

# Modélisation et Calibration Ajustement du Urban Growth Moland Model à l'espace transfrontalier franco-italo-monégasque

## 3.1. Détermination de l'aire de la modélisation

En 2000, l'aire d'étude globale soumise à la modélisation est égale à 1 région d'étude, 3 Etats, 183 communes et 558 114 cellules réparties en 19 classes d'occupation du sol (cf. figure.3.1). La délimitation de l'aire réelle de la modélisation (cf. figure 3.2) est basée sur les territoires urbains susceptibles d'abriter les futures gares TGV à l'horizon 2020. Ces villes sont Cannes et Nice du côté français et Vintimille du côté italien (dans une perspective de desserte du projet de transport LGV PACA de la frontière italienne). Toutefois, l'aire réelle de la modélisation a été volontairement étendue aux communes voisines de Cannes, Nice et de Vintimille, dans la perspective de pouvoir détecter l'effet de diffusion des phénomènes susceptibles d'émerger au lendemain de la mise en service de la LGV PACA.

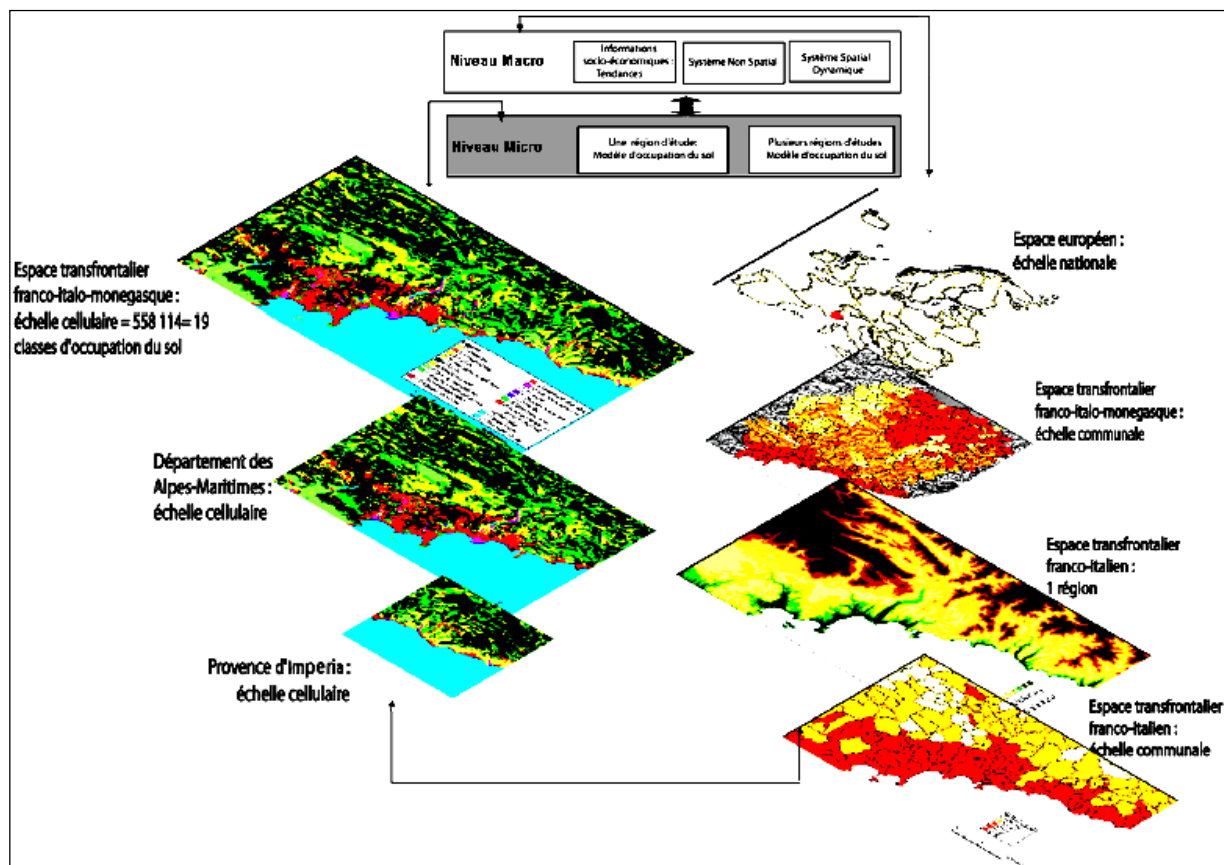


Figure 3.1 : Différents niveaux territoriaux composant l'aire d'étude

Avec 183 communes au total réparties sur une superficie d'environ 4262,64 km<sup>2</sup> et une population totale en 1999 de 1 190 819 habitants, l'aire de la modélisation regroupe la totalité des communes de la « bande côtière », c'est-à-dire du littoral et du moyen pays, soit environ 133 communes sur les 163 que compte le département des Alpes-Maritimes et 32 sur

les 67 que compte la province d'Imperia. Dix-huit petites communes de montagne du département du Var frontalières des Alpes-Maritimes, viennent compléter l'aire d'étude. Ces petites communes du Var jouxtant l'agglomération cannoise enregistrent en 1999 des densités de population faibles, variant entre 1,47 habitants au km<sup>2</sup> pour la commune de Val-de-Chalvagne et 153 habitants au km<sup>2</sup> pour la commune de Fayence. Ces communes représentent moins de 1% des 898 441 habitants que compte la population du département du Var en 1999.

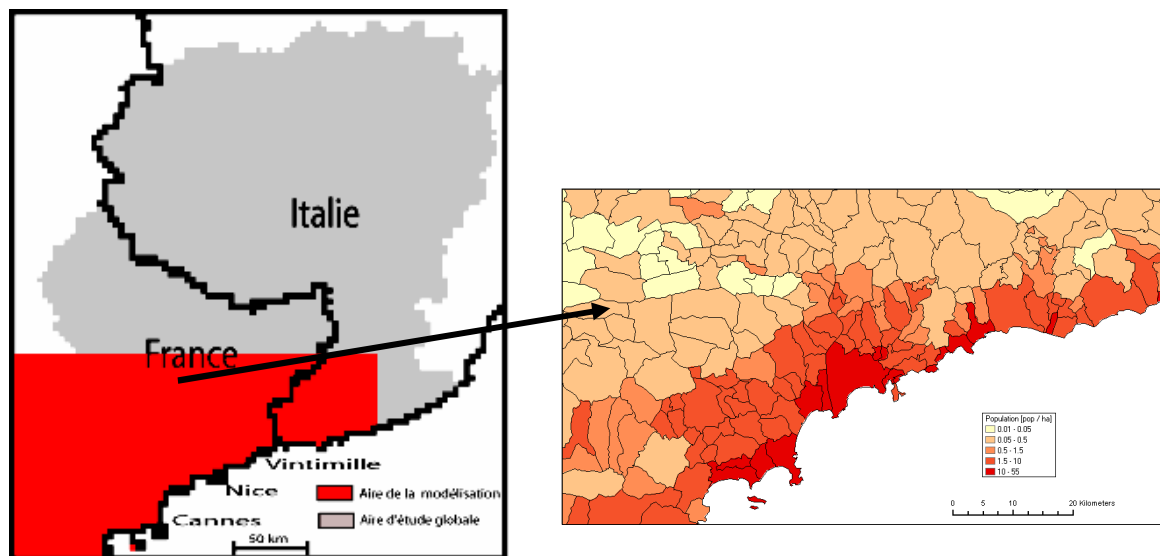


Figure 3.2. : Aire réelle de la modélisation en 2000 : Alpes-Maritimes, Principauté de Monaco et San Remo dans la Province d'Imperia

Source : EUROSTAT- GISCO, (base de données mise à jour 2008)

## 3.2. La construction du modèle

### 3.2.1. Des contraintes d'ordre méthodologique et/ou technique

Il existe quatre contraintes d'ordre méthodologique qu'il faut prendre en compte pour assurer le bon fonctionnement du modèle. Il s'agit :

(1) de la question de l'harmonisation des données d'entrée du modèle (inputs). En effet, pour garantir le bon fonctionnement du modèle, toutes les données d'entrée doivent au moins être disponibles pour plusieurs dates. Le modèle exige une année de base 1 (année historique de référence représentant le passé) et une année de base 2 (année représentant le présent, la réalité actuelle). De préférence, 10 à 15 années doivent séparer ces deux années de référence afin de mieux évaluer l'évolution du système étudié.

