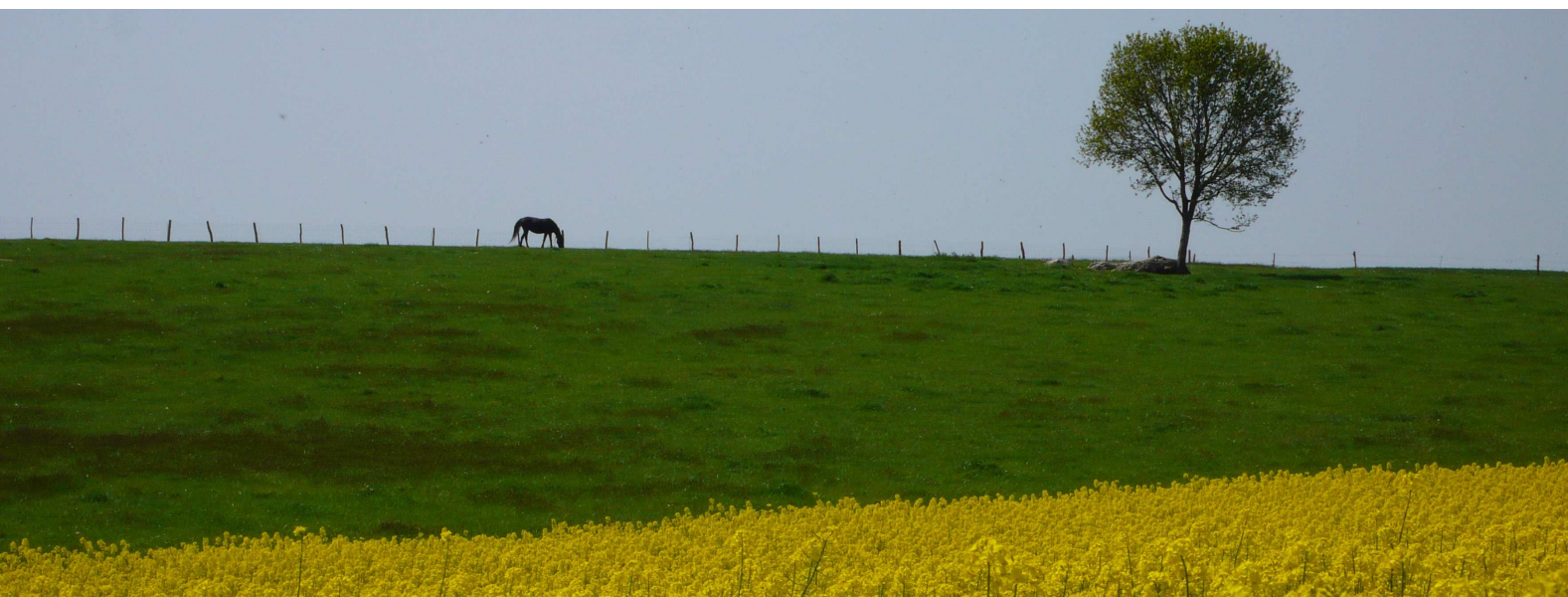
Remerciements

Les auteurs remercient l'ensemble des personnes ayant contribué à la réalisation de cette étude : la Direction Régionale de l'Agriculture et de la Forêt de Midi-Pyrénées pour leur collaboration et la mise à disposition des données, et tout particulièrement Jean-Claude Teurlay, Chef de la Division Statistique, Isabelle Seiller, Déléguée à l'Ingénierie des Formations ; Mohammed Hilal, Chercheur de l'UMR CESAER de l'INRA-Dijon pour la mise à disposition des données distance-temps ; la société Géosignal du groupe BVA pour son assistance informatique ; le Conseil Régional de Midi-Pyrénées et l'INRA pour le financement du programme de recherche sur la représentation et la simulation de données spatiales; ainsi que Georges Maffre, Thomas Gagnaire et Honoré Komol pour leur contribution.

Précision

Cette étude a, par ailleurs, fait l'objet d'une publication scientifique sous la référence bibliographique suivante :

MAGRINI M.-B., BONNEU F., THOMAS-AGNAN C., COELHO S., 2011, "Educational planning: a simulation approach for creation or closure of school classes", Environment and Planning B : Planning and Design, 38.



Introduction

L'organisation de la carte scolaire des formations est un problème récurrent pour l'administration publique française. Ce problème renvoie le plus souvent à des choix d'ouverture et de fermeture de classes, voire d'établissements. Pour adapter la carte scolaire, les acteurs publics utilisent le plus souvent des règles de décision fondées sur des seuils d'effectifs qui permettent d'ajuster annuellement l'offre scolaire à la demande. Les établissements de formation sont ainsi annuellement exposés à de possibles ouvertures ou fermetures de classes en fonction de l'évolution des inscriptions.

Ces choix restent difficiles car ils sont soumis à des fluctuations dans le temps de la demande de formation, dont les déterminants ne sont pas facilement identifiables et prévisibles. Pour autant, il existe un déterminant essentiel : la distance géographique séparant le domicile parental de l'établissement. Toutes les études sur le sujet s'accordent en effet à reconnaître son rôle majeur dans le choix des élèves d'un établissement de formation. Son influence est d'autant plus forte que la formation choisie est courante et/ou de niveau général.

Alors que la littérature est relativement développée sur les mobilités individuelles (Lemistre et Magrini, 2011) et sur les problèmes de localisation dans l'espace, notamment pour les activités industrielles (Alberada-Sambola, Fernandez et Laporte 2007), il existe néanmoins très peu d'études sur la question de la localisation des formations ; et celle-ci est essentiellement abordée au travers du problème des ouvertures/fermetures de classes (Müller, Tscharaktschiew, Haase, 2008 ; Hastings, Kane, Staiger, 2005 ; Cahuzac et Tahar, 2000).

Or, ce problème d'ouverture et de fermeture de classes est d'autant plus réactualisé que les contraintes budgétaires sont grandissantes, amenant les pouvoirs publics à rationaliser l'organisation des formations, et plus précisément, à « optimiser » leur répartition spatiale. Tel est l'enjeu notamment pour l'enseignement agricole qui doit faire face désormais à une baisse régulière de son budget, liée en partie à la baisse de ses effectifs depuis une dizaine d'années. Après des années marquées par une forte croissance de ses effectifs (plus de 178 000 élèves de la 4^{ème} aux classes BTSA et préparatoires en 1999), l'enseignement agricole semble aujourd'hui traverser une période de crise qui l'amène à réviser en profondeur la carte scolaire de ses établissements (rapport Férat, 2006).

La présente étude propose d'appréhender ce problème de choix de localisation des fermetures et ouvertures de classes sous l'angle d'un indicateur de distance peu employé dans la littérature : celui du temps d'accès routier entre le domicile de l'élève et le site de formation, dans le cadre de formations initiales courtes de niveau baccalauréat et post-baccalauréat de l'enseignement agricole français. Cette distance-temps permet, *a priori*, de mieux décrire les conditions réelles d'accessibilité de l'élève à son établissement.

Après un retour sur la littérature concernant le rôle de la distance dans les mobilités liées aux études et une présentation de nos données, nous proposons un programme mathématique de résolution du problème d'ouverture et de fermeture de classes, à partir de simulations de populations d'élèves issues d'un modèle ajusté sur les localisations observées des élèves. L'originalité de l'étude réside avant tout dans l'utilisation d'un modèle mathématique définissant une répartition aléatoire de points dans l'espace : la méthode des Processus Ponctuels (Diggle,

2003 ; Bonneu et Thomas-Agnan, 2008), pour simuler la localisation des élèves. Notre modèle, estimé à partir des localisations observées des élèves, permet de générer plusieurs scénarios de population d'élèves et de résoudre le problème d'ouverture/fermeture sur chacun d'eux. L'optimisation basée sur un critère de minimisation de distance-temps globale est réalisée à l'aide d'un algorithme heuristique fournissant des solutions approchées. Cette démarche présente l'avantage de prendre ainsi en compte le caractère aléatoire de la localisation des élèves afin de fournir une indication sur la variabilité des positions optimales du lieu d'ouverture ou de fermeture de classe. Ces problèmes sont résolus pour deux cas de formations soumis par la Direction Régionale de l'Agriculture et de la Forêt de Midi-Pyrénées.

1. Offre scolaire et arbitrage formation-distance

Les quelques études présentes dans la littérature montrent que le processus de choix des formations renvoie à deux considérations majeures¹. La première concerne la proximité géographique du site où est dévolue la formation, ou son accessibilité. La seconde considération a trait aux critères de réussite et d'excellence que peut présenter l'établissement et qui seront différemment perçus selon l'origine sociale des élèves, leurs résultats scolaires antérieurs, etc. (Manski et Wise, *op. cit.*). Le poids de chacun de ces critères dans la décision dépend essentiellement du type de formation choisie. Plus la formation visée est sélective et plus le jeune et sa famille seront prêts à parcourir des kilomètres supplémentaires pour s'offrir le maximum de chance de réussite. Plus au contraire la formation est généraliste et/ou de faible niveau et moins le jeune élève et sa famille sont enclins à choisir un site éloigné du lieu de résidence.

Ce type d'arbitrage renvoie directement au raisonnement de la théorie du capital humain de Becker (1964). Dans cette théorie, un étudiant est indifférent entre poursuivre une année supplémentaire d'études ou entrer sur le marché du travail si le coût de cette année d'étude est égal au rendement salarial (actualisé dans le temps) du niveau de formation actuellement atteint. Un raisonnement similaire peut être appliqué ici : une famille est indifférente entre inscrire son enfant sur un site proche ou distant si les coûts supplémentaires d'accès au site éloigné sont compensés par l'espérance d'une réussite future plus grande.

Cependant, on ne saurait conclure que les établissements présentant une proportion non négligeable d'élèves éloignés de chez eux seraient nécessairement de « bons » établissements. En effet, les procédures d'affectation peuvent être relativement complexes lorsque les établissements peuvent recruter sur la base d'une sélection des élèves. Certains élèves de « mauvais » niveau peuvent alors être refusés dans des établissements proches de chez eux et être contraints de s'éloigner du domicile familial.

