



HAL
open science

URBANSIMUL 1. Un modèle prototype de simulation de l'occupation du sol à l'échelle parcellaire

Ghislain Geniaux

► **To cite this version:**

Ghislain Geniaux. URBANSIMUL 1. Un modèle prototype de simulation de l'occupation du sol à l'échelle parcellaire. [Rapport Technique] 2011. hal-02811415

HAL Id: hal-02811415

<https://hal.inrae.fr/hal-02811415>

Submitted on 6 Jun 2020

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.

URBANSIMUL 1

Un modèle de simulation de l'occupation du sol à l'échelle parcellaire

Ghislain Geniaux*

March 18, 2011

Abstract

On propose dans ce document de travail un modèle de simulation de l'évolution de l'occupation du sol et du nombre de logements à différentes échelles intitulé URBANSIMUL 1. Ce modèle est élaboré à l'échelle parcellaire sur 3 SCOT de l'OUEST du Vaucluse à partir de modélisation en économétrie spatiale. On s'intéresse essentiellement au caractère bâti ou non bâti des parcelles et au nombre de logements qu'elles peuvent accueillir. Le modèle est dit a-spatial dans les zones déjà constructibles des POS/PLU et procède à une allocation des nouveaux logements à l'échelle d'un découpage des zonages POS/PLU par niveau d'inconstructibilité (indice construit par croisement des parcelles avec les zonages environnementaux, les zonages risques, le bati et les infrastructures, voir Geniaux et al. (2009a)). Hors des zonages constructibles, le modèle est spatialisé et s'appuie sur un approche probabiliste à l'échelle parcellaire mobilisant des estimations de type logit GAM géoadditifs (Wood, 2006; Kammann and Wand, 2003). Les probabilités sont estimées en fonction d'une gamme importante de facteurs ayant joué sur le processus d'urbanisation des parcelles dans les 20 dernières années (facteurs physiques, économiques, sociaux, environnementaux et politiques publiques renseignés dans un SIG historique à l'échelle Parcellaire sur PACA). A partir des résultats de ce modèle probabiliste, un modèle de simulation de l'extension à moyen terme de l'urbanisation paramétrable en fonction

*INRA (Institut National de la Recherche Agronomique, Avignon FRANCE), GREQAM-IDEP (Groupement de Recherche en Economie Quantitative d'Aix-Marseille, Institut d'Economie Publique, Marseille). Author Adress for correspondence : Ghislain Geniaux, INRA Ecodeveloppement, Site Agroparc Domaine St Paul, 84914 Avignon Cedex 9 FRANCE. email: geniaux@avignon.inra.fr

de différents critères a été calibré et testé. Plusieurs de ces critères, dont les niveaux initiaux sont issues du modèle probabiliste, comme l'incidence de la distance au noyau villageois, de la distance au réseau routier et de la densité lissée dans un voisinage de 500m, peuvent être modifiés pour représenter différents scénarii de politiques d'urbanisme, notamment du point de vue des efforts de densification de l'habitat. Il est également possible de choisir la date de projection, les modalités de calcul des projections démographiques, le niveau de préférence entre conservation des zonages naturels versus agricoles, ainsi que différents critères de calcul des seuils sur l'implantation de l'habitat par type de zone POS/PLU. Les paramètres initiaux utilisés pour la construction de ces seuils sont estimés à partir d'une analyse géostatistique de l'évolution du bâti résidentiel entre 1950 et 2008 sur l'ensemble de la région PACA, que permet la le couplage de la base de données foncières MAJIC (utilisation de la date de construction des éléments bâtis) avec les cadastres numérisés. Le modèle de simulation permet d'évaluer les conséquences de différentes options de politiques publiques en termes d'effort de densification sur de larges territoires et de les représenter cartographiquement, fournissant ainsi un outil de prospective sur l'urbanisation et de veille sur les lieux de tensions foncières.