(2) Les données doivent être strictement comparables. Ce qui constitue une difficulté majeure dans un contexte transfrontalier, où la question de la disponibilité et de l'harmonisation des données est plus qu'ailleurs une limite. Ceci explique donc d'exclure certaines informations comme les « zoning statuts ». Les « zoning statuts » sont les politiques

qui visent à réguler l'aménagement urbain des villes. En France, elles existent sous forme de PLU, de DTA, de SCOT, etc. En Italie, elles sont connues sous le nom de *piani regolatori ou du PTCP (Piano Territoriale di Coordinazione del Paesaggio)*. Les « zoning statuts » sont l'exemple type de données non comparables et non exploitables directement car, selon le pays, les objectifs fixés sont différents et l'échelle de réalisation des documents n'est pas la même. À ce titre, en France, les PLU sont réalisés à l'échelle des villes/quartiers (UMR ESPACE, 2004) et en Italie, les informations les plus exhaustives et certainement les plus « fiables » se situent plutôt à l'échelle des provinces. Mais plus encore, les programmes de régulation (documents d'aménagements divers) s'ils sont logiques (parce que s'inscrivant dans un contexte de protection de l'environnement et du développement durable), sont loin d'être efficaces car leur application reste limitée. Souvent en effet, nombre de ces documents d'aménagement sont en total décalage avec la réalité. Une raison suffisante pour ne pas introduire un document d'aménagement, existant certes, mais inopérant dans le modèle au risque de rendre encore plus incertains les résultats des différentes prévisions futures. Aussi, au moment d'intégrer ce type de document dans le processus de la simulation, le modélisateur se doit de se demander si la réalité de l'espace est effectivement traduite par les documents d'aménagement en tant que tels. Le facteur "zoning statut" représenté par la fonction  ${}^tZ_{k,x,y}$  de la règle de transition (cf. l'Équation 3) quand il n'est pas intégré dans le processus de calibration du modèle est égal à 1. En revanche, quand il est une composante des entrées principales du modèle, il est calé avec chacune des différentes classes d'occupation du sol.

(3) Les données d'entrée doivent être identiques en taille (c'est-à-dire couvrir la même aire/échelle spatiale), en résolution et en origine/source. Autrement dit, elles doivent être enregistrées dans une base de données commune et couvrir complètement l'ensemble de la région étudiée. Pour répondre à cette exigence du modèle, nous avons mobilisé la base de données de *Corine Land Cover* pour l'occupation du sol, la base de données *GISCO*, pour les informations relatives à la population par exemple, la base de données *Télé-Atlas Europe* pour les informations relatives au réseau de transport, et la base de données *Digital Elevation Model (DEM)* ou Modèle Numérique de Terrain (MNT) qui représente la topographie d'une surface terrestre donnée. Un DEM peut couvrir différentes échelles spatiales, depuis la couverture globale de la plus petite échelle, jusqu'à la couverture locale à la plus grande échelle. Le DEM formule de façon quantitative les caractéristiques physiques d'un territoire en fournissant des informations détaillées et spatialisées permettant de visualiser le profil topographique du terrain étudié.

Les différentes bases de données mobilisées lors de l'étape de la calibration du modèle ont en commun d'être harmonisées de la plus petite à la plus grande échelle. Elles ont pour avantage de couvrir toute l'Europe et donc de répondre aux exigences d'une thématique transfrontalière.

(4) La dernière exigence du modèle porte sur les données d'usage des sols qui doivent être modélisées au niveau cellulaire (à l'échelle micro du modèle).

### 3.2.2. *Les inputs du modèle*

La construction des données d'entrées du modèle peut être déclinée en quatre grandes étapes. Dans une première étape, nous avons pris en considération les quatre contraintes majeures du modèle décrites dans la section précédente (cf. section, 3.2.1). Dans une seconde étape, nous nous sommes appuyée fortement sur les résultats issus du retour d'expérience du TGV (cf. partie 1 chapitre 3) et du diagnostic spatial transfrontalier (cf. partie 1 chapitre 2) qui a montré l'importance des contraintes physiques, les fortes relations spatiales entre la France, la Principauté de Monaco et l'Italie en dépit de la présence de la frontière, ainsi que les enjeux de l'urbanisation et de l'accessibilité de cette région transfrontalière, préfigurant les hypothèses suivantes : les effets du TGV sur la croissance urbaine et sur l'accessibilité du territoire à partir de 2020. Dans une troisième étape, nous avons procédé à une collecte de l'information disponible. Cette phase est complexe car elle nous amène à résoudre différentes questions : les questions relatives à la disponibilité des données et plus largement les questions relatives à l'harmonisation des informations entre la France et l'Italie, les questions afférentes à l'échelle et à la taille des cellules, mais aussi la question de la crédibilité des informations concernant le tourisme et les réglementations urbaines. Sur la base des trois premières étapes, la quatrième et dernière étape porte sur la validation de la fiabilité des entrées définitives du modèle, phase primordiale à prendre en compte pour l'assurance d'une bonne calibration. Quatre facteurs principaux constituent les entrées du modèle. Le premier facteur est l'occupation du sol (basée sur la classification de Corine Land Cover) à différentes dates. C'est au niveau de l'occupation du sol que se situent les vraies interactions spatiales et par conséquent donc, que se situent les effets de voisinages entre différentes catégories de classes d'occupation du sol. Le second facteur est le facteur *suitability* (ou adaptabilité du territoire) calculé au moyen du *Digital Elevation Model* (DEM) à partir d'une résolution de 100 m. De cette résolution est produite l'inclinaison des pentes (the steepness of slope) en pourcentage. Pour obtenir ensuite ce que l'on nomme les « *suitability maps* », les données relatives à l'aspect physique du territoire sont croisées avec les données d'occupation du sol. En effet, le facteur *suitability* est la probabilité d'apparition des cellules, en fonction de la pente de chaque cellule/classe d'occupation du sol dans l'ensemble de l'espace cellulaire. C'est ce facteur qui va permettre d'évaluer l'impact du déterminisme physique (l'inertie spatiale en quelque sorte) sur l'évolution de la croissance urbaine. Le troisième facteur est le réseau de transport qui va nous permettre d'évaluer l'accessibilité (cf. figure 3.3.). Le dernier facteur est le paramètre de perturbation aléatoire  $\alpha$ . À travers ce facteur, c'est une « chaîne de hasard » qui est introduite dans le modèle car, un modèle ne pouvant tout prévoir, les interactions et rétroactions entre éléments du système, les « entrelacs » des causes et des effets produits par l'effet de voisinage, vont être la source de phénomènes spatiaux émergents, totalement inattendus.