Par ailleurs, d'autres critères peuvent entrer en compte dans le choix d'un établissement, tels que, par exemple, la présence d'un internat. Ce critère réduit d'autant l'enjeu de la distance puisque les déplacements ne seront plus quotidiens. D'autres critères en termes de parcours scolaire peuvent aussi influencer le choix. On peut choisir un site éloigné parce qu'il offre des formations que l'élève souhaite poursuivre ultérieurement. Il peut alors préférer être déjà présent dans l'établissement afin d'être sélectionné, peut-être plus facilement, dans la poursuite de ses études.

Les déterminants des choix des formations et de leur établissement semblent donc nombreux. Mais au vu des résultats de la littérature, nous proposons ici une brève synthèse qui insiste,

1. Manski et Wise, *op. cit.* pour étude de cas allemand ; Cahuzac et Tahar *op. cit.* pour étude de cas français ; Hoxby 2003, Manski et Wise 1983 pour des études nord-américaines.

avant tout, sur le rôle crucial de la distance dans ces choix indépendamment de ces autres considérations.

1.1. Choix de formation et choix d'établissement : l'importance de la distance dans le processus de décision

Les quelques articles existant sur le sujet tendent à distinguer deux types de choix non indépendants : le choix d'une formation ou d'un diplôme, et le choix d'un établissement qui dispensera la formation. Le problème reste de savoir si l'un de ces choix précède l'autre. Est-ce que l'élève et sa famille choisissent d'abord une formation, puis regardent quels sont les établissements qui permettent de suivre cette formation ? Ils peuvent alors choisir l'établissement le plus proche. Ou au contraire, est-ce que le choix se porte d'abord sur l'établissement, selon un éventuel critère de distance, puis l'élève et sa famille choisissent parmi les formations offertes par l'établissement retenu ? Les résultats de la littérature, montrent que l'un ou l'autre de ces scénarios s'opère. Mais le premier scénario semble dominant, bien qu'*in fine* il reste difficile de distinguer les deux.

1.1.1. Choix d'établissement, puis choix de formation

Dans le cadre des formations professionnelles initiales, l'étude de Larceneux (1997) sur les vœux des parents montre une certaine « inertie » dans les mobilités d'études dans la mesure où les parents tendent d'abord à choisir un établissement proche, puis à choisir parmi les formations que l'établissement propose. Il semble donc que la distance joue un tout premier rôle dans ces choix.

Ce constat est corroboré par l'étude d'Arrighi (2004) qui montre que pour les jeunes ruraux leur orientation scolaire commence très tôt en fonction de l'offre territoriale de formation. Ils apparaissent en effet plus nombreux dans des filières professionnelles, alors que « *si seul jouait le critère du niveau scolaire, les jeunes ruraux seraient en proportion aussi nombreux dans les filières générales que leurs homologues urbains* »². L'auteur explique ce constat par l'importance qu'accordent les familles à la « *proximité future des fratries* ». Du fait que certains diplômes ne sont vraiment valorisables qu'en milieu urbain, le choix de formation des familles s'oriente vers l'offre professionnelle locale plus adaptée au milieu rural : « *ces formations constituent une "offre de proximité"... qui prépare à un avenir "de proximité"* »³. Ainsi les formations permettant d'accéder aux emplois dits « métropolitains supérieurs » qui sont concentrés dans les plus grandes aires urbaines, loin des campagnes, n'attirent pas les familles rurales et leurs enfants qui souhaitent favoriser une proximité géographique avec leurs origines. Les études de la sociologue Grelet (2006, 2004) confirment également cette forte influence de la proximité géographique en montrant que les parcours scolaires des jeunes ne sont que le reflet de l'offre de formation locale.

1.1.2. Choix de formation, puis choix d'établissement

Dans la lignée de Ourliac et Fourcade (1983), qui ont aussi montré l'importance de « micro-mobilités » pour évaluer les aires d'influence des établissements, l'étude de Cahuzac et Tahar (*op. cit.*) s'intéresse plus particulièrement au cas de la formation BEP (formation accessible à l'issue de la classe de 3ème de collège) dans la région Midi-Pyrénées en France. Ils observent environ 7 000 jeunes âgés en moyenne de 16 ans et répartis sur 94 établissements de la région. Ils montrent une forte interaction entre choix de formation et choix d'établissement par rapport

au critère de distance (distance évaluée à « vol d'oiseau »). Ils constatent que 68% des élèves sont inscrits sur un site à moins de 20km de chez eux et 83% à moins de 30km. Mais leur étude va au-delà de l'observation des effectifs, en s'intéressant également aux voeux émis par les élèves avant leur inscription définitive. Lorsque les voeux sont émis pour une même formation, il ressort que pour près de 50% des élèves le site du 1er voeu est plus proche que celui donné en second voeu et pour 30% les distances entre les sites des deux premiers voeux sont équivalentes. Si les élèves ont émis des voeux de formations différentes, pour plus de 70% les distances entre les sites de chacune des formations choisies sont équivalentes et seulement 13% ont choisi une formation en 1er voeu dont la distance est plus grande que pour la formation choisie en 2ème voeu. Ainsi, qu'il s'agisse d'observer les voeux des élèves ou leur inscription finale, on constate qu'une large majorité d'élèves préfèrent les sites de formation les plus proches. Mais surtout, les auteurs concluent qu'une fois la formation choisie, les élèves choisissent en priorité le site le plus proche.

Cette idée se retrouve également dans l'étude de Müller *et al.* (*op. cit.*) en Allemagne, qui observent que la majorité des élèves âgés entre 10 et 19 ans sont inscrits dans le collège le plus proche de chez eux⁴ pour un type donné d'établissement, ie. de cursus offert par le site de formation : « *Students attending schools with a common profile mostly are located in the surrounding catchment area. Thus, one could assume that proximity is an important factor for school choice* »⁵. Alors qu'il n'existe pas sur leur territoire d'observation de règles d'affectation par les autorités administratives de la ville (la ville de Dresden comptant 14 000 élèves sur 23 collèges). Le constat qu'un grand nombre d'élèves ne choisissent pas finalement le collège le plus proche de chez eux (80% des élèves), mais celui qui est le plus proche parmi les collèges de même profil (70% de ces élèves) renforce l'idée que les familles et leurs enfants choisissent d'abord un type de formation, puis un établissement. Mais aussi, le constat de faibles distances de séparation laisse présager que le choix de formation a également été établi à partir de celles offertes par l'ensemble des collèges situés en deçà d'une distance jugée comme maximale pour assurer des trajets quotidiens. Il reste donc difficile de supposer que le premier choix sur les formations s'établisse indépendamment de la distance.

Ainsi, au côté du critère de sélection, la distance apparaît donc comme un déterminant majeur. En effet, tous ces résultats montrent bien que la distance ou la proximité géographique, joue un rôle fondamental dans le choix d'une formation et/ou d'un établissement, quelque soit l'ordre de ces choix. Cette influence de la distance trouve bien sûr des éléments d'explication liés aux coûts de déplacements. Rappelons que Combes et Lafourcade (2005) ont montré que la distance géographique ou le temps de trajet sont très fortement corrélés aux coûts de transport (coefficient de corrélation de 0.99) à l'échelle de la France entière. Cette dimension est d'autant plus renforcée aujourd'hui que nos économies développées traversent une crise énergétique sans précédent.

L'étude qui nous a été soumise reposant sur des formations agricoles de niveau général et faiblement sélective, il est donc attendu que la distance joue un rôle d'autant plus important dans le choix des élèves. Nous proposons, dans la section suivante, d'explorer les données qui nous ont été fournies et de mesurer la proportion d'élèves ayant choisi des sites de formations les plus proches de chez eux. Pour cela, si la distance géographique euclidienne est souvent employée comme proxy, nous choisissons ici de raisonner sur des distances-temps. En effet, comme Hilal (2007) le met en avant, les populations qui se déplacent accordent plus d'importance au temps d'accès aux services qu'à la distance kilométrique.

4. Dans leur étude, le critère de distance observé est la distance métrique du plus court chemin.

5. Page 5

2. Données : la distance dans la distribution des élèves aux sites d'enseignement agricole

La distance apparaissant comme un déterminant majeur du choix du site de formation, particulièrement dans le cadre de formations des cycles secondaires et post-baccalauréat court, nous proposons de vérifier cette hypothèse sur la base de simples statistiques sur les données fournies. Avant de présenter les distances-temps des élèves de l'enseignement agricole de la région, nous commençons par donner une brève description de l'enseignement agricole en France et dans la région Midi-Pyrénées. Nous terminerons en détaillant les deux formations retenues pour résoudre des problèmes d'ouverture et de fermeture de classe.

2.1. L'enseignement agricole : une structure distincte de l'éducation nationale

2.1.1. L'enseignement agricole en France

L'enseignement agricole (EA) est le fruit d'une longue tradition française qui remonte à 1848 avec la création des premières fermes-écoles. Rattaché au Ministère de l'Agriculture, l'EA a été, et continue d'être, un moteur essentiel de modernisation et de développement de l'agriculture française. Longtemps destiné aux enfants d'agriculteurs, cet enseignement a permis d'apporter au sein des exploitations agricoles, bien souvent familiales, un renouvellement progressif des compétences. A partir des années 1950 cet enseignement a su diversifier son offre de formation et renouveler son public pour parer à la baisse progressive du nombre d'agriculteurs. Il couvre désormais une gamme complète de formations touchant à la fois à la production, à la transformation et aux services, et s'étendant de la classe de 4^{ème} jusqu'au 3^{ème} cycle de l'enseignement supérieur. Il est dispensé en formation initiale, mais également par la voie de l'apprentissage et de la formation continue. L'enseignement initial qui regroupe les formations du second degré et post-baccalauréat est généralement dénommé « enseignement agricole technique » et se distingue de l'enseignement supérieur comprenant les établissements classés dans la catégorie « grandes écoles », telles que les écoles nationales supérieures d'agronomie (ENSA). Notre étude ne s'intéressant qu'à la première catégorie d'enseignement, nous utiliserons ici le terme plus simple d' « enseignement agricole ».