Keywords: land use model, urban growth simulation

Introduction

Longtemps les disciplines s'intéressant à la modélisation du processus d'urbanisation ont dû se contenter de données satellitales sur l'occupation du sol et se sont appuyées dans leurs modèles sur des données à des mailles grossières, disponibles sur peu de dates et négligeant quasi-systématiquement la régulation publique. On note en effet très peu de tentatives s'appuyant sur la délimitation réelle des propriétés foncières comme les contours des parcelles (Irwin, 2010) pour des raisons tenant aussi bien à la disponibilité pour la recherche de ce type de données relativement nouvelles sur de larges étendues, qu'à la capacité des équipes à disposer de temps et d'outils permettant le traitement statistique de millions de parcelles, qu'il faut croiser avec différentes sources de données sur les activités, sur les usages du sol et leurs réglementations. Travailler à cette échelle est d'ailleurs essentiel pour les sciences sociales dans la mesure où cela peut permettre de disposer d'informations de sources diverses sur les propriétaires, leurs activités et leur environnement, et incidemment sur les déterminants de leurs décisions. Les données géographiques sur les zonages des politiques publiques sont également relativement nouvelles et ne couvrent pas encore l'ensemble du territoire français pour ce qui est des plans d'urbanisme. Les travaux de modélisation en économie régionale et urbaine anglo-saxons, qui constituent la masse de la production scientifique sur ce sujet, négligent ce type de données pour des raisons tenant aux modèles théoriques de référence qui supposent très majoritairement que le marché dicte les usages du sol et que la politique publique ne fait qu'entériner à travers ses zonages les préférences individuelles en matière de localisation (Evans, 1999). Pour ceux qui s'intéressent véritablement au rôle des zonages et des plans d'urbanisme (McDonald and McMillen, 2004; Munneke, 2005; Irwin et al., 2003; Cho and Newman, 2005; Wu and Cho, 2007), la question de la disponibilité des données géographiques est déterminante et la plupart des travaux utilisent des données d'enquête qui identifient l'existence d'une série de dispositifs de régulation dans chaque municipalité, sans disposer de leur organisation spatiale.

Dans le cadre de travaux d'étude et de recherche menés en collaboration entre l'INRA les services déconcentrés de l'État en région PACA, la région PACA et le CETE sud-est, l'élaboration d'un modèle de simulation sur l'extension de l'organisation a été engagé dans la zone des trois SCOT ouest du Vaucluse. Dans ce cadre, il nous a été possible sur cette zone de disposer depuis quatre ans de l'ensemble des données foncières, des POS/PLU numérisés, des réseaux d'infrastructure et de l'ensemble des cadastres numérisés disponibles¹ pour procéder à une série de travaux géostatistiques novateurs à cette échelle et à cette étendue géographique, qui aujourd'hui s'étend sur

¹Il est à noter que sur cette zone l'ensemble des communes dispose de cadastres numérisés et de POS/PLU numérisés.

l'ensemble de la région PACA (Vimal et al., 2011).

L'article suit le plan suivant. Dans une première partie nous présentons les données utilisées et proposons une synthèse des traitements géomatiques permettant de construire les endogènes, les variables explicatives et les seuils de référence utilisés dans les modélisations et simulations. La seconde partie présente l'architecture globale du modèle de simulation et décrit en détail la partie déterministe du modèle concernant les espaces constructibles des plans d'urbanisme. Dans une troisième partie, nous détaillons le modèle d'estimation des probabilités de conversion au sein des espaces non constructibles des plans d'urbanisme. Dans une quatrième partie, nous présentons et discutons les résultats d'une simulation menée sur un horizon de 20 ans sur la base d'un scénario calé sur les paramètres par défaut du modèle visant à produire une prolongation de tendance. La conclusion revient sur l'utilisation de cet outil par les partenaires publics du projet, et esquisse les principales voies d'amélioration retenue à ce stade.

1 Données

La zone géographique d'études englobe les trois SCOT ouest du Vaucluse (Avignon, Carpentras, Cavaillon). Les principales sources de données géographiques utilisées dans le cadre de ce modèle sont les suivantes : la matrice littérale du cadastre (MAJIC II, 2005), les plans cadastraux informatisés (PCI 2006), les POS/PLU généralisés 2006, BD TOPO®, les zonages environnementaux et de gestion des risques (DIREN PACA, MNHM), les historiques de population et du nombre de logements communaux des recensements de l'INSEE. Trois communes parmi les 72 de la zone d'étude ont été exclues car elles ne disposent d'aucun plan d'urbanisme.

Sur cette zone d'une surface de 160000 ha, on comptait en 2006 environ 400 000 parcelles, 270 000 éléments bâtis pour 230 000 logements. Le tableau 1 indique pour chaque type de zonage la surface concernée dans la zone d'étude.