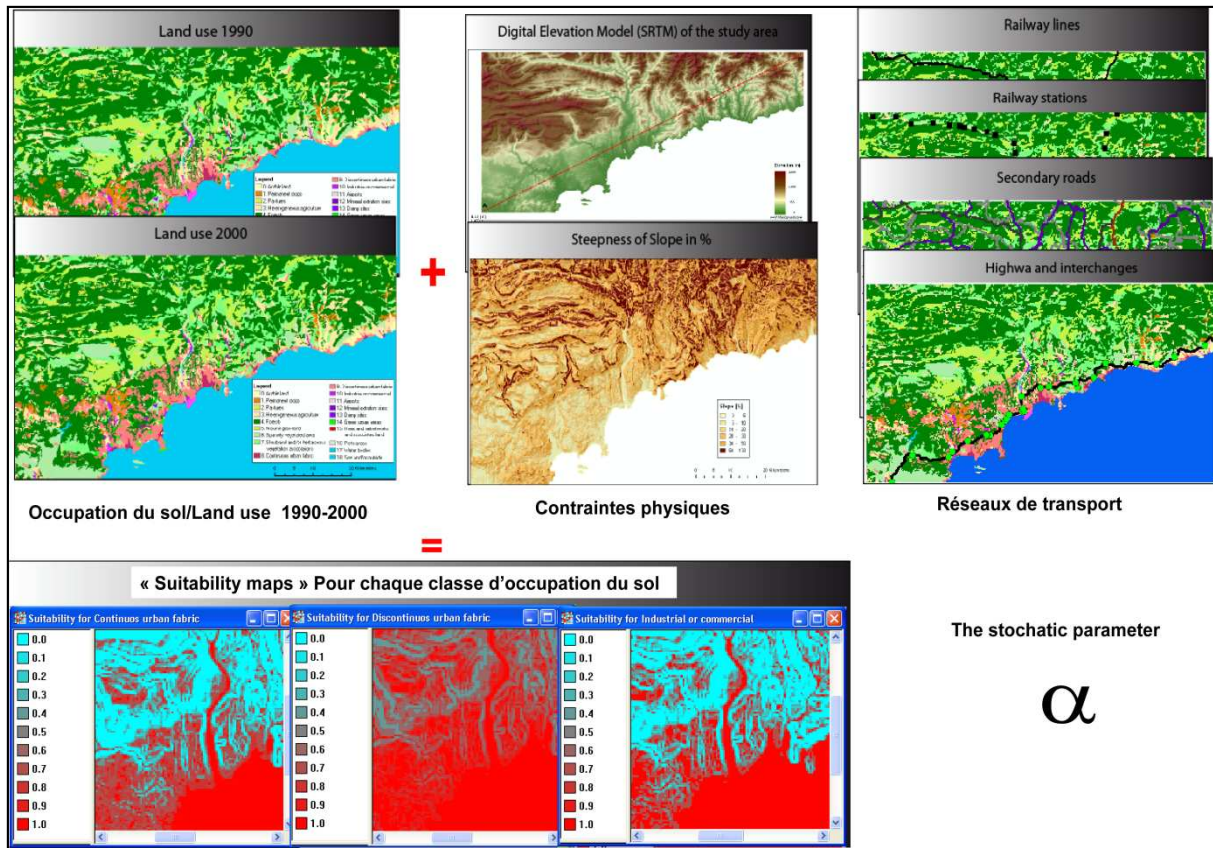


Figure 3.3 : Les inputs du modèle

Nous l'avons donc remarqué, le nombre d'inputs est volontairement limité. Seulement quatre ont été sélectionnés. Cette démarche est peu commune mais s'explique par le fait que nous réfutons l'idée selon laquelle une bonne modélisation/simulation est celle qui multiplie les entrées dans un modèle. Il en va tout autrement, et les modèles macro-économiques en sont l'illustration type. Plus les entrées seront multipliées, plus les résultats de la modélisation seront flous, et il devient plus compliqué, au terme de la simulation, d'attribuer à un facteur plus qu'à un autre, les changements survenus dans le système étudié. Mais en plus, ce choix d'entrées limitées se justifie dans un contexte de recherche ayant pour support l'infrastructure grande vitesse ferroviaire où il est déjà difficile d'isoler les facteurs socio-économique et sociodémographique impactés de façon certaine par l'arrivée du TGV (cf. partie 1 chapitre 3). Mais aussi, avec des hypothèses de recherche telles que la croissance urbaine et l'accessibilité, il convient d'avoir pour entrées des facteurs sûrs, quantifiables et mesurables, suffisamment représentatifs de la réalité des territoires afin de détecter avec précision là où les changements devraient se produire. Cependant, il n'est pas dit ici que nous écartons toutes les études apportant des éclaircissements sur les possibles retombées socio-économiques du TGV, à savoir des retombées sur l'emploi ou encore sur le tourisme, etc. Il s'agit seulement de prendre du recul avec ce type de recherches dans la phase de modélisation. Nous pensons en effet que ces dernières s'avèrent plus utiles et plus intéressantes quand elles interviennent en aval de la modélisation, autrement dit quand elles sont comparées avec les outputs. Notons

que tous les inputs sont construits dans un premier temps à l'extérieur du modèle, plus précisément dans un système d'Information Géographique (ARC-GIS en l'occurrence) avant d'être introduits définitivement dans le Moland pour être ajustés au modèle. Couplés généralement avec les modèles basés sur les automates cellulaires, les SIG contribuent grandement au réalisme de ces modèles de simulation. L'interaction entre un SIG et Moland est presque « naturelle » car, l'espace cellulaire de l'automate cellulaire est structurellement non différent de la représentation cellulaire de l'espace dans un SIG (Clarke and Gaydos, 1998). Ils forment tous deux essentiellement une grille de cellules partitionnant une zone géographique. Cette similarité permet une liaison facile entre le SIG et Moland d'un point de vue conceptuel et technique en s'assurant que les systèmes de référence à la grille du modèle et ceux du SIG correspondent et que, la relation une à une entre les cellules des deux outils est établie.

### 3.2.3. Calibration générale du modèle

Il s'agit dans cette section, de décrire les procédures de calage du modèle Moland. Moland, par son comportement complexe, requiert des procédures heuristiques difficilement disponibles et un temps nécessairement plus long pour être calé et exécuté que d'autres modèles de la même famille. En effet, contrairement à une technique purement statistique ou algorithmique, le modèle sollicite fortement « l'intuition » du scientifique réalisant le calage (Straatman *et al.*, 2004). Ce qui peut sembler suggérer que le calage en soi n'est pas « scientifique » et que le modèle calibré est sans valeur crédible, or c'est tout le contraire, l'appel à l'intuition est l'indication même que la structure du modèle à l'échelle micro est le miroir de la complexité des systèmes, et que seule une maîtrise du fonctionnement de l'espace étudié peut garantir les résultats d'une bonne calibration. La calibration du modèle s'effectue à partir des quatre grandes entrées. D'abord, le calage des 19 classes d'occupation du sol, autrement dit, l'effet de voisinage de l'espace cellulaire. Ensuite, le calage du facteur « *Suitability* », et pour finir, celui du facteur accessibilité et de la perturbation aléatoire.

#### 3.2.3.1. Ajustement de l'espace cellulaire

Caler l'effet de voisinage de l'espace cellulaire du modèle Moland (*the neighbourhood space effect*), revient à définir les interactions et interconnexions spatiales entre différentes cellules. En des termes plus concrets, il s'agit de déterminer les échanges entre cellules ainsi que l'importance de la dépense spatiale entre chaque cellule/lieu/pixel. Le modèle Moland, comme l'essentiel des modèles basés sur les automates cellulaires, obéit à l'hypothèse selon laquelle tout lieu est fortement déterminé par une certaine forme de dépendance (Tobler, 1970,1979). En effet, en fonction de la localisation géographique de chaque cellule et de la similitude ou non de ses cellules voisines dans l'environnement de l'espace cellulaire du modèle, est déterminé l'effet de répulsion ( $W_{K,L,C} < 0$ ) ou d'attraction (

$w_{K,L,C} > 0$ ) (cf. équation 3) qui s'exerce entre chaque classe d'occupation du sol. La distance définie par la fonction  $w_{K,L,C}$  de la règle de transition (cf. équation 3) exprime les contraintes globales qui s'exercent sur les relations entre une cellule appartenant à une classe d'occupation du sol donnée, et les autres cellules voisines qui peuvent être de la même classe d'occupation du sol qu'elle ou différentes d'elle (cf. la figure 3.4). Pour exemple, une cellule de la classe "industrie et commerce" résistera généralement mieux à la pression d'une cellule voisine de la classe "urbain discontinu" pour la simple raison que pour rester attractive, les zones urbaines ont besoin que les activités industrielles et commerciales se forment à leurs proximités pour assurer aux populations l'accès aux opportunités économiques. En revanche, une cellule appartenant à la classe "agriculture" résistera difficilement à la pression des cellules urbaines car la croissance urbaine s'effectue généralement sur d'anciens espaces agricoles considérés par les cellules de la classe "urbain discontinu" comme des zones très attractives (espace disponible) donc propices à leur expansion. D'une manière générale donc, le calage de l'effet de voisinage est basé sur une procédure interactive dans laquelle chaque cellule est mise en relation avec les autres.

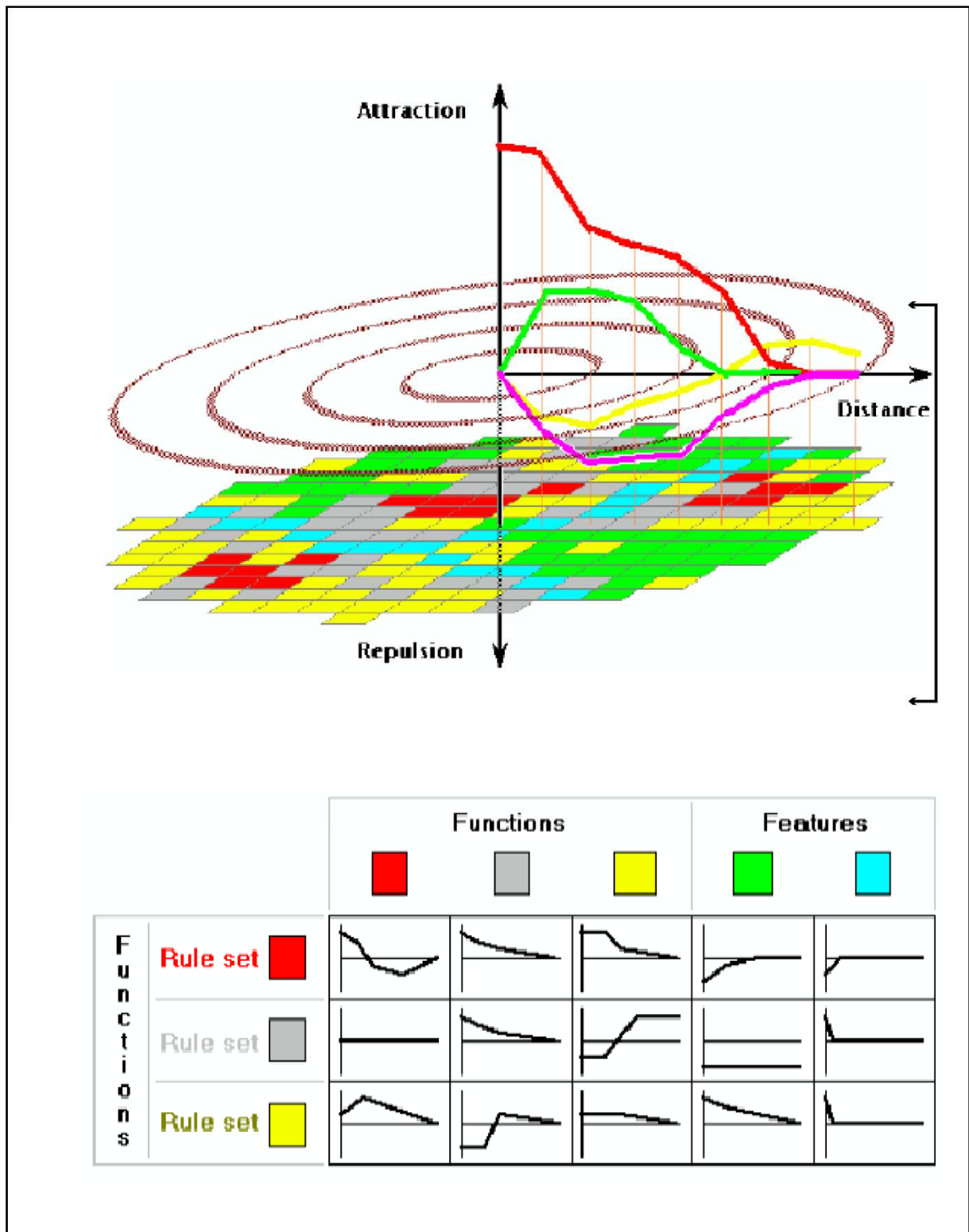


Figure 3.4 : L'environnement du voisinage cellulaire de Moland :  
la détermination des effets d'attraction et de répulsion