2.1.2. Évolution des effectifs et offre de formation dans un monde agricole en déclin démographique

La capacité de l'EA à s'adapter aux nouveaux métiers, fait que malgré un contexte de déclin démographique du monde agricole, cet enseignement a su globalement maintenir une forte expansion de ses effectifs jusqu'à la fin des années 1990⁶. En effet, rappelons que le nombre d'actifs agricoles a été divisé par deux entre 1970 et 2000. La part de la population active agricole représente aujourd'hui moins de 3,5% de la population active française. Pour autant, les effectifs d'élèves et d'étudiants de l'enseignement agricole ont constamment progressé, passant de 130 000 élèves au début des années 1980 à plus de 178 000 élèves à la fin des années 1990. Mais depuis le début des années 2000, cet enseignement est touché par un renversement de tendance : les effectifs sont marqués par une baisse progressive et ils se situent désormais autour de 175 000 élèves. Aujourd'hui, les élèves se répartissent sur 850 établissements dans l'ensemble de la France⁷, avec une part toujours aussi importante de l'enseignement privé qui représente près de 62% des effectifs et 75% des établissements.

6. « L'enseignement agricole a investi de nouveaux créneaux de formation, en amont ou en aval de son « coeur de métier », et recrute désormais au-delà de son vivier « traditionnel » issu du monde agricole. » (Rapport Férat, *op. cit.*, page 16).

7. Comparativement, le second cycle professionnel relevant de l'éducation nationale comprend près de 724 000 élèves à la rentrée 2005, du CAP au bac professionnel ; s'y ajoutant plus de 230 000 étudiants en sections de techniciens supérieurs (rapport du Sénat 2006).

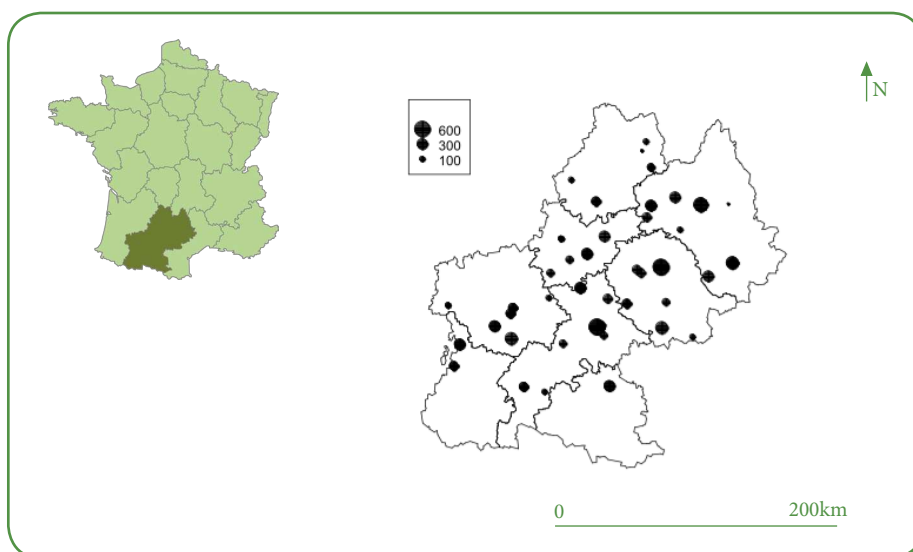
8. Les calculs de distance seront ainsi plus homogènes car quelques élèves peuvent venir de très loin, notamment de l'étranger, ce qui augmente considérablement le critère de somme des distances que nous utilisons par la suite.

Précisons ici que l'attractivité de l'enseignement agricole a pu se maintenir grâce à un élargissement du vivier de recrutement. Alors que les enfants d'agriculteurs ou de salariés agricoles représentaient 55% des élèves en 1975, leur part est tombée à 17% aujourd'hui. Les enfants d'agriculteurs restent bien sûr relativement plus nombreux que dans les établissements relevant de l'éducation nationale où ils représentent moins de 3% des effectifs. Cet attrait s'explique alors essentiellement par la diversification de l'offre de formation qui est passée d'un pôle exclusivement « agricole » au développement de métiers relevant autant de l'agriculture, que de l'alimentation, de l'environnement que des services, particulièrement au sein des territoires ruraux. L'enseignement agricole prépare ainsi à quatre grandes familles de métiers liés à : la production, la transformation, l'aménagement et les services en milieu rural.

Néanmoins, la légère baisse actuelle des effectifs accompagnée d'une prise de conscience collective de la nécessité de réduire les dépenses publiques s'est traduite par une restriction des dépenses de l'enseignement agricole et une accélération des fermetures de classes dont le nombre dépasse désormais celui des ouvertures. Ainsi, entre 2004 et 2006 on comptabilise plus de 200 fermetures nettes. Si la région Midi-Pyrénées est l'une des régions ayant un enseignement agricole important, elle n'échappe pas à cette logique de rationalisation budgétaire.

2.2. La répartition des effectifs de l'enseignement agricole dans la région Midi-Pyrénées

La région Midi-Pyrénées fait partie des régions de France ayant des effectifs de l'enseignement agricole parmi les plus importants. La région est composée de 3 019 communes dont 43 possèdent un site d'enseignement agricole initial et dont un peu plus de la moitié sont privés. Notons ici que nous considérons dans cette étude l'ensemble des sites de formation, qu'ils soient publics ou privés. En effet, la forte complémentarité de ces structures et leur collaboration étroite avec les autorités administratives fait qu'il est difficile de les distinguer et qu'ignorer les sites privés conduirait à occulter près de la moitié des élèves. Au total l'ensemble de ces communes accueillent à la rentrée 2007 un peu plus de 10 000 élèves dont près de 90% sont originaires de la région⁸. Dans ce qui suit nous considérerons uniquement les élèves résidant dans la région⁸. La carte suivante représente la répartition des élèves inscrits sur des communes ayant un établissement d'enseignement agricole.



Carte 1.
Répartition des élèves sur les 43 communes d'enseignement agricole en 2007 de la région Midi-Pyrénées

L'ensemble des sites de formation accueille au total 74 formations diplômantes différentes. Chaque formation est identifiée comme un diplôme type de l'Education Nationale (CAP, BEP, Bac, BTS, etc.) avec une spécialité développée par l'enseignement agricole. Par exemple, dans le cadre des deux formations qui sont étudiées : la formation « Bac Pro Services en Milieu Rural » relève du diplôme de type « Bac Pro » avec comme spécialité « les Services en Milieu Rural » ; la formation « BTSA ACSE » relève du diplôme BTS Agricole avec comme spécialité « Analyse et Conduite de Système d'Exploitation ».

Concernant plus précisément les temps d'accès aux sites (tableau 1), nous constatons que pour les trois-quarts des élèves, il existe une commune ayant un site d'enseignement agricole situé à moins de 25 minutes. Néanmoins, ces fréquences ne tiennent pas compte de la spécificité des enseignements dispensés sur les sites comme nous allons l'aborder dans ce qui suit.

Tableau 1.
Proportion d'élèves ayant un site d'enseignement à proximité

Existence d'un site situé :	% cumule
Sur la même commune	13
à moins de 5 min	15
à moins de 10 min	24
moins de 15min	41
moins de 20 min	60
moins de 25 min	75
moins de 30 min	83
moins de 35 min	89
moins de 40 min	96
moins de 50 min	99

2.2.1. L'importance de la distance dans le choix des élèves de Midi-Pyrénées

Au vu des résultats de la littérature présentés précédemment, nous proposons d'observer la répartition des effectifs selon des classes de distance et d'évaluer la proportion d'élèves inscrits :

- i) sur le site le plus proche de chez eux, sans considérer la formation choisie ;
- ii) sur le site le plus proche parmi les sites dispensant la formation choisie par l'élève.

Nous présentons ces résultats selon deux critères de distance : une distance géographique kilométrique évaluée à vol d'oiseau et une distance-temps évaluée comme le temps de trajet routier. Ces deux indicateurs sont évalués entre deux points centraux des communes du site de formation et de résidence de l'élève (ie. du domicile parental) que sont les mairies. Plus précisément ici, les distances temps représentent le temps de trajet routier en heures de pointe⁹. Si l'établissement est situé sur la même commune de résidence de l'élève, alors la distance géographique ou le temps de trajet valent 0.

A priori, cette distance routière en temps apparaît comme un meilleur proxy que la distance kilométrique évaluée à vol d'oiseau et généralement employée dans les études sur le sujet. La distance temps révèle mieux en effet les conditions réelles d'accessibilité qui peuvent varier selon le type de commune considérée. De faibles distances géographiques à vol d'oiseau pour relier deux communes urbaines peuvent cacher des temps de trajet relativement longs du fait de problèmes d'engorgement routier. Il en est de même pour deux communes rurales : il suffit qu'elles soient séparées par un obstacle géographique important (par exemple, une montagne) pour que leur liaison routière soit également très longue.

9. Ces distances routières exprimées en temps sont obtenues par une recherche du plus court chemin minimisant le temps de déplacement entre deux communes sur l'aller et le retour. Ce temps de déplacement entre communes prend en compte l'environnement géographique, les conditions de circulation etc. Pour plus de détails, voir Hilal (2007) page 44.

10. Un tiers des effectifs de ces formations mono-sites ne sont pas originaires de la région Midi-Pyrénées contre un taux moyen de 10% pour les autres formations.

Ensuite, le calcul de la proportion d'élèves inscrits sur le site le plus proche pour la formation choisie n'a de sens que si la formation considérée est dispensée sur différents sites. Auquel cas, le jeune est amené à choisir parmi eux. Le tableau suivant précise le nombre de formations présentes sur un seul site et celles qui existent sur plusieurs sites. Il ressort que très peu d'élèves n'ont finalement qu'un seul choix possible pour des formations présentes sur un unique site (8,5% des inscrits de Midi-Pyrénées). Il s'agit de formations très spécifiques (spécialité ou option particulière) et qui, d'ailleurs, accueillent une forte proportion d'élèves nonoriginaires de la région¹⁰.