Un premier travail préalable d'identification des espaces ouvrables à l'urbanisation, mené de 2005 à 2009, nous permet de disposer pour chaque parcelle de la zone d'étude d'un indice de constructibilité "théorique" en 5 niveaux qu'il est possible de décliner dans les zones réglementairement constructibles ou inconstructibles des plans d'urbanisme. Cet indice de constructibilité "théorique" est élaboré à partir de l'intégralité des données géographiques susceptibles d'être mises à disposition des collectivités territoriales. Un premier type de contraintes considérées dans cet indice concerne les zonages réglementaires, notamment les zonages environnementaux et les zonages de gestion des risques. Une classification en 5 niveaux d'inconstructibilité en fonction des règles de ces zonages du

Table 1: Surface et type de zonages des POS/PLU généralisées sur la zone d'étude

Label	Zonage	Surface en ha
<i>Zonages constructibles des POS/PLU</i>		
Zone d'habitat dense	GHAB	9373
Zone d'équipement Urbain	GE	2182
Zone à Urbaniser pour l'habitat	GNA	3327
Zone à Urbaniser pour les activités	GNAE	3268
Zone d'habitat diffu	GNB	1734
<i>Zonages non constructibles des POS/PLU</i>		
Zone agricole	GND	42465
Zone naturelle	GNC	69651
<i>Hors POS/PLU, zone d'application du MARNU</i>		
Hors POS/PLU	HP	7420

point de vue l'habitat nouveau a été élaborée en partenariat avec la DRE, la DIREN et le CETE². Un second type de contraintes concernant les caractéristiques physiques des parcelles (bâti/pas bâti, part entre surface bâti et non bâti, proximité à des infrastructures, pente, etc.) a permis d'exclure les parcelles et partie de parcelles non constructibles. Le tableau 2 liste les types de zonages et de contraintes retenus dans (Geniaux et al., 2009a). Un second travail préalable a été mené sur l'évolution de l'implantation du bâti dans les 20 dernières années sur l'ensemble de la région PACA à partir d'un traitement géo-statistique des dates de construction des éléments bâtis de la matrice cadastrale (MAJIC) afin d'identifier les entités spatiales au sein desquels on observe une certaine homogénéité des densités de bâti et de leur évolution récente. Ces densités ont été estimées dans chaque zone POS/PLU de chaque commune en PACA. On a identifié 3 facteurs d'organisation des densités de logement de ces polygones. Naturellement, le type de zonage POS/PLU est le principal facteur de structuration des densités d'habitat et de leur évolution. Ensuite, les densités au sein d'un même type de zonage varient essentiellement en fonction de la population de la commune et de la population cumulée des communes voisines. On note en effet des processus de densification très divergents entre communes de même classes de population selon qu'elle sont au coeur des agglomérations, en première couronne, en deuxième couronne ou en zone rurale (Geniaux et al., 2005, 2009b). Ainsi, pour des communes de même classe de population, les zonages urbains denses de

²Le CETE sud-est utilise désormais cette classification dans ses travaux statistiques.

deuxième couronne peuvent afficher des densités d’habitat comparables à celle du zonage d’habitat diffus en coeur d’agglomération ou en première couronne. A partir de ces trois facteurs, on a construit une typologie de communes en une vingtaine de classes : la population communale en 5 classes et la population lissée dans un voisinage de 50 km en 4 classes³. Cette typologie de commune est ensuite croisée avec le type de zone POS/PLU en 7 classes (zonage dit ”généralisé de niveau 2” : urbain dense, urbain équipement, urbanisation future, urbanisation future pour activités, zones d’habitat diffus, zones agricoles, zones naturelles). Par exemple, on considère comme des entités spatiales comparables les zones POS/PLU de type *zone agricole* dans les communes de 0 à 500 habitants situées en première couronne des agglomérations. A partir de traitements statistiques des densités d’habitat dans ces entités spatiales homogènes sur l’ensemble de la région PACA, une série de seuils sur les distances minimales entre bâtis, les densités d’habitat moyennes, médianes et maximales ont été estimées pour caler la partie déterministe du modèle URBANSIMUL1 et servir de paramètre de contrôle.

La méthodologie d’estimation des indices de constructibilité des parcelles et des statistiques sur les densités par entités spatiales homogènes ont fait l’objet de plusieurs publications (Vimal et al., 2011; Geniaux et al., 2009a,b). Le tableau 2 permet de lister pour chaque type de zonages et de contraintes, le niveau d’inconstructibilité que lui attribue la méthode de classification proposé dans (Geniaux et al., 2009a). Pour chaque parcelle, le niveau d’inconstructibilité est soit nul, soit gradué dans une échelle allant de un à quatre. Le niveau 1 correspond à une inconstructibilité totale ; les niveaux 2 à des contraintes fortes, le niveau 3 à des contraintes moyennes, et le niveau 4 à des contraintes faibles. Le tableau 3 liste les superficies dans chaque niveaux de contrainte selon le caractère constructible ou pas de la zone POS dans lequel se trouve la parcelle.