Source : White et Engelen, 2000

Le calage se fait donc en mode pair afin de déterminer s'il y a attraction et/ou répulsion entre les types d'occupation du sol. Une fois que toutes les cellules sont ajustées en



fonction des relations qu'elles entretiennent (attraction ou répulsion) avec les cellules voisines et avec elles-mêmes parce qu'on admet l'idée qu'un lieu puisse s'influencer lui-même, au même titre que son voisinage d'influence le modèle est exécuté plusieurs fois dans le but d'assurer sa robustesse.

### 3.2.3.2. Ajustement du facteur Suitability

Caler le facteur suitability  $tS_{k,x,y}$  (cf. équation 3) au modèle, c'est introduire en quelque sorte de l'inertie spatiale dans le modèle afin de pouvoir analyser dans quelle mesure celui-ci prend part dans les dynamiques urbaines. En effet, dans le cas de notre zone d'étude où l'on dénote de très fortes inclinaisons des pentes, ce facteur se révèle être très important dans le cadre de l'analyse de son évolution. Pour être calées au modèle, les « suitability maps » sont normalisées à des valeurs situées entre 0 et 1, représentant la probabilité d'apparition d'une cellule à un lieu/endroit plus qu'à un autre. Plus cette valeur est proche de 0, plus la probabilité d'apparition de la cellule dans l'espace cellulaire sera nulle et plus elle sera proche de 1, plus elle aura d'opportunité d'apparaître et d'évoluer dans le voisinage de la cellule (cf. figure 3.3.).

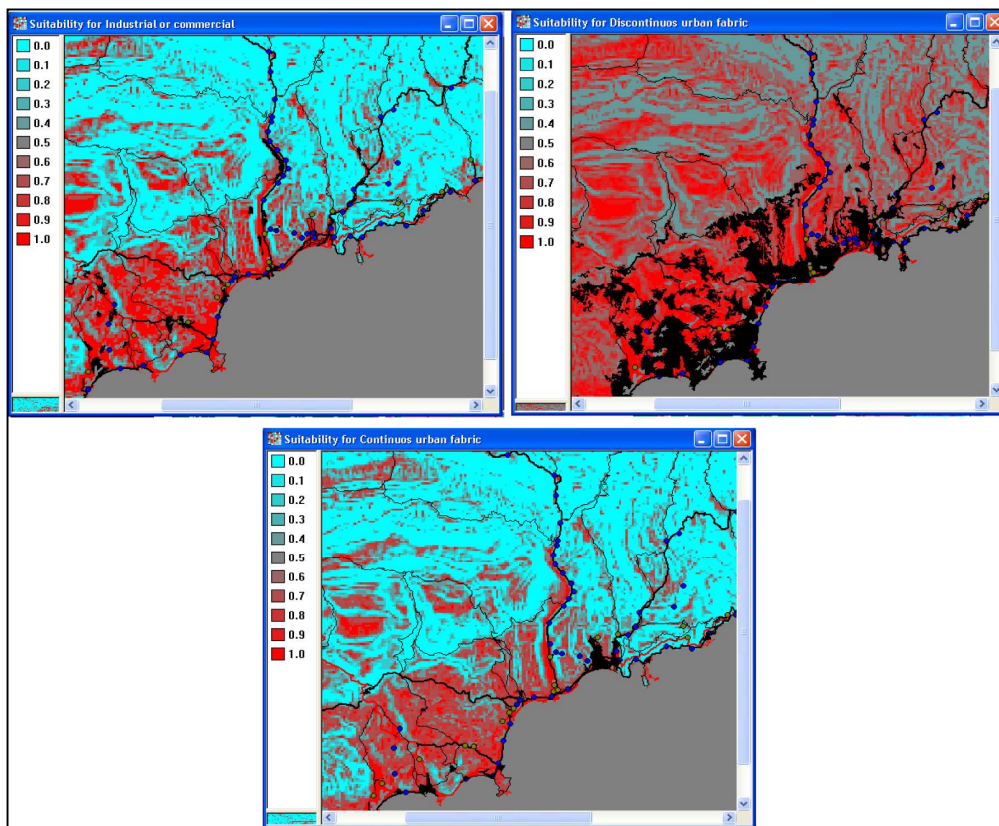


Figure 3.5 : Exemple de carte de *suitability* pour trois classes d'occupation du sol : (en noir la classe d'occupation du sol, en lignes ou points les réseaux de transport)

### 3.2.3.3. Ajustement des infrastructures de transport et détermination de l'accessibilité des lieux

Caler le facteur accessibilité, c'est mettre en évidence les interactions probables entre deux lieux. Il est calculé pour chaque classe d'occupation du sol dite active (*continuous or discontinuous urban fabric and industrial or commercial land use classes*) en fonction de chaque type d'infrastructure de transport et de l'effet de la distance.

Dans le processus de calage du facteur accessibilité, deux principaux paramètres sont pris en compte par le modèle : les paramètres de poids du réseau et les paramètres d'accessibilité eux-mêmes. Les paramètres de poids du réseau rangés entre 0 et 1 déterminent l'importance relative de l'accessibilité d'un mode de transport en question pour un usage du sol particulier. Quand ce poids est proche de 0, l'accessibilité est jugée faible, en revanche, plus ce poids est important, c'est-à-dire proche de 1 ou égal à 1, plus l'accessibilité sera bonne et plus l'attraction qu'exerce l'infrastructure de transport sur la classe d'occupation du sol en question est grande. Pour exemple, une ligne TGV serait égale à 0 pour une cellule appartenant à la classe agriculture et égale à 1 pour une cellule appartenant à la classe des activités industrielles et commerciales. Les paramètres d'accessibilité spécifient le fait que les cellules les plus proches du réseau bénéficient d'une forte accessibilité alors que celles éloignées de ce dernier seront caractérisées par une très faible accessibilité. Dans la figure ci-après par exemple, l'accessibilité est calculée à partir des classes *discontinuous urban fabric* et *industrial or commercial land use classes*. Ici, les zones les plus sombres désignent les lieux les moins accessibles qui en l'occurrence se trouvent être les plus éloignés du réseau de transport, en revanche, les zones les plus claires, là où l'accessibilité est optimale, sont celles qui sont plus ou moins proches du réseau.