Nombre de sites	Part des formations concernées	Part des élèves inscrits
1 seul site possible	46%	8,5%
2 sites possibles	21%	9%
Entre 3 et 8 sites possibles	17%	27,5%
Entre 9 et 35 sites possibles	10%	55%

Tableau 2.
Nombre de sites par formation et part d'élèves de Midi-Pyrénées inscrits

N.B. : le nombre maximum de sites différents pour une même formation est de 35

Par ailleurs, une caractéristique importante de l'enseignement agricole par rapport à l'éducation nationale est d'offrir des possibilités d'internat sur tous ses sites. Ainsi, dans l'ensemble 56% des élèves inscrits sont internes (54% dans le public et 60% dans le privé). Néanmoins, le recours à l'internat varie selon le diplôme préparé. Par exemple, les élèves en BTSA sont nettement moins nombreux en internat (13% seulement) alors que les élèves préparant un CAP sont à 70% internes. Il est probable que ces différences s'expliquent par l'âge : les premiers ont une moyenne d'âge de 20 ans contre 17 ans pour les seconds. On observe au final que trois-quarts des internes ont 19 ans ou moins. On peut penser que l'internat modifie, *a priori*, l'influence du critère de distance puisque les déplacements ne sont plus quotidiens. Néanmoins, il est fréquent que les élèves rentrent chez eux tous les week-end. Aussi l'influence de la distance reste présente, mais en moindre mesure par rapport à des élèves demi-pensionnaires ou externes. C'est ce que révèlent les résultats suivants (tableau 3).

	Distance à vol d'oiseau			Distance-temps		
	Tous	DP ou externe	Interne	Tous	DP ou externe	Interne
Part des élèves situés sur le site le plus proche de chez eux	27%	46%	13%	26%	44%	12%
Part des élèves situés sur le site le plus proche de chez eux pour la formation choisie (calcul uniquement pour les élèves dont les formations existent sur plusieurs sites)	60%	79%	46%	61%	79%	47%

Tableau 3.
Part des élèves inscrits sur le site le plus proche

N.B. : DP = Demi-pensionnaire

Ces résultats montrent d'abord que, globalement, l'usage de l'un ou l'autre de ces indicateurs de distance (à vol d'oiseau ou en temps de trajet) conduit à des conclusions comparables et confirment les résultats de la littérature. En effet, pour une formation donnée, les élèves sont inscrits, pour la plupart, sur le site le plus proche de chez eux. Ce taux étant particulièrement important pour les élèves non-internes (79%).

Cependant, l'indicateur des distances temps reste un meilleur indicateur individuel des conditions d'accessibilité (Hilal, *op. cit.*), nous poursuivrons donc nos commentaires et nos calculs avec uniquement les distances-temps.

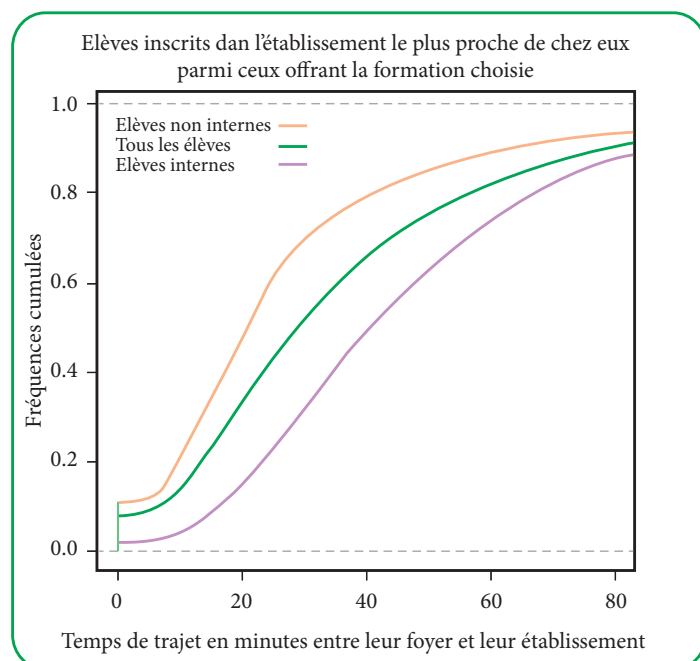
Au-delà de ce critère global, nous observons quelques variabilités dans l'influence de la distance selon d'autres critères qui méritent d'être mentionnés.

D'abord, nous constatons que le critère de distance semble d'autant plus important que l'élève réside dans un espace peu concentré. En effet, la part d'élèves situés sur le site le plus proche pour la formation choisie tombe à 58% en milieu urbain contre respectivement 62% et 68% pour les élèves du périurbain et du rural. Ce dernier résultat se rapproche de celui de Mayer *et al.* (*op. cit.*) qui constatent également que plus l'élève réside dans un espace concentré et plus la distance perd de son influence car le jeune a alors accès à une plus grande offre de transports collectifs. Ensuite, les élèves inscrits dans un établissement privé sont moins nombreux que ceux du public à se situer sur le site le plus proche pour la formation choisie (respectivement 60% et 67%). Ceci peut s'expliquer par la recherche pour certaines familles de caractéristiques spécifiques d'enseignement liées au privé que nous ne pouvons expliquer ici.

2.2.2. Les temps de trajet : le choix d'un critère de distance-temps

Si l'on considère plus précisément les temps d'accès des élèves ayant choisi le site le plus proche pour la formation choisie, la moyenne est de 37 minutes (30 et 47 minutes respectivement pour les non-internes et les internes). Plus précisément, la moitié des élèves sont situés sur un site à moins d'une demi-heure de chez eux, ce taux augmentant considérablement pour les élèves non-internes (70%) alors qu'ils ne sont qu'un tiers chez les élèves internes (graphique 1).

Graphique 1.
Fréquences cumulées des
temps de trajet



11. Ces temps de trajet ont été calculés par l'auteur sur la base des déclarations des habitants sur leur commune de résidence et celle d'exercice de leur activité quotidienne lors du recensement de la population en 1999. Pour les jeunes, l'activité correspond au lieu d'études.

12. Comme nous l'avons exposé dans la section précédente, il est difficile d'établir les raisons des choix d'élèves pour telle ou telle formation et/ou pour tel ou tel établissement en dehors du critère de distance. Les explications peuvent être nombreuses et nous ne disposons pas d'autres informations à ce jour.

Ces résultats peuvent être rapprochés de ceux de Hilal (*op. cit.*). Son étude montre en effet que le temps d'accès moyen pour les déplacements domicile-études des jeunes âgés entre 7 et 18 ans est de 23 minutes en France (dont 9 minutes pour les 7-15 ans)¹¹.

L'auteur observe également que les jeunes scolarisés de plus de 15 ans tendent à se déplacer plus loin que les actifs, ces derniers ayant un temps moyen de déplacement de 17 minutes. Enfin, il montre aussi que plus une commune est éloignée du pôle urbain qui lui est le plus proche et plus le temps d'accès moyen aux équipements augmente. Il constate ainsi que le temps d'accès moyen des scolaires est d'une vingtaine de minutes pour les élèves des pôles urbains contre une trentaine de minutes pour ceux des communes rurales.

Bien que le site de formation choisi soit ici le plus proche, nos résultats montrent que les temps de trajet sont plus longs. Ceci est probablement dû à l'importance des effectifs provenant de zones rurales (47% des inscrits). Néanmoins, les temps de trajet moyens restent ici similaires quelque soit le type de commune de résidence de l'élève, du fait probablement aussi de l'importante présence d'établissements d'enseignement agricole en zone rurale qui permet d'amortir les temps de trajet de ces espaces peu concentrés.

Quant aux élèves qui n'ont pas choisi le site le plus proche pour la formation choisie, nous observons des écarts de temps qui restent faibles particulièrement pour ceux qui ne sont pas internes. Pour ces derniers, la moitié d'entre-eux ont un temps de trajet supplémentaire de moins de 15 minutes.

Au final, bien que d'autres considérations mériteraient des investigations supplémentaires et nécessiteraient des enquêtes spécifiques¹², il ressort, comme dans la plupart des études sur le sujet, que la majorité des élèves sont inscrits sur le site le plus proche de chez eux pour la formation choisie. C'est sur la base de cette hypothèse que nous proposons de résoudre le problème d'ouverture et de fermeture de classes pour deux cas de formation soumis par la Direction Régionale de l'Agriculture et de la Forêt de Midi-Pyrénées et présentés ci-après.

2.3. Présentation des deux problèmes de fermeture et d'ouverture de classes étudiés

2.3.1. Un problème de fermeture pour la formation « BTSA ACSE »

La formation BTSA ACSE est une formation accessible après le bac qui s'étale sur deux années. Elle permet l'obtention du Brevet de Technicien Supérieur Agricole pour l'Analyse et la Conduite de Système d'Exploitation. Ces diplômés se destinent principalement à être des agriculteurs ou des conseillers techniques dans une exploitation ou un organisme professionnel. Cette formation accueille au total en Midi-Pyrénées un peu plus de 300 élèves, dont plus de 80% sont originaires de la région. Elle est dispensée sur 8 sites différents parmi l'ensemble des sites d'enseignement agricole. Chaque site (ou commune) possède une seule classe égale au nombre d'inscrits. Cette formation concernant essentiellement des élèves majeurs, peu d'élèves sont internes (18% pour les 1ères années). Plus de la moitié des élèves ont choisi le site le plus proche de chez eux, particulièrement pour les élèves qui ne sont pas internes (voir tableaux en annexe pour plus de détails). Cette formation qui existe depuis plus de 10 ans a vu ses effectifs régulièrement augmenter jusqu'à un relatif tassement depuis le début des années 2000, et même une légère baisse de 2% entre 2006 et 2007. Ce relatif manque d'attractivité par rapport à d'autres formations en pleine expansion, conduit la DRAF à décider de fermer cette formation sur l'un des sites de la région.

2.3.2. Le problème d'ouverture pour la formation « Bac Pro Service en Milieu Rural »

Le Bac Pro Service en Milieu Rural (dénommé Bac Pro dans ce qui suit) est une formation qui fait suite à une classe de 1ère de la même spécialité en lycée. Elle est accessible après une seconde générale (de l'éducation nationale ou agricole). Pour plus de clarté, nous considérons donc que la formation Bac Pro est une formation en deux années : la classe de Première, puis la classe de Terminale. La moyenne d'âge des élèves en première est de 18 ans et celle des élèves en terminale de 19 ans. Au total cette formation compte un peu moins de 500 élèves originaires de la région et répartis sur 8 sites. Les élèves sont plus jeunes et la moitié d'entre eux sont internes. La proportion d'élèves situés sur le site le plus proche est ici des plus élevées : 85% des élèves de 1ère et de 2ème année (voir tableaux en annexe pour plus de détails). Cette formation a été ouverte pour la 1ère fois en 2005 sur 6 sites, puis sur deux sites supplémentaires en 2006. La forte croissance de ses effectifs conduit ici au contraire la DRAF à privilégier l'ouverture de cette formation sur un nouveau site.