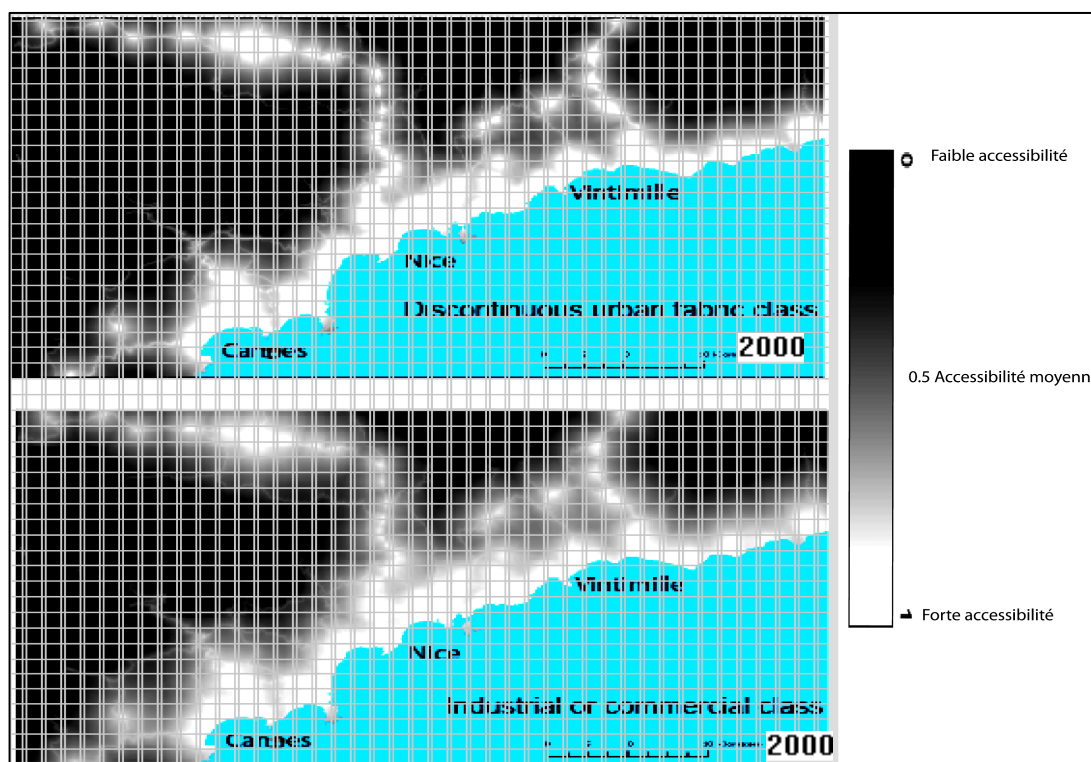


Figure 3.6. : Calage du facteur accessibilité

### 3.2.3.3. Ajustement du paramètre de perturbation aléatoire $\alpha$

Le paramètre de perturbation aléatoire conceptualisé dans la figure 3.7 suivant par Benenson et Torrens dans leur ouvrage « Geosimulation : Automata-based modelling of urban phenomena » contrôle le degré de stochasticité du modèle MOLAND. Ce paramètre indiqué par la fonction  ${}^tV$  de l'équation de transition (cf. équation 3 du chapitre 2 section 2.3.3.1. de cette deuxième partie) est l'expression de la perturbation stochastique au temps  $t$  ; définie par  $V=1+(-\ln R)^\alpha$ , où ( $0 < R < 1$ ) est une fonction aléatoire uniforme et  $\alpha$  est un paramètre autorisant l'ajustement de la taille de la perturbation stochastique. Le paramètre  $\alpha$  est indispensable au bon fonctionnement du modèle car sans stochasticité il est impossible de générer des structures d'usage du sol réalistes (cf. figure 3.8) et avec une stochasticité trop élevée, le modèle aura tendance à s'éloigner de la réalité (cf. figure 3.9). Le travail ici consiste donc à ajuster ce paramètre (en faisant tourner plusieurs fois le modèle si nécessaire) de façon à s'approcher le plus possible de la structure réelle des milieux urbains étudiés. Plus largement, ce paramètre détermine (1) le gradient de densité urbain (par exemple, la différence de densité/compacité des cellules urbaines entre le centre des agglomérations et la périphérie/moyen pays/arrière pays), (2) l'émergence de nouveaux pôles urbains/clusters à proximité des infrastructures de transport par exemple et (3) les limites/étendues/surfaces occupées par ces nouveaux pôles/clusters dans l'espace urbain. Dans cette calibration, la valeur de  $\alpha = 0.5$  fournit de bons résultats au regard des scores de la simulation (carte observée et carte simulée figure 3.11).

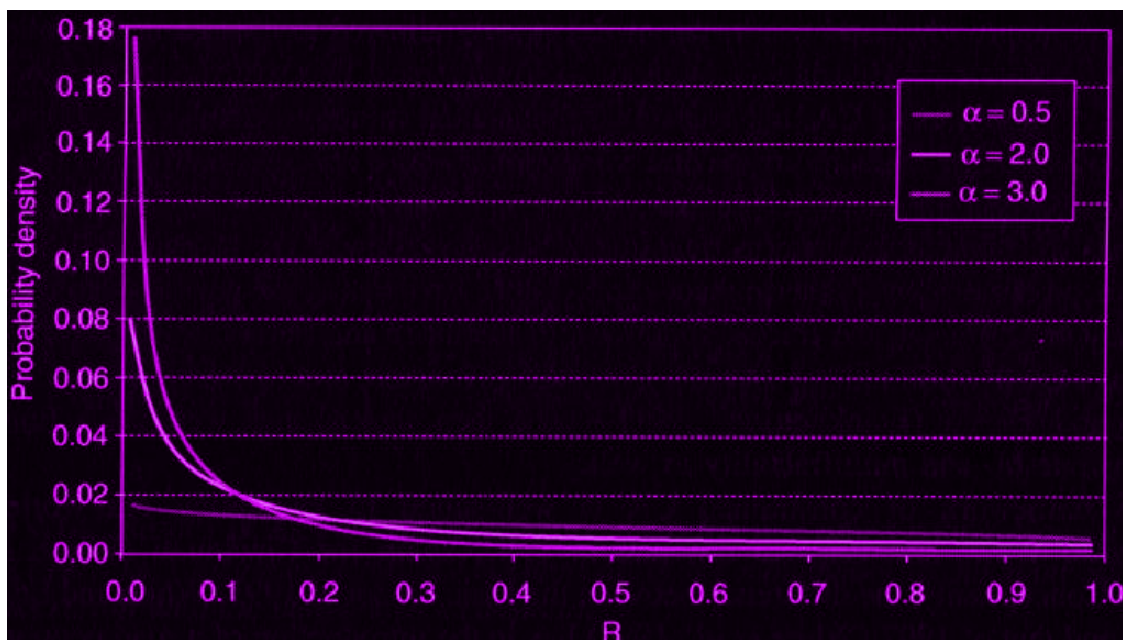


Figure 3.7 : “Distribution of stochastic multiplier of White and Engelen’s transition function for different values of parameter”