Dans la mesure où une logique d'ouverture et de fermeture n'a de sens que si elle a lieu sur une classe de première année¹³, nous proposons d'étudier ces deux problèmes sur la base de l'observation des élèves inscrits en 1ère année uniquement et d'utiliser un critère de minimisation des distances-temps pour évaluer le site de fermeture ou d'ouverture. La section suivante expose la méthodologie employée et les résultats obtenus. La méthode employée propose de traiter ces problèmes d'ouverture et de fermeture de classes en intégrant une dimension aléatoire dans la localisation des élèves.

3. Modélisation et simulation de la répartition spatiale des élèves sur le territoire

Cette section présente la méthodologie employée pour décider sur quel site fermer et ouvrir les deux formations considérées. Nous proposons de fonder ce choix sur un critère de minimisation de la somme des distances-temps des élèves aux sites. Ce critère est appliqué sur une série de populations d'élèves simulées. En effet, le choix de minimiser la somme des distances perd de son intérêt si ce critère ne s'applique que sur l'observation des effectifs d'une année donnée. Ce choix apparaîtrait relativement arbitraire. Nous proposons donc de générer les positions de différentes populations d'élèves, à partir d'un modèle de répartition aléatoire de points dans l'espace qui est estimé sur les données réelles des élèves inscrits. Des approches similaires ont été conduites pour d'autres problèmes de localisation dans l'espace, par exemple, pour le choix d'emplacement d'une nouvelle caserne de pompiers (Bonneu et Thomas-Agnan, 2008). Parmi les modèles requis pour ce type d'approche (Diggle, 2003), nous avons choisi celui d'un processus ponctuel de Poisson.

Certains problèmes d'affectation optimale dans l'espace peuvent être résolus à l'aide de logiciels de géo-marketing¹⁴. Il s'agit, par exemple pour un commercial, de déterminer l'itinéraire optimal de ses visites de magasins compte tenu de leur localisation. Mais la résolution de ces problèmes est établie « au jour le jour » sur des données réelles et non sur des données stochastiques qui rendraient compte de leur variabilité. Or dans notre problème, le choix d'une ouverture ou d'une fermeture de classe apparaît comme une solution unique qui s'appliquera pour au moins quelques années, alors que le nombre et la position des élèves peuvent varier d'une année sur l'autre. L'enjeu ici est donc de déterminer de solutions « optimales » en tenant compte du

13. En effet, dans le cadre d'une fermeture, si celle-ci a lieu sur une classe de deuxième année, alors les élèves de 1ère année n'auraient pas la possibilité d'achever leur formation. Il convient donc de commencer par fermer une classe de 1ère année. De même, dans le cadre d'une ouverture, une classe de 2ème année n'a de légitimité sur un même site que s'il existe une classe de 1ère année. Il convient donc aussi de commencer par ouvrir une classe de 1ère année.

14. Notamment avec le logiciel GEOCONCEPT (Da Silva et Daniaud 2004).

caractère aléatoire de la localisation des élèves.

La prise en compte de la nature stochastique de cet ensemble de points dans l'espace s'effectue au travers de l'estimation d'un modèle probabiliste. Notre jeu de données fournit une localisation des élèves approximative par le biais de leur commune d'appartenance. Cette spécificité nous a entraînés à considérer deux méthodes de modélisation possibles.

Une première approche discrète considère pour chaque commune que le nombre d'élèves est la réalisation d'une variable aléatoire de Poisson. Une estimation naturelle du paramètre d'espérance de cette variable serait alors le nombre d'élèves qui est observé sur la commune. Cette méthode présente le désavantage majeur de ne pas permettre d'obtenir, lors des simulations, d'élèves localisés dans les communes où aucun élève n'aurait été localisé en 2007.

La seconde méthode que nous employons se base sur la théorie des processus ponctuels spatiaux qui sont souvent utilisés pour modéliser un ensemble aléatoire de points dans un espace Ω de dimension $d \geq 2$. Dans notre cas, l'espace d'observation Ω est la région Midi-Pyrénées avec $d=2$ (chaque observation ou point, i.e. chaque élève, est repéré sur la carte par deux coordonnées spatiales : la longitude et la latitude). Contrairement au cadre très répandu où les localisations sont considérées comme plusieurs réalisations indépendantes entre elles et issues d'une même variable aléatoire, les processus ponctuels spatiaux présentent deux avantages : la considération de l'aléa sur le nombre de points et l'éventuelle auto-corrélation entre ces points. Ces caractéristiques et le nombre important de modèles de processus ponctuels spatiaux possibles justifient leur utilisation croissante dans des domaines comme l'astronomie, l'écologie, la gestion forestière, l'épidémiologie spatiale, l'analyse d'image...

La famille de processus ponctuels la plus élémentaire, et à laquelle nous avons ici recours, est celle des processus ponctuels de Poisson qui modélisent une répartition aléatoire des points sans interaction (i.e. les points sont localisés indépendamment des uns des autres¹⁵). Ces processus sont identifiés par leur fonction intensité λ ou, plus précisément, par la mesure de cette intensité μ définie pour tout sous-espace d'observation B inclus dans Ω comme :

$$N(B) \sim P(\mu(B)), \forall B \subset \Omega$$

De manière heuristique, $\lambda(s)ds$ représente la probabilité d'apparition d'un point dans un espace sphérique de centre s et de volume ds . $\mu(B)$ correspond ainsi à l'espérance du nombre d'apparition de points dans la sous-région B définissant ainsi une « intensité » d'observation des points.

Pour $0 < \mu(B) < \infty$, un processus ponctuel de Poisson X est caractérisé par une variable aléatoire $N(B)$ du nombre de points dans B pour le processus X distribué selon une loi de Poisson de moyenne $\mu(B)$

$$N(B) \sim P(\mu(B)), \forall B \subset \Omega$$

15. Hypothèse de « Complete Spatial Randomness ».

Dans le cadre de ce processus ponctuel de Poisson, l'intensité peut être reliée à la fonction de densité d'une variable aléatoire. En effet, $(X_B | N(B) = n)$ est tout simplement un processus de n réalisations indépendantes et identiquement distribuées issues de la variable aléatoire de densité $f(s) = \lambda(s)/\mu(B)$ pour tout $n \in \mathbb{N}$ et $B \subset \Omega$ (où X_B est le processus ponctuel restreint au sous-ensemble $B \subset \Omega$).

La dernière propriété montre le lien qui existe entre l'intensité d'un processus ponctuel de Poisson et la densité d -dimensionnelle f définissant la loi de distribution des localisations S_i conditionnellement au nombre de points N :

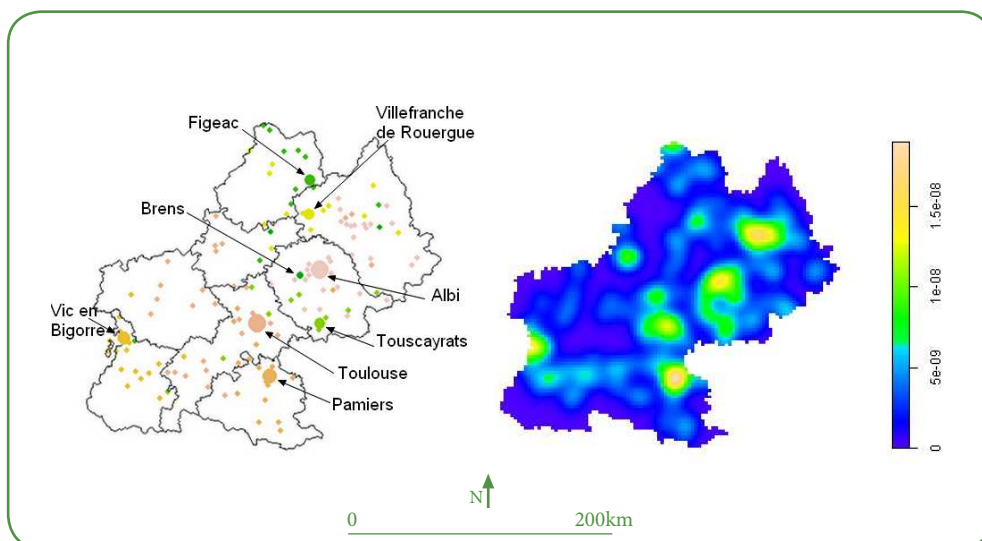
$$\forall s \in \Omega, f(s) = \frac{\lambda(s)}{\int_{\Omega} \lambda(s) ds}$$

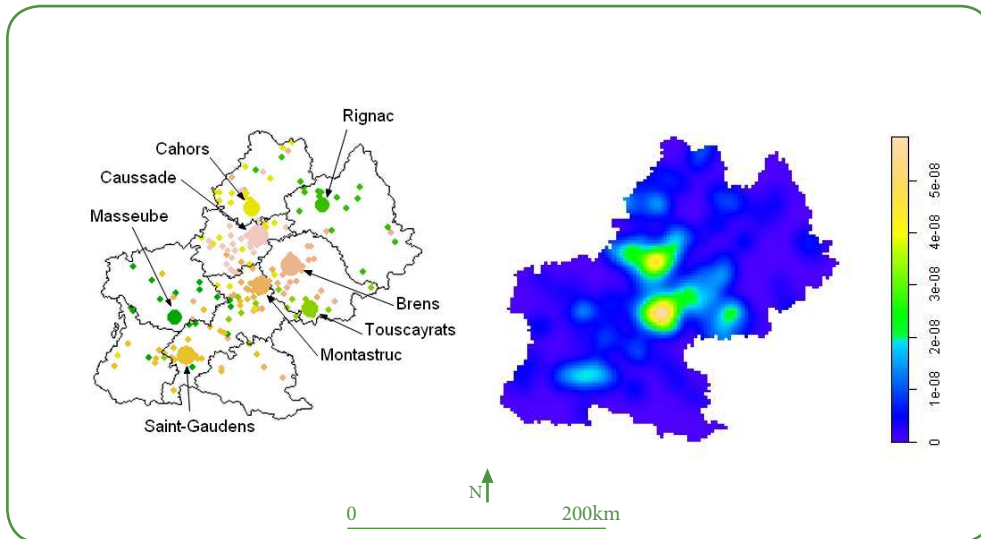
Indiquons ici que les processus ponctuels de Poisson appartiennent à deux familles de processus ponctuels : les processus ponctuels de Cox et les processus ponctuels de Markov. Ces deux familles possèdent des processus permettant de modéliser les interactions (agrégation ou répulsion) entre les points. Mais comme la prise en compte d'interactions n'avait pas de finalité ici, nous avons fait le choix de n'utiliser que le processus ponctuel de Poisson qui ne prend pas en compte les interactions (voir pour une application : Moller et Waagepetersen, 2004, Diggle *op. cit.*, ou Bonneu, 2007).