Source : Benenson et Torrens, 2004, page 117



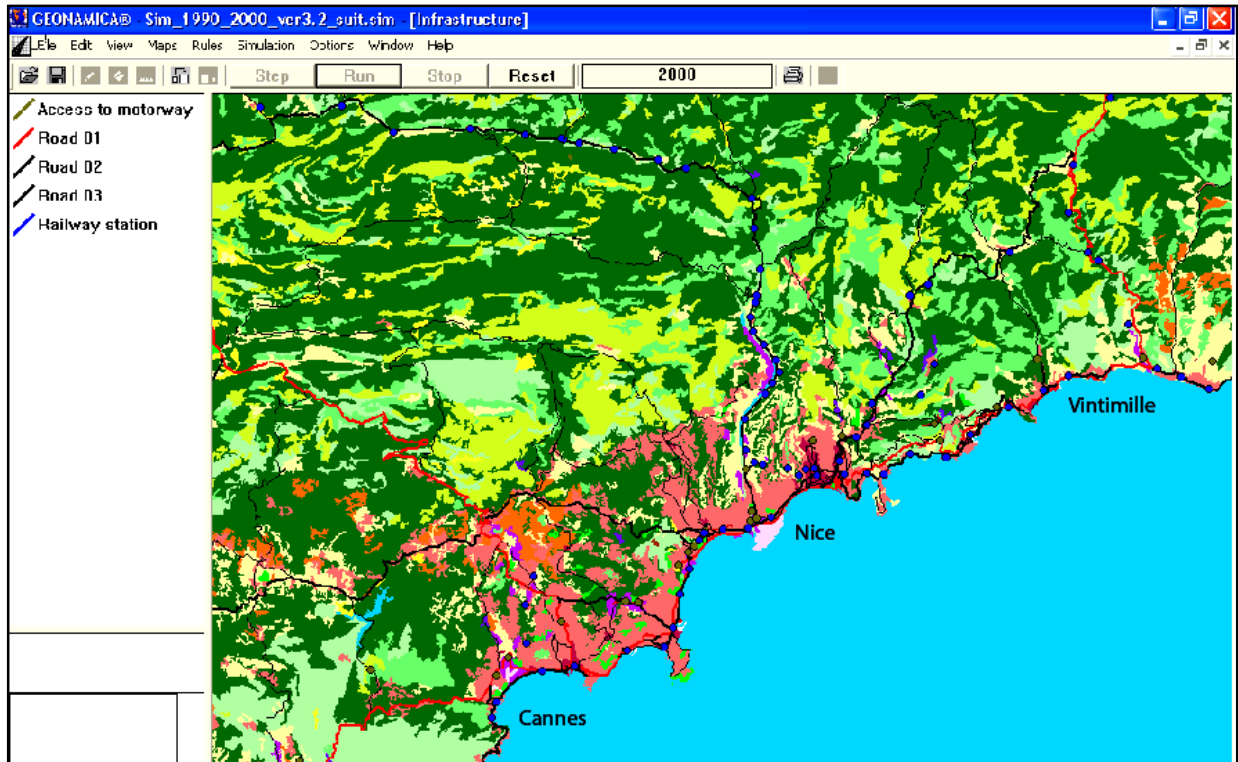


Figure 3.8 : Structures d'usage des sols générées avec une faible stochasticité

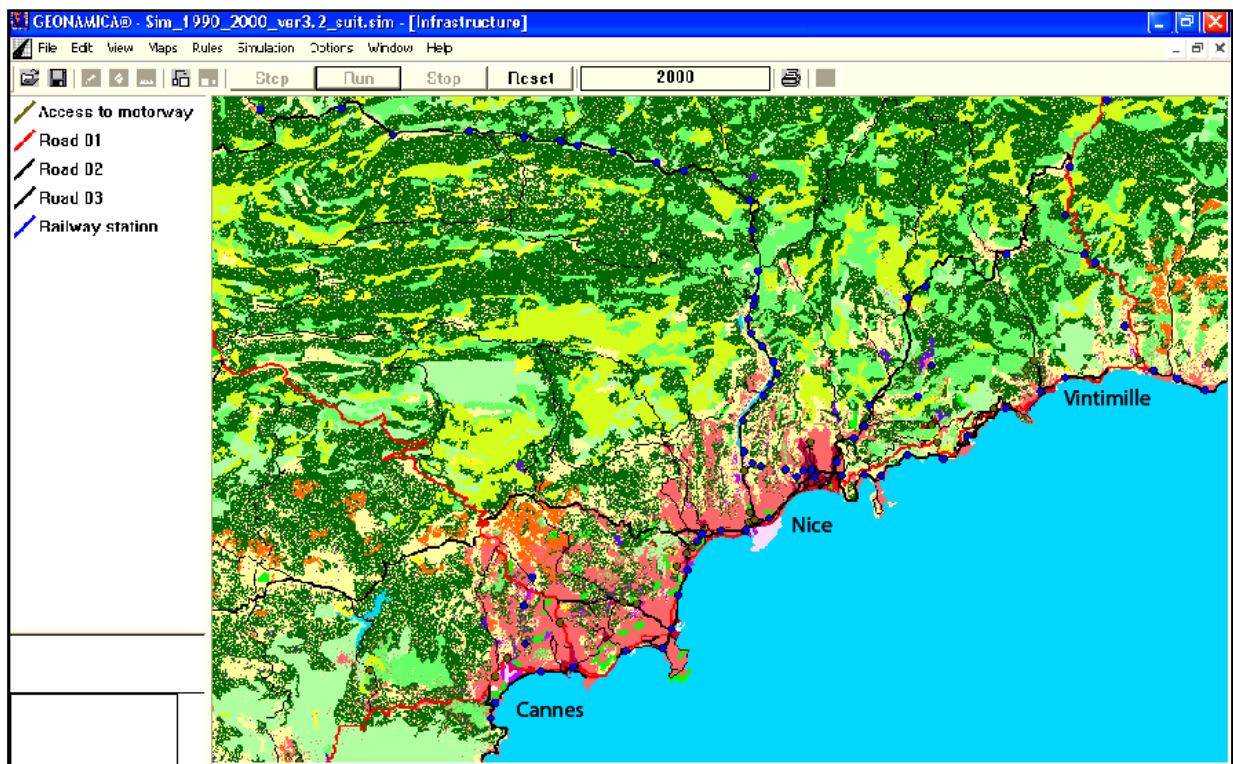


Figure 3.9 : Structures d'usage des sols générées avec une stochasticité exagérée

### **3.3. Validation du modèle : « Fuzzy Kappa Method »**

#### **3.3.1. Description de la méthode Fuzzy Kappa et fondements théoriques**

C'est Zadeh qui fut le premier en 1965 à introduire la théorie des ensembles flous (*Fuzzy sets* en anglais) permettant de prendre en compte et de décrire les « vagues d'imprécision » qui subsistent dans le raisonnement humain (Zadeh, 1965). La base de cette théorie est la notion d'imprécision. Celle-ci est directement liée aux données et aux mesures et donc à l'attribut lui-même. Aussi, dans le cadre du traitement de l'information provenant de sources différentes, cette notion ouvre des voies pour traiter et appréhender les limites que comporte inévitablement une source d'information, que les modèles spatiaux rigides, en ignorant les erreurs qui surviennent dans les procédures de mesure, ne sont pas toujours capables d'appréhender. La méthode "Kappa flou" ou "Fuzzy Kappa" est un procédé qui sert à comparer les cartes catégorielles en tenant compte à la fois des relations de proximité et du fait que certaines paires de catégories sont plus semblables que d'autres (Hagen, 2003 ; White, 2006).

#### **3.3.2. Domaine d'application de la méthode Fuzzy Kappa**

La théorie des ensembles flous prend une part de plus en plus importante pour les chercheurs qui s'emploient à décrire, expliciter et évaluer le devenir des systèmes complexes dans un temps relativement long (25 à 30 ou plus ans). Nombre d'études, utilisant différents moyens de classification d'images spatiales et spectrales sensibles ont montré que l'application de la méthode Fuzzy Kappa permet de traiter les imprécisions qui se trouvent au niveau des informations et données provenant souvent de sources différentes, et ceci grâce aux fonctions mathématiques et/ou l'algorithme de groupage de la méthode qui assigne plusieurs membres à un pixel représentant différents types d'utilisations des sols, ainsi que d'autres conditions intermédiaires. Cette méthode a également fait l'objet d'applications dans les SIG et plus spécifiquement, dans l'analyse de la propagation des incertitudes qui surviennent pendant les différentes opérations de mise en cohérence des bases de données (Burrough, 1995). C'est à partir de cette méthode que nous allons valider notre modèle.

#### **3.3.3. Validation des résultats de la calibration : la méthode Fuzzy Kappa**

Dans le cadre de la validation des modèles dynamiques et spatialement explicites, la méthode *Kappa statistic* (cf. voir le cadre suivant) est l'une des plus reconnue et son intérêt est sans cesse grandissant auprès de la communauté scientifique (Pontius, 2004 ; Shahumyan *et al.*, 2009).



**Kappa Statistic equations** (Hagen, 2004)

$$K_{Location} = \frac{P(A) - P(E)}{P(\max) - P(E)} \quad \text{Équation (4)}$$

$$K_{histo} = \frac{P(\max) - P(E)}{1 - P(E)} \quad \text{Équation (5)}$$

$$K = K_{histo} * K_{Location} \quad \text{Équation (6)}$$

$$Fuzzy\ Kappa = K_{Fuzzy} = \frac{S - E}{1 - E} \quad \text{Équation (7),}$$

avec E, calculant la similarité attendue pour chaque cellule et S la similarité globale des cellules.

Le premier facteur, K-location (cf. équation 4), est une mesure de la similarité de l'allocation spatiale des catégories de cellules des différentes cartes comparées. Le second facteur, K-histo (cf. équation 5), mesure le degré de similarité des différentes cartes. La fonction Kappa est définie comme le produit de deux facteurs. P(A) représente la fraction de concordance qui est corrigée pour obtenir la fraction de concordance statistiquement prévue à partir des « délocalisations » aléatoires de toutes les cellules dans les cartes. Ainsi, cette concordance prévue se base sur des emplacements aléatoires dans la limite de la distribution observée, c'est P(E).

Les résultats de la calibration indiquent un très bon score de la simulation car la différence entre la carte simulée et la carte observée n'est pas significative. La comparaison entre les données de 1990 et de 2000 permet d'acquérir des connaissances exactes sur l'évolution d'une classe d'occupation du sol donnée entre deux dates (cf. figure 3.10)

Les résultats indiqués dans le tableau 9 pour l'ensemble des périodes sont satisfaisants et indiquent que le modèle correspond à la réelle distribution spatiale des classes d'usage du sol dans l'aire d'étude. Les scores au niveau du Kappa, Klocation, montrent que le modèle peut fidèlement rendre compte de la réalité observée en 2000. Les résultats présentés dans ce tableau 9 sont généralement proches de 1 et de ce fait, montrent que la situation observée en 2000 est presque identique à la situation simulée pour la même période. Ces scores (cf. tableau 9) sont trop élevés et il reste encore difficile d'expliquer pourquoi (Shahumyan, White, et al., 2009). Faut-il y voir là le résultat d'une « sur-calibration » ? ou est-ce simplement une limite de la méthode fuzzy kappa ? Ce qui est une évidence à ce stade de la réflexion c'est que nous disposons maintenant d'un modèle suffisamment robuste et fiable pour être testé dans sa capacité à reproduire de façon pertinente des situations probables et réalistes. L'objectif de la calibration est donc atteint car la similitude entre la carte observée et la carte simulée comme l'indique la figure 3.11 est réaliste.

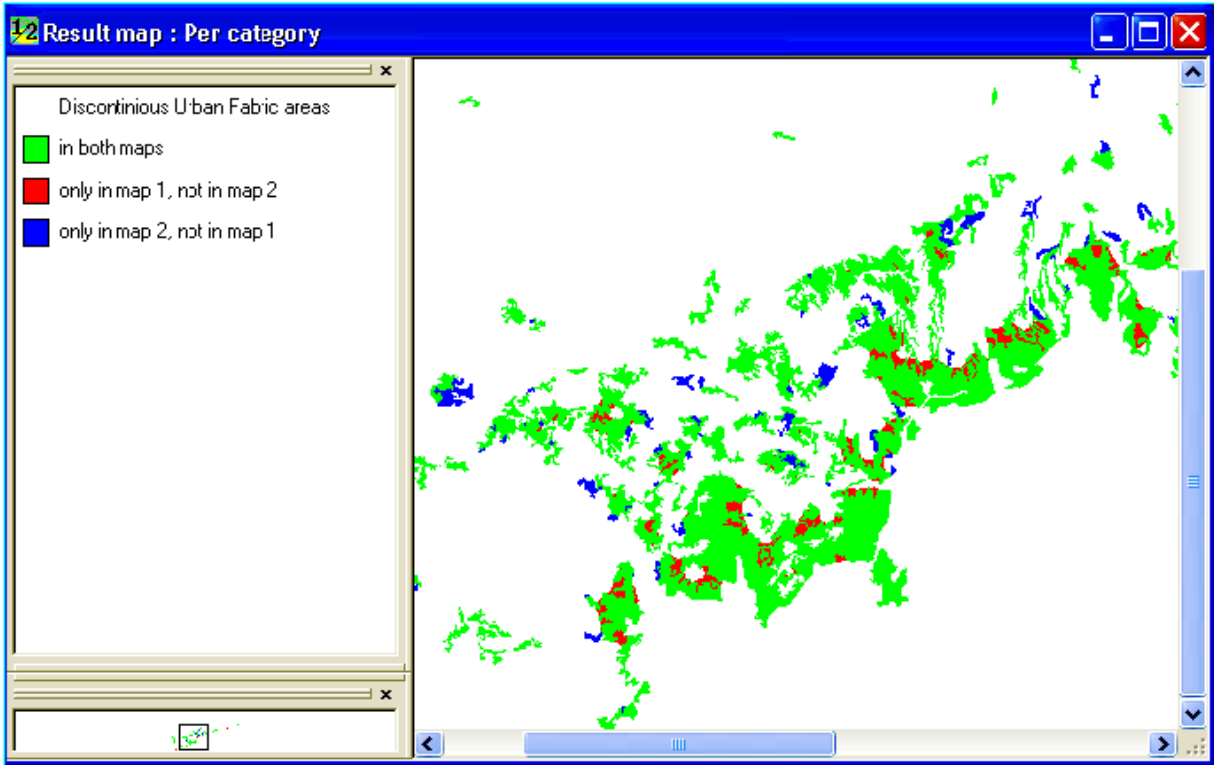


Figure 3.10 : Comparison of simulated and actual location of discontinuous urban fabric, 2000

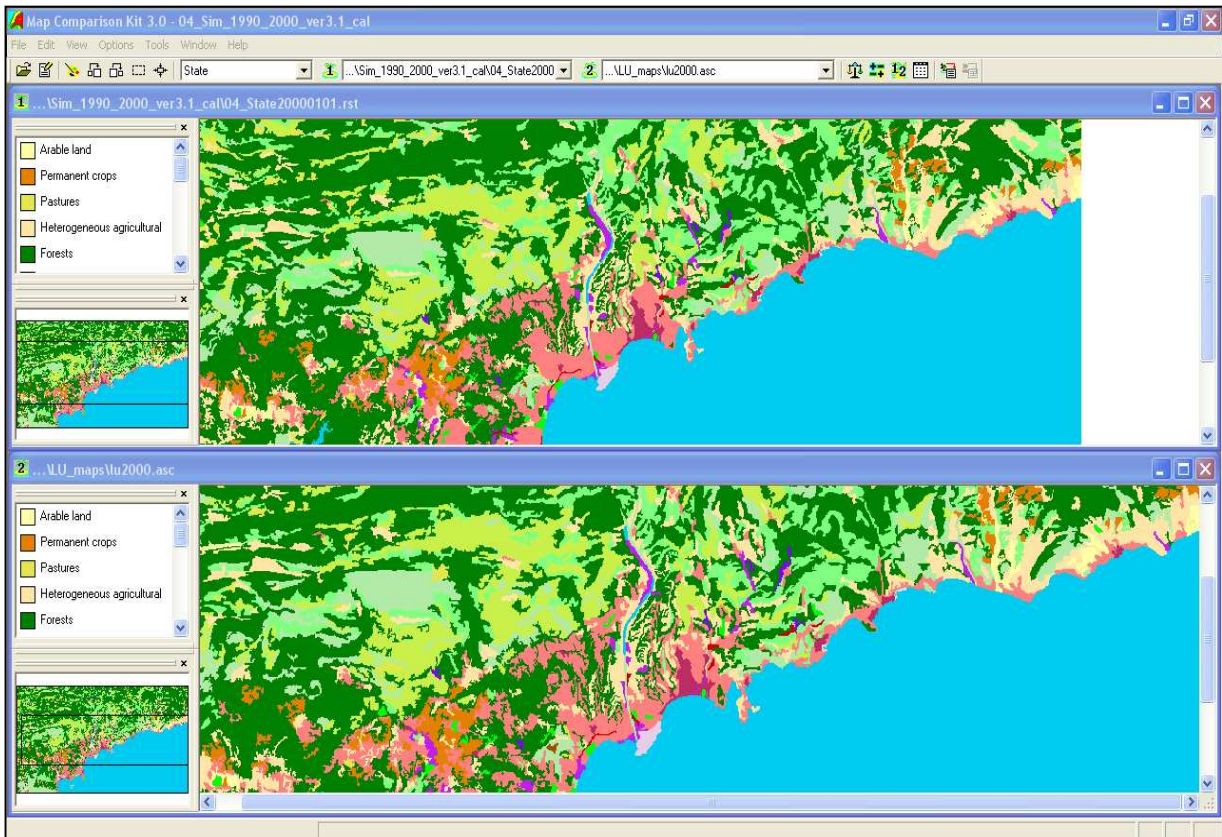


Figure 3.11 : Situation simulée en 2000 et situation réelle en 2000

Tableau 9: Validation globale de la calibration : Fuzzy Kappa comparison method (1990-2000)

	Kappa	Klocation	Khisto
<b>Overall/résultats Globaux de la calibration</b>	<b>0.98</b>	<b>0.99</b>	<b>0.99</b>
<b>Résultats détaillés des catégories du model</b>	Kappa	Klocation	Khisto
Arable land	0.91	0.98	0.93
Permanent crops	0.98	0.99	0.99
Pastures	0.99	0.99	0.99
Heterogeneous agriculture areas	0.94	0.95	0.98
Forest	0.98	0.99	0.98
Natural grassland	0.99	1.00	0.99
Sparsely vegetated area	0.99	0.99	0.99
Shrubland and/or herbaceous vegetation	0.95	0.98	0.96
Continuous urban fabric	0.99	0.99	1.00
Discontinuous urban fabric	0.92	0.92	1.00
Industrial or commercial	0.87	0.87	1.00
Airports	0.97	1.00	0.98
Mineral extraction sites	0.88	1.00	0.88
Dump sites	0.64	1.00	0.64
Green urban areas*	0.89	0.96	0.92
Road and railnetworks and associated land	1.00	1.00	1.00
Port areas	1.00	1.00	1.00
Water bodies*	1.00	1.00	1.00
Sea	1.00	1.00	1.00

### Conclusion du chapitre 3

À l'examen de ce chapitre dédié au calage et à l'évaluation des mesures d'ajustement, il semble nécessaire de faire une distinction entre le calage du modèle et la validation du modèle. Le calage se réfère au processus de création d'un modèle de telle sorte qu'il soit cohérent avec les données utilisées pour créer le modèle (Verburg et al, 2006). En revanche, l'objectif de la validation du modèle consiste à en améliorer la robustesse et l'acceptabilité en explorant à partir de méthodes de calcul basées sur la méthode Fuzzy Kappa, les erreurs et les imprécisions liées aux données et/ou mesures d'entrée ou encore, en vérifiant le degré d'influence des paramètres (effet de voisinage, accessibilité et « Suitability ») sur la performance du modèle. Les résultats des kappa statistic trop élevés invitent à une amélioration de la méthode et donc de l'outil MCK (Map Comparison Kit). Mais plus que des outils, c'est la question récurrente de la validation des modèles en général qui émerge aussi au terme de ce chapitre. La question de la validation des modèles demeure en effet un vaste champ de recherche en géographie en général et plus particulièrement en modélisation et simulation.

Par la procédure de calibration, les règles de transition ont été adaptées aux données observées en 2000 en définissant les différents niveaux de paramètres qui reflètent le mieux possible la réalité.

Au final, l'effet de la frontière n'a pas été intégré véritablement dans le modèle, puisqu'aucun paramètre n'a été déterminé en fonction de cela. Il n'en demeure pas moins que les territoires sont tellement différents de part et d'autre de la frontière, que leurs caractéristiques suffisent à intégrer l'effet frontière.

## Conclusion de la deuxième partie

Les systèmes socio-économiques du monde réel se développent dans des espaces géographiques avec des *hétérogénéités* et des *idiosyncrasies* à tous les niveaux de détail et sont formés par des processus d'interactions qui se logent à différentes échelles de l'espace géographique (Wolfram, 1994 ; Couclelis, 1988). Dans le cadre de la modélisation et de la simulation, certains de ces systèmes obéissent à des logiques « micros » et se situent ainsi dans le périmètre de l'espace cellulaire tandis que d'autres sont au-delà de l'espace cellulaire du système modélisé car obéissant à des logiques « macros ». Aussi, les modèles basés sur les automates cellulaires ayant pour but de représenter les systèmes spatiaux de manière authentique, doivent s'accommoder de ces aspects de la réalité. C'est dans ce contexte qu'il faut replacer cette partie plus mathématique et technique que la précédente mais, tout aussi méthodologique car, il y a été discuté et développé les voies que peuvent emprunter les modèles basés sur les automates cellulaires pour incorporer des processus d'interactions spatiales aboutissant à reproduire fidèlement la structure et le fonctionnement réel de l'espace transfrontalier franco-italo-monégasque.