Pour notre problématique, nous supposons ainsi que le lieu de résidence des élèves est la réalisation d'un processus ponctuel de Poisson permettant de simuler les positions des élèves indépendamment les uns des autres¹⁶ à partir de l'intensité estimée de la population des élèves. L'intensité est estimée par une méthode non paramétrique à noyau (Diggle, 2003) qui permet de prendre en compte la répartition inhomogène (irrégulière) des élèves sur le territoire. Bien que le choix du paramètre de lissage soit en général primordial par rapport à celui de la fonction noyau, le choix d'une fonction à support infini donne un avantage considérable par rapport à l'approche discrète. En effet, l'intensité peut-être faible mais jamais nulle sur des communes où aucun élève n'a été localisé en 2007. Cette spécificité est importante car elle correspond bien à une réalité lorsque l'on considère l'évolution de la localisation des élèves. Les cartes suivantes (cartes 2 et 3) rendent compte de la localisation actuelle des élèves de 1ère année pour chacune des formations étudiées et de l'intensité estimée.

16. Le processus ponctuel de Poisson comprend une hypothèse d'indépendance. Cette hypothèse correspond à l'hypothèse vraisemblable que l'inscription d'un élève sur un site n'est pas influencée par celle d'autre élève proche de son domicile. En outre, cette hypothèse permet le choix d'un modèle plus simple qui évite des complications calculatoires importantes.

Carte 2
Sites et origines des
étudiants en 2007
pour le BTSA ACSE





Carte 3.
Sites et origines des
étudiants en 2007
pour le Bac Pro

Cette méthode nous permet ainsi de simuler de nouvelles populations d'élèves sur lesquelles sont évalués différents scénarios d'ouverture et de fermeture de classe comme exposé ci-après.

4. Méthodologie du choix d'ouverture et de fermeture de classes à partir des simulations

Différents scénarios d'ouverture et de fermeture sont évalués selon un critère de minimisation de la somme des distances-temps du domicile des élèves aux sites de formation. Cette somme des distances est calculée pour chaque génération aléatoire d'élèves par le processus ponctuel de Poisson estimé. Le calcul des distances-temps sur les données simulées nécessite une étape intermédiaire, entre la génération de la population d'élèves et le calcul de la fonction objectif de l'étape d'optimisation.

4.1. Estimation de la distance-temps d'une localisation simulée d'un élève

Les distances-temps ont été préalablement calculées entre toutes les communes de Midi-Pyrénées et les communes des sites de formation et sont stockées dans une matrice de taille 3019×43 , à partir d'un point central de chaque commune (la mairie). Or la localisation simulée d'un élève est identifiée par les coordonnées géographiques latitude et longitude, et non par un code de géo-référencement que serait la commune. Par conséquent, nous avons choisi de considérer la distance-temps entre la commune la plus proche « à vol d'oiseau » de la position géographique simulée de l'élève et la commune la plus proche disposant de la formation considérée. Cette approximation apparaît naturelle et se justifie aisément. Par exemple, pour une ville très étendue comme Toulouse, cette hypothèse correspond au fait qu'un élève en périphérie est souvent plus proche d'une commune de banlieue que du centre ville de sa commune de résidence.

4.2. Minimisation de la somme des distances et fermeture

A priori, chacune des 8 communes accueillant actuellement la formation BTSA est susceptible de subir une fermeture. Nous évaluons donc la somme des distances-temps des élèves aux sites, pour chaque cas possible de fermeture sur l'un des 8 sites accueillant actuellement la formation. Cette somme des distances est évaluée après avoir réaffecté chacun des élèves du site de fermeture au site le plus proche parmi les 7 sites restants. La DRAF nous ayant précisé qu'elle souhaitait que les établissements convergent vers des seuils d'effectifs par classe de 32 élèves au maximum,

nous n'avons pas affecté d'élève au site le plus proche si celui-ci avait déjà atteint, voire dépassé, ce seuil¹⁷. L'élève est donc affecté au site le plus proche respectant cette contrainte de seuil.

4.3. Minimisation de la somme des distances et ouverture

Dans le cas du choix d'une nouvelle ouverture pour la formation Bac Pro Services en Milieu Rural, le critère d'évaluation reste la somme des distances. Actuellement, cette formation est dispensée sur 8 communes. Le coût de création d'un nouveau site de formation étant très élevé, le choix de la DRAF est de privilégier des ouvertures sur les sites existants. Notons que l'ouverture pourrait aussi être faite sur un site bénéficiant déjà de cette formation, un site disposerait alors de deux classes. Mais dans un souci d'équité territoriale, la DRAF préfère privilégier une ouverture sur un nouveau site. Il s'agit donc de déterminer sur quelle commune parmi les 35 autres ouvrir une nouvelle classe. Pour cela, nous évaluons la somme des distances-temps des élèves aux sites d'affectation pour chaque scénario possible d'ouverture parmi les 35 sites candidats de la manière suivante.

L'ouverture de la formation sur un nouveau site entraîne nécessairement une réaffectation de l'ensemble des élèves sur les 9 sites (les 8 sites d'origine et le nouveau site). Lors de cette réaffectation le critère maximum de 32 élèves doit être respecté. En revanche, le seuil d'ouverture d'une nouvelle classe est fixé par la DRAF à 12 élèves. Il est donc nécessaire d'atteindre ce seuil pour valider le choix d'ouverture. Nous avons choisi de minimiser la somme des distances parmi les sites qui atteignent ce seuil de 12 élèves. Notons que pour cette formation nous avons observé des effectifs qui dépassent pour un grand nombre de sites le seuil des 32 élèves. Cette formation en pleine expansion a effectivement conduit les services administratifs à accepter ces sureffectifs dans l'attente d'une nouvelle ouverture. Lors de l'évaluation des différents scénarios, la DRAF peut alors privilégier le site d'ouverture qui à la fois minimise la somme des distances et équilibre mieux la répartition des effectifs sur les sites.

4.4. Significativité des solutions et classement des choix d'ouverture/fermeture

A l'issue de ces simulations, nous appliquons un test statistique pour déterminer si la fréquence d'apparition d'un site comme solution de fermeture ou d'ouverture est significativement différente du hasard. Le test requis est un test de proportion où la proportion est la fréquence d'apparition du site comme solution au problème parmi les n simulations effectuées. Nous supposons que la loi théorique suivie par cette proportion est une loi normale de moyenne p et d'écart-type $\sqrt{\frac{p(1-p)}{n}}$

avec n est la taille de l'échantillon observé qui correspond au nombre de simulations générées.

Dans le cas du choix d'une fermeture, nous vérifions que cette fréquence est significativement différente du choix purement aléatoire d'un « site à fermer » parmi les sites accueillant la formation considérée. Cette probabilité théorique de choix purement aléatoire est notée p_0 . Nous comptons 8 sites accueillant une formation de BTS ASCE, soit $p_0 = 0,125$: chaque site a 1 chance sur 8 d'être tiré au sort pour être le site de fermeture lors d'un tirage aléatoire.

Pour le problème d'ouverture, la probabilité théorique change. Il s'agit de déterminer sur quel site ouvrir une nouvelle formation de Bac Pro Service parmi les 35 sites n'accueillant pas la formation. Chacun de ces sites a une probabilité théorique p_0 de 0.028 (1/35) d'être tiré au sort comme « site à ouvrir ».

17. L'observation des élèves inscrits en 2007 montre que certains sites ne respectent pas ce seuil. Il n'y a en effet pas d'obligation pour l'établissement de respecter ce seuil qui est plus une norme définie par les autorités administratives pour assurer un enseignement de qualité. Dans la pratique, il s'avère que les établissements dépassant largement ce seuil divisent les élèves de la classe en groupes pour certaines matières, notamment des TP.

18. Concernant ce dernier point nous n'avons pas mené d'investigation sur comment choisir une solution en fonction de l'ensemble des rangs de la solution. Une réflexion sur l'utilisation d'une méthode d'évaluation de type scoring sera propice pour la poursuite des travaux.

19. <http://cran.r-project.org/>

Les fréquences d'apparition des sites comme solution au problème sont ensuite classés selon deux rangs : au rang 1 quand la solution est celle qui minimise le plus la somme des distances et au rang 2 pour la deuxième solution qui minimise, juste après la 1ère solution, la somme des distances. La DRAF peut alors choisir parmi les solutions ayant des fréquences d'apparition significatives les plus élevées au rang 1 ou à la fois aux rangs 1 et 2¹⁸.

4.5. Les résultats pour les deux problèmes étudiés

Suivant cette méthodologie, l'évaluation des différents scénarios d'ouverture et de fermeture a été réalisée sur la base de 100 populations d'élèves simulées à partir du modèle ajusté de processus ponctuel de Poisson. Les algorithmes de calcul créés ont été implémentés avec le logiciel R¹⁹ (Baddeley et Turner, 2006).

Solution pour la fermeture sur un site de la formation BTSA ACSE

La formation BTSA ACSE est dispensée sur 8 sites agricoles. Nous avons évalué chacun des 8 sites ; c'est-à-dire que nous avons supprimé chacun d'entre eux l'un après l'autre puis nous avons réaffecté les élèves aux sites et calculé la somme des distances temps. Le tableau 4 et le graphique 2 présentent les résultats obtenus après 100 simulations, en 1ère et 2nd rangs (i.e. pour chaque simulation nous avons relevé les 2 sites dont la fermeture génère les deux plus petites sommes des distances). Le test de proportion est appliqué pour chacune des fréquences d'apparition. Nous avons vérifié ces résultats sur la base de nouvelles simulations au nombre de 200 et les mêmes solutions ont été trouvées.