Ces dernières années ont été marquées par un intérêt croissant pour l'exploration de nouvelles méthodes consacrées à l'analyse des incertitudes dans la perspective d'offrir une alternative à l'analyse d'erreurs subsistant dans les données et mesures d'entrée. Parmi ces méthodes, se distingue la théorie des ensembles flous. Ainsi, le fait de combiner le modèle Moland et la méthode des ensembles flous « Fuzzy Kappa » à travers l'utilisation du modèle MCK offre plusieurs avantages. Cette démarche permet d'abord de mieux localiser les erreurs produites et les omissions lors de l'étape du calage et ainsi de les corriger. Ensuite, ce couplage est une occasion d'identifier les écarts entre les résultats de la simulation et la réalité observée. Enfin, il permet de produire une simulation la plus proche possible de la réalité et de connaître la valeur de chaque cellule ainsi que les interdépendances dans le cadre des relations de voisinage. L'objectif de l'utilisation des méthodes de calcul basées sur la théorie des ensembles flous est de répartir le risque que nous assumons lorsque nous utilisons les sorties obtenues avec l'analyse des résultats d'un modèle donné dans un processus de décision. Ce type d'analyse est nécessaire pour les approches de scénarios essayant d'explorer les chemins futurs qui sont impossibles à valider complètement (Verburg *et al*, 2006a).

Dans les deux premières parties de cette étude, nous avons posé les jalons servant à définir de façon pragmatique « l'horizon du connu » du territoire transfrontalier franco-italo-monégasque en édifiant de nouvelles bases de connaissances, soit à partir de la méthode du diagnostic spatial transfrontalier - un exercice qui s'est efforcé de montrer forces, faiblesses et opportunités réelles de développement de cet espace - soit par l'étape du calage et de la validation du modèle Moland qui a permis l'acquisition de nouvelles données et par voie de conséquence, la consolidation des connaissances sur le fonctionnement du territoire transfrontalier. Dans la dernière partie, il s'agira de prédire « l'horizon de l'inconnu » à travers une approche de prospective territoriale. Cet « horizon de l'inconnu » se doit d'être bien illustré afin d'apporter la dernière pierre à l'édifice d'une recherche qui a pour vocation

première de décrire les différentes explorations qui devront impérativement être “efficacement” menées pour expliquer les changements intervenus en termes d’organisation des territoires étudiés. À ce titre, il semble, à la lecture des retours d’expérience sur les effets TGV, que l’approche prospective soit la seule voie scientifique convaincante et pertinente pour appréhender et illustrer les possibles impacts territoriaux d’un projet dont l’évolution dans le débat politique, tout dernièrement, semble interdire toute forme de certitude.



***PARTIE 3 : Scénarios géoprospectifs et simulation spatiale : Intégrer l'occupation du sol et le réseau de transport dans la modélisation d'un système complexe***

*La pensée, on l'oublie trop souvent, est un art, c'est-à-dire un jeu de précision et d'imprécision, de flou et de rigueur. (Morin, 1990).*

## Introduction

Il est particulièrement difficile, voire impossible de prédire avec exactitude le devenir d'un territoire à un horizon plus ou moins lointain, notamment quand celui-ci est un espace transfrontalier où les hétérogénéités sont plus criantes que sur un territoire non frontalier, beaucoup plus homogène. Aussi, ce qui attire l'attention à la lecture de cette partie, c'est certainement la difficulté pour le chercheur de s'adonner à l'exercice géoprospectif. Cette difficulté s'accroît dans notre cas du fait de l'absence d'informations statistiques harmonisées de part et d'autre de la frontière, informations qui s'avèrent nécessaires pour évaluer le degré de confiance à accorder aux différents facteurs à introduire dans le modèle. Elle s'accroît également au contact d'incertitudes (localisation des gares TGV par exemple, ou date de mise en service de la ligne, 2020 ? ou 2023 ?) tournant autour de la réalisation du projet LGV PACA.

Dans un tel contexte, l'exercice de la prospective vs Géoprospective ici se présente comme une tâche particulièrement délicate et complexe dans sa réalisation. Mais, faut-il pour autant renoncer à cet exercice ? À l'évidence non, car si l'on se réfère aux résultats issus de la modélisation et de la simulation, même s'ils sont loin d'être la réalité, ils s'en approchent néanmoins et apparaissent, en l'occurrence, comme de nouvelles bases de connaissances pour la compréhension des processus complexes d'urbanisation future de l'espace transfrontalier franco-italien. Par conséquent, l'exercice géoprospectif à travers la modélisation et la simulation, est considéré dans cette recherche comme l'occasion d'ouvrir les champs des futurs possibles dans le cadre de l'aménagement de l'espace transfrontalier à l'horizon 2020-2040. Par ailleurs, les résultats issus de la modélisation, dans un contexte de concertation publique, peuvent constituer de nouvelles bases de connaissances pour l'aide à la décision. L'objectif d'aide à la décision prend tout son sens à l'heure où aménagement durable des territoires implique nécessairement protection de l'environnement.

La démarche de géoprospective territoriale permet de se projeter dans l'avenir et d'imaginer le devenir de l'occupation du sol des espaces à un horizon plus ou moins lointain. L'idée de se projeter dans le futur est d'ailleurs conforme aux recommandations des documents et ou lois d'aménagement des territoires (SRU, LOADDT), qui considèrent *l'exercice prospectif comme une nécessité dans la gestion des territoires*. En effet, ces documents insistent en grande partie sur *l'idée de projet* en admettant au préalable que *c'est en connaissant ce vers quoi l'on veut tendre demain que l'on peut aujourd'hui prévoir et réglementer*. C'est ici le point de départ pour bon nombre de scientifiques ayant comme angle d'attaque, la prospective. Aussi, pour répondre aux exigences de l'exercice prévisionnel, les scénarios géoprospectifs proposés dans cette recherche, tous confondus, s'appuient sur la méthode consistant à prendre en compte l'évolution passée et présente du territoire pour prédire ses dynamiques futures. La mise en œuvre de cette démarche, au moyen d'un automate cellulaire (Moland-Model) permet d'offrir d'une part une vision diachronique de

l'évolution de l'occupation du sol de l'espace transfrontalier franco-italien, et d'autre part des images de son futur en termes de développement urbain et d'accessibilité. Trois chapitres sont proposés dans cette partie.

Dans le premier chapitre, sont déclinés les fondamentaux de la démarche de géoprospective territoriale. Ici, est mise en évidence toute la difficulté quant à l'application de cette démarche. Dans le même chapitre, un éclairage constant est apporté quant à l'élaboration des hypothèses et la construction des scénarios géoprospectifs qui se résument aisément en deux grandes familles. La première propose une image du territoire à l'horizon 2040, il s'agit du scénario tendanciel appelé aussi *Business As Usual* (BAU). La seconde, visualise des dynamiques spatiales dans un contexte de ligne à grande vitesse. Autour de cette dernière famille de scénarios s'articulent deux sous-scénarios. Le premier part de la localisation des possibles gares TGV en centre des agglomérations, et le second s'appuie sur une localisation des futures gares TGV en périphérie urbaine.

Dans le second chapitre, sont simulées et visualisées les dynamiques urbaines, sur la base de l'occupation du sol. À partir de là, et dans le cadre de l'occupation du sol des systèmes urbains transfrontaliers, nous sommes en mesure de proposer des réponses aux multiples questions qui ont jalonné notre cheminement : quels sont les espaces qui seront les plus dynamiques en 2040 dans un contexte de ligne à grande vitesse ou en son absence ? Pourquoi particulièrement ces lieux et pas d'autres ? Quelles sont les entrées du modèle qui ont le plus déterminé la trajectoire de ces espaces urbains ? En dehors du modèle, quels autres facteurs (non inclus dans le modèle) pourraient expliquer ces dynamiques ?

Enfin, dans un dernier et troisième chapitre, nous insistons sur les conséquences du TGV sur l'accessibilité des lieux en général, et sur le nombre de personnes qui, grâce à la nouvelle infrastructure, devraient ou non être affectées par les limites/contours des isochrones des cartes d'accessibilité. Parallèlement, nous soumettons trois méthodes destinées à valider les résultats issus des scénarios. Une première qui consiste à confronter ces résultats et des documents d'aménagements ou autres études prospectives réalisées sur l'aire d'étude. Une seconde méthode qui consiste à comparer ces mêmes résultats aux différentes tendances déjà mises en évidence dans l'étape du diagnostic spatial. Et une troisième et dernière méthode, plus quantitative, qui consiste à mesurer le degré de validité des résultats des scénarios géoprospectifs à partir de la méthode Fuzzy Kappa.