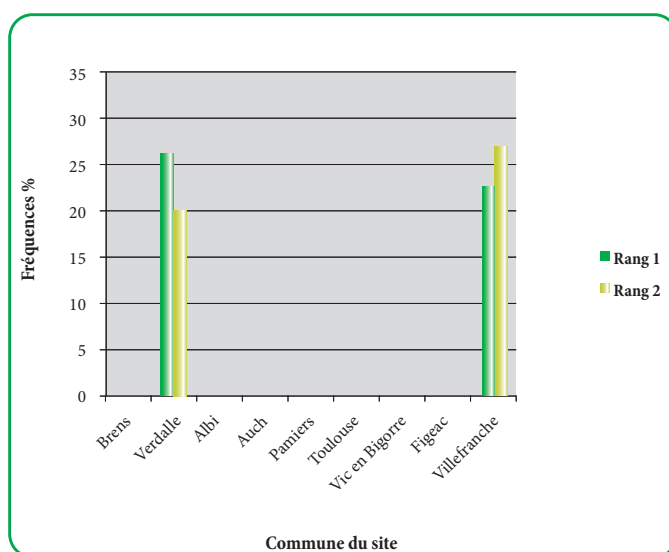
Commune du site de formation à fermer	Rang 1		Rang 2	
	Fréquence d'apparition	p-value	Fréquence d'apparition	p-value
Brens	8%	0.2265	8%	0.2265
Verdalle	26%	8.465e-05	20%	0.03129
Albi	13%	1	11%	0.7624
Pamiers	0%	2.851e-04	0%	2.851e-04
Toulouse	15%	0.5453	17%	0.2265
Vic en Bigorre	0%	2.851e-04	0%	2.851e-04
Figeac	15%	0.5453	17%	0.2265
Villefranche	23%	0.002497	27%	2.304e-05

Tableau 4. Fréquences d'apparition aux rangs 1 et 2 des communes des sites de formation pour la fermeture d'une classe BTSA ACSE (100 simulations).

On peut observer que les sites de Brens, d'Albi, de Toulouse et de Figeac ne ressortent pas comme des solutions significatives. Parmi les solutions significatives, la commune de Verdalle a un très fort pourcentage d'apparition au premier rang, avec un écart très faible avec celui de Villefranche. Il est intéressant de remarquer qu'au second rang les deux mêmes sites ont les pourcentages d'apparition les plus forts. En revanche au second rang la fréquence du site de Villefranche est supérieure à celle de Verdalle. *In fine*, ce sont donc ces deux sites qui apparaissent comme deux solutions de prédilection pour fermer la formation BTSA ACSE.

Il est aussi intéressant de relever que certains sites n'apparaissent jamais comme des sites « solutions » (communes de Pamiers et de Vic en Bigorre). Il s'agit ici de sites ayant une localisation géographique privilégiée aux extrémités du sud de la région. Ils disposent ainsi d'un large bassin de recrutement non concurrencé par d'autres sites.

Graphique 2.
Fréquences significatives
au seuil de 1%



Solution pour l'ouverture sur un site de la formation BAC PRO

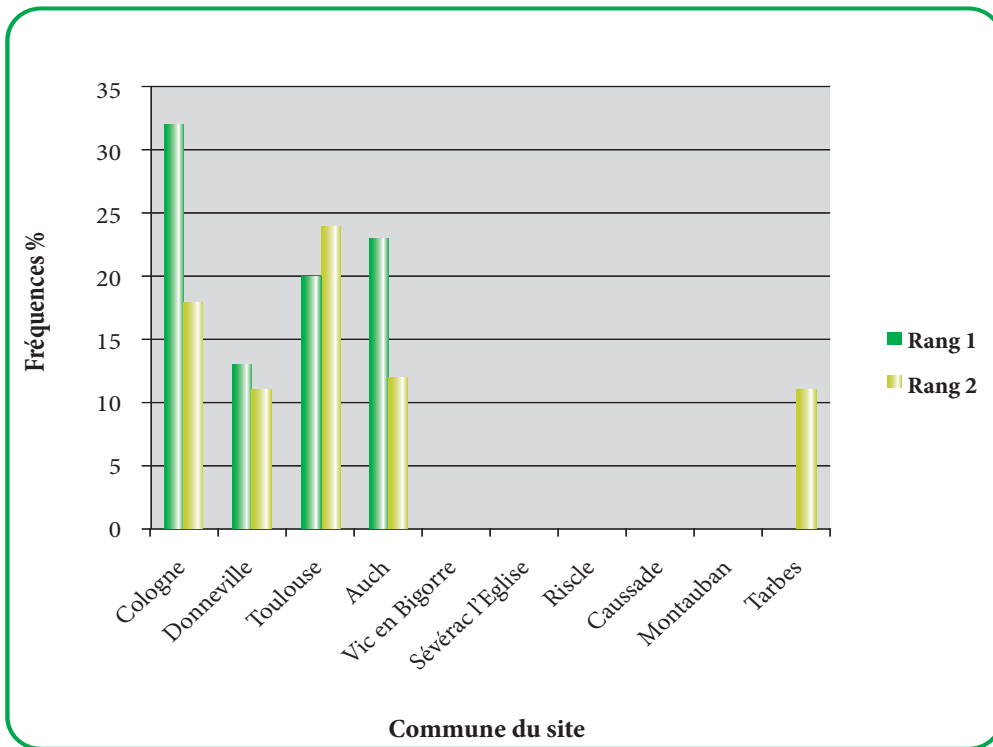
La formation BAC Pro Services est dispensée sur 8 sites. Nous avons évalué à tour de rôle les 35 sites restants pouvant accueillir la formation lorsqu'on réaffecte les élèves sur les 9 sites pour 100 simulations. A chaque simulation nous n'avons retenu que les cinq premiers sites minimisant la somme des distances-temps. Le tableau 5 présente les fréquences d'apparition aux rangs 1 et 2 des sites qui sont apparus au moins une fois parmi les 5 meilleures solutions lors des simulations. On constate globalement que pour le rang 1, ce sont 8 mêmes sites qui reviennent comme solution. Pour le rang 2, nous retrouvons ces sites et deux sites supplémentaires qui ne sont jamais apparus au rang 1. 200 nouvelles simulations ont également conduit aux mêmes résultats.

Tableau 5.
Fréquences
d'apparition aux
rangs 1 et 2 des
communes des sites
ouvrant une classe
BAC PRO SERVICE
pour 100 simulations

Commune du site de formation à fermer	Rang 1		Rang 2	
	Fréquence d'apparition	p-value	Fréquence d'apparition	p-value
Cologne	32%	< 2.2e-16	18%	0.2265
Donneville	13%	3.229e-14	1%	0.03129
Toulouse	20%	< 2.2e-16	24%	0.7624
Auch	23%	< 2.2e-16	12%	2.851e-04
Vic en Bigorre	2%	0.8303	6%	0.2265
Sévérac l'Eglise	6%	0.1127	7%	2.851e-04
Riscle	2%	0.8303	2%	0.2265
Caussade	2%	0.8303	4%	2.304e-05
Montauban			5%	0.3241
Tarbes			11%	4.484e-06

Certains sites n'apparaissent jamais comme des solutions significatives (Vic en Bigorre, Riscle, Caussade). Parmi les solutions significatives (graphique 3), les sites de Cologne et de Toulouse apparaissent comme les meilleures solutions avec des fréquences d'apparition élevées. Viennent ensuite les sites d'Auch et de Donneville.

Les quatre sites ayant les fréquences d'apparition les plus élevées forment géographiquement un triangle délimité par Cologne-Donneville-Auch incluant Toulouse. Ce triangle semble révéler la force d'attraction du poids démographique de l'agglomération toulousaine. Rappelons en effet que la population de la région Midi- Pyrénées est celle ayant le plus fort déséquilibre démographique territorial en France. La Haute-Garonne rassemble à elle seule 42% de la population régionale, contre 5 à 13% pour chacun des sept autres départements de la région. Avec plus d'un million d'habitants, l'aire urbaine de Toulouse représente à elle seule 37% de la population régionale. Dès lors, parmi les sites de prédilection d'ouverture de formation, il n'est pas étonnant de trouver comme solution le site de Toulouse et d'autres communes situées à proximité.



Graphique 3. Fréquences significatives au seuil de 1%.

Conclusion

Au-delà des résultats obtenus qui sont spécifiques au problème posé, nous souhaitons insister sur la finalité de ces travaux. Ces outils restent à nos yeux des « aides à la décision » et ne sont pas substituables à une décision d'acteurs. Seuls ces derniers peuvent en effet prendre en compte d'autres critères moins « objectivables » ou « quantifiables » dans les techniques quantitatives que nous employons. L'appréciation du paysage des sites d'enseignement agricole par l'ensemble des parties prenantes du monde éducatif reste cruciale car l'outil ne sert à éclairer qu'une dimension de cette réalité. Les résultats obtenus n'ont donc pas pour vocation de donner « la » solution au problème, mais d'éclairer une dimension pouvant aider les services de l'Etat dans leur décision d'ouvrir ou de fermer telle ou telle classe. En l'occurrence, nous avons choisi ici de mettre en avant un critère de distance, plus précisément de temps parcours, qui constitue un déterminant majeur du choix d'un site de formation par les élèves et leur famille.

La poursuite de ce travail pourra donc envisager d'intégrer d'autres critères afin d'améliorer la prise en compte de cette « réalité » dans le choix de solutions pour les futures ouvertures et fermetures. Parmi eux nous pouvons citer : la présence ou non d'un internat, la proximité du site à une station de transports en commun, les logiques de parcours (le site permet de poursuivre une formation d'un niveau supérieur à l'issue de la formation actuelle), la localisation des enseignants (la fermeture d'une formation suppose la réaffectation de l'enseignant sur un autre site). Tous ces critères sont susceptibles de modifier les présentes solutions, ou au contraire, de les renforcer. Pour cela il convient de réfléchir à la construction d'indicateurs qui puissent être intégrés aux algorithmes générés et d'envisager un système de pondération dans l'appréciation de l'importance de chacun d'entre eux. Ces travaux d'évaluation multi-critères sont, bien entendu, plus complexes à mener, et surtout, ils exigent des recueils de données supplémentaires qui nécessitent de mener des enquêtes sur les déterminants des choix des élèves.

Un intérêt majeur de ce travail réside dans le recours à la méthode des Processus Ponctuels pour prendre en compte le caractère aléatoire des données. D'autres améliorations peuvent être apportées dans ces travaux. Une première voie d'amélioration réside dans l'appréciation de la répartition des effectifs sur les sites au travers par exemple de l'usage des indices de Gini ou d'indice d'entropie. Une deuxième voie d'amélioration consisterait à utiliser un processus ponctuels marqué. Dans ce cas, à chaque localisation peut-être affectée un ensemble de variables de différents types (réelles, entières, booléennes...). Une amélioration serait peut-être d'affecter à chaque élève des caractéristiques comme les vœux formulés, la localisation du lieu de travail des parents ou la volonté d'être en internat. Ces éléments supplémentaires nous amèneraient alors à considérer l'optimisation d'une fonction plus réaliste que celle basée sur les seules distances-temps.

Enfin, une autre voie peut être envisagée dans le couplage de ces techniques de simulation à d'autres méthodes de résolution comme l'AGMC – Algorithmes Génétiques Multi-Critères – développés par l'IRIT (Institut de Recherche en Informatique de Toulouse), dont on peut trouver un exemple d'application dans Berro et Leroux (2006). Ces techniques par apprentissage permettent d'explorer des espaces de solution beaucoup plus vastes, notamment dans le cadre

de la résolution de problème de type « réaffectation d'une formation sur l'ensemble des sites ». Le nombre de scénarios possibles devient en effet trop important pour poursuivre avec les méthodes d'optimisation employées ici.

Les perspectives de travaux futures sont donc nombreuses et ces premiers résultats n'apparaissent ainsi que comme les premiers prémices dans un domaine de recherche très peu investi jusqu'à présent.

Bibliographie

ALBERADA-SAMBOLA M., FERNANDEZ E., LAPORTE G., 2009, "The capacity and distance constrained plant location problem", *Computers & Operations Research*, 36, 597-611.

ARRIGHI J-J., 2004, "Les jeunes dans l'espace rural : une entrée précoce sur le marché du travail ou une migration probable", *Formation et Emploi*, 87, 63-88.

BADDELEY A., TURNER R., 2006, "Modeling spatial point patterns in R", *Case studies in spatial point process modeling*, 185, 23-74.

BECKER G., 1964, *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, University of Chicago Press.

BERRO A., LEROUX I., 2006, "A simulation of strategic bargaining within biotechnology clusters", in *Handbook of Research on Nature Inspired Computing for Economics and Management*, Eds J-P Rennard (Idea Group, XXIII) vol. I, 335-351.

BONNEU F., THOMAS-AGNAN C., 2008, "Spatial point process models for locationallocation problems", *Computational Statistics & Data Analysis*, 53, 3070-3081.

BONNEU F., 2007, "Exploring and Modeling Fire Department Emergencies with a Spatio-Temporal Marked Point Process", *The Case Studies in Business, Industry and Government Statistics (CSBIGS) Journal*, 1(2), 139-152.

CAHUZAC E., TAHAR G., 2000, "Offre scolaire et arbitrage formation-distance : le cas des BEP en Midi-Pyrénées", *Les Notes du Lirhe*, 311.

COMBES P-P., LAFOURCADE M., 2005, "Transport Costs: Measures, Determinants and Regional Policy Implications for France", *Journal of Economic Geography*, 5(3), 319-349.

DIGGLE P. J., 2003, *Statistical analysis of spatial point patterns*, Oxford University Press, 2nd edition.

FÉRAT F., 2006 "L'enseignement agricole une chance pour l'avenir des jeunes et des territoires", *Rapport d'information du Sénat n°27* http://www.senat.fr/rap/r06-027/r06-027_mono.html.

FOURCADE B., OURLIAC G., 1983, "Localisation des formations, recrutement et insertion des élèves: un essai de mise en relation, le cas des formations techniques courtes en Midi-Pyrénées", *Economie Méridionale*, 123.

GRELET Y., 2006, "Des territoires qui façonnent les parcours scolaires des jeunes", *Bref*, 228, 1-4.

GRELET Y., 2004, "La reproduction sociale s'inscrit dans le territoire", *Formation et Emploi*, 87, 79-98.

HASTINGS J., KANE T., STAIGER D., 2005, "Parental preferences and school competition, evidence from a public choice program" Technical Report, National Bureau of Economic Research.

HILAL M., 2007, "Temps d'accès aux équipements au sein des bassins de vie des bourgs et petites villes", *Economie et Statistique*, 402, 41-56.

HOTELLING H., 1939, "The General Welfare in relation to Problems of Taxation and of Railway and Utility Rates", *Econometrica*, 242-269.

HOXBY C., 2003, *The Economics of School Choice*, The University of Chicago Press.

LARCENEUX A., 1997, "Mécanisme de décision et contexte local : l'offre de formation professionnelle initiale", *Formation et Emploi*, 59, 57-70.

LEMISTRE P., MAGRINI M.-B., 2011, "Job Qualification, Distance between Towns and Geographic Relocation for French Youth", *Urban Studies*, Volume 48, Issue 10.

MANSKI C., WISE D., 1983, *College choice in America*, Harvard University Press.

MOLLER J., WAAGEPETERSON R., 2004, "Statistical inference and simulation for spatial point processes", 100, Chapman & Hall/CRC.

MÜLLER S., TSCHARAKTSCHIEW S., HAASE K., 2008, "Travel-to-school mode choice modelling and patterns of school choice in urban areas", *Journal of Transport Geography*, 16, 342-357.

Sites de formation	Secteur	Nombre d'élèves de la région	Part d'internes	Taux d'élèves sur le site le plus proche pour cette formation		Nombre d'élèves non originaire de la région
				D.P. ou externe	Interne	
Pamiers	Public	20	10%	78%	100%	5
Villefranche de Rouergue	Privé	14	79%	100%	54%	1
Toulouse	Public	31	26%	35%	37%	6
Figeac	Public	12	0%	58%	–	6
Vic en Bigorre	Public	18	0%	91%	–	2
Albi	Public	34	15%	45%	40%	2
Brens	Privé	5	20%	0%	0%	1
Verdalle	Privé	13	0%	38%	–	8
Total		147	18%	56%	48%	31

Source : DRAF

D.P. : Demi-pensionnaire

Tableau 1. Effectifs de BTSA ACSE de 1ère année (2007)

Sites de formation	Secteur	Nombre d'élèves de la région	Part d'internes	Taux d'élèves sur le site le plus proche pour cette formation		Nombre d'élèves non originaire de la région
				D.P. ou externe	Interne	
Pamiers	Public	18	5%	82%	100%	7
Villefranche de Rouergue	Privé	15	80%	100%	75%	1
Toulouse	Public	13	8%	50%	0%	4
Figeac	Public	5	0%	80%	–	2
Vic en Bigorre	Public	10	0%	100%	–	2
Albi	Public	26	0%	46%	–	0
Brens	Privé	3	100%	–	35%	8
Verdalle	Privé	15	0%	60%	–	3
Total		105	16%	66%	65%	27

Source : DRAF

D.P. : Demi-pensionnaire

Tableau 2. Effectifs de BTSA ACSE de 2^e année (2007)

Sites de formation	Secteur	Nombre d'élèves de la région	Part d'internes	Taux d'élèves sur le site le plus proche pour cette formation		Nombre d'élèves non originaire de la région
				D.P. ou externe	Interne	
Rignac	Privé	21	76%	80%	69%	0
Montastruc	Privé	37	22%	83%	62%	1
Saint-Gaudens	Public	32	44%	83%	57%	0
Masseube	Privé	19	68%	100%	70%	1
Cahors	Public	30	80%	83%	37%	2
Brens	Privé	43	63%	75%	22%	1
Verdalle	Privé	25	60%	80%	40%	1
Caussade	Privé	44	25%	90%	55%	0
Total		251	51%	85%	47%	6

Source : DRAF

D.P. : Demi-pensionnaire

Tableau 3. Effectifs de Bac Pro Services en Milieu Rural de 1ère année (2007)

Sites de formation	Secteur	Nombre d'élèves de la région	Part d'internes	Taux d'élèves sur le site le plus proche pour cette formation		Nombre d'élèves non originaire de la région
				D.P. ou externe	Interne	
Rignac	Privé	17	29%	58%	100%	1
Montastruc	Privé	26	15%	82%	25%	0
Saint-Gaudens	Public	25	52%	92%	31%	0
Masseube	Privé	16	37%	70%	33%	0
Cahors	Public	29	62%	91%	44%	1
Brens	Privé	60	55%	81%	36%	3
Verdalle	Privé	23	26%	100%	83%	2
Caussade	Privé	43	20%	91%	22%	0
Total		239	39%	85%	42%	7

Source : DRAF

D.P. : Demi-pensionnaire

Tableau 4. Effectifs de Bac Pro Services en Milieu Rural de 2^e année (2007)





Dans le cadre du Programme de recherche Pour et Sur le Développement Régional (PSDR3) en Midi-Pyrénées, le Projet COMPTER bénéficie du soutien financier de :



Plus d'informations : www.inra.fr/psdr-midi-pyrenees

RÉSUMÉ

Ce Document d'analyse PSDR s'intéresse au choix de fermeture/ouverture de classes, à partir d'un critère de distance peu employé : le temps de trajet entre le domicile de l'élève et son école. L'originalité de l'étude repose sur l'utilisation d'un modèle mathématique prenant en compte la variabilité stochastique de la position géographique des élèves, pour ainsi fournir une indication de la variabilité des positions optimales du lieu d'ouverture / fermeture de classe minimisant le critère distancetemps. La résolution de ces problèmes est proposée pour deux cas de formation de l'enseignement agricole français.

MOTS-CLÉS – Distance, Processus Ponctuel, Simulation, Localisation Optimale, Déplacements domicile-école, Enseignement Agricole

SUMMARY – Spatial Applied Simulation and Optimization Methods on schools' creating or closing : a case study for the French Agricultural Education.

This document deals with the problem of creation or closing of school classes from the perspective of a distance indicator seldom used: the time needed to travel by road from the family home to the training establishment. The originality of this work lies in the use of a mathematical model which takes into account the randomness of students' locations to provide an indication of the variability of the optimal positions for the classes' creation/closing solutions based on a minimization of the distance-time. These problems are solved for two case studies of agricultural education in France.

KEY-WORDS – Distance, Point Process, Simulation, Optimal Location, Home-to-School Commuting, Agricultural Education