Enfin, un travail de géomatique a été mené pour construire une série de covariables candidates pour le modèle probabiliste. Pour construire les variables d’environnement des parcelles, on s’appuie sur une grille de 500m par 500m soit 2.5 ha (la surface médiane des parcelles est 0.1 ha) pour estimer des indicateurs sur l’urbanisation dans le voisinage proche des parcelles. L’accès à la matrice littérale du cadastrale, nous donne ici la possibilité d’estimer le nombre réel de logements et pas uniquement le nombre d’éléments bâtis comme on le retrouve généralement dans ce type de modélisation en géographie s’appuyant sur des données comme BD TOPO® ou les informations graphiques des PCI, ou encore dans le pire des cas sur des données de télédétection qui n’arrive à distinguer que des taches d’urbanisation. On estime également pour chaque parcelle une série de distance minimale

³Les quatre classes de population lissée spatialement ($popliss_i = \sum_j \exp^{-0.1dist_{ij}} POP_j \quad \forall j \text{ avec } dist_{ij} < 50km$) permettent de distinguer les communes situées dans le coeur des agglomérations, la première couronne des agglomérations, une deuxième couronne plus large, et les zones rurales.

Table 2: Zonages environnementaux et risques, et indice de constructibilité

Layers	Niveau de contrainte			
	1	2	3	4
Bâti , Sources MAJICII, PCI				
Parcelles considérées comme déjà bâties*	X			
Emprises du bâti sur les parcelles considérées comme partiellement bâties*	X			
Infrastructures et réseaux , Sources BD TOPO ®				
Aires de triage	X			
Cimetières	X			
Maisons	X			
Réservoirs	X			
Terrains de sport	X			
Parkings, péages, etc.	X			
Pistes d'aérodrome	X			
Emprise des voies ferrées*	X			
Zonages environnementaux , Sources DIREN ⁴ , MNHM ⁵				
Arrêtés de protection de biotope	X			
Terrains du conservatoire du Littoral	X			
Parcs nationaux (zone centrale)	X			
Réserves naturelles nationales	X			
Sites classés	X			
Réserves de la biosphère (zone centrale)		X		
Sites Ramsar		X		
Réserves naturelles régionales		X		
Sites inscrits		X		
Sites d'intérêt Communautaires (SIC natura 2000)			X	
Réserves biologiques			X	
Réserves biogénétiques du Conseil de l'Europe			X	
Réserves nationales de chasse et faune sauvage			X	
Zones de protection du patrimoine architectural, urbain et paysager (ZPPAUP)			X	
Parcs nationaux (zone périphérique)			X	
Zones de protection spéciale (ZPS)			X	
Zones spéciales de conservation (ZSC)			X	
Zones d'importance pour la conservation des oiseaux (ZICO)			X	
Réserves de la biosphère (zone tampon)				X
Réserves de la biosphère (aire de transition ou de coopération)				X
Périmètres de protection autour des réserves naturelles géologiques (PPRNG)				X
Projet d'intérêt général (PIG)				X
ZNIEFF 1				X
ZNIEFF 2				X
Zones sensibles				X
Zones vulnérables				X
Parcs naturels régionaux				X
Prévention des risques et bruit , Sources DIREN ⁶ , MNHM ⁷				
PPRI rouge	X			
PPRI orange			X	
PPRI rouge rayé (urbain)			X	
PPRI orange rayé (urbain, centre ancien)			X	
PPRI jaune			X	
Routes soumises à la loi Barnier*	X			

(*) la couche d'origine a subi des transformations géométriques : sélection, buffer ou découpage.

Table 3: Superficie par niveaux de contrainte de constructibilité et type de zone POS/PLU

Zone POS/PLU	Contrainte	label	surface en ha
Urbanisable	NULL	Disponible à l'urbanisation nouvelle, sans contrainte	10137
Urbanisable	4	Disponible à l'urbanisation nouvelle, avec contraintes faibles	3052
Urbanisable	3	Disponible à l'urbanisation nouvelle, avec contraintes moyennes	5669
Urbanisable	2	Disponible à l'urbanisation nouvelle, avec contraintes fortes	0
Urbanisable	1	Non disponible, contraintes très fortes	8304
Non Urbanisable	NULL	Disponible à l'urbanisation nouvelle si évolution du POS/PLU, sans contrainte	21162
Non Urbanisable	4	Disponible à l'urbanisation nouvelle si évolution du POS/PLU, avec contraintes faibles	20918
Non Urbanisable	3	Disponible à l'urbanisation nouvelle si évolution du POS/PLU, avec contraintes moyennes	57237
Non Urbanisable	2	Disponible à l'urbanisation nouvelle si évolution du POS/PLU, avec contraintes fortes	0
Non Urbanisable	1	Non disponible, contraintes très fortes	16857

bord à bord entre la parcelle et :

- le noyau villageois (coeur urbain dense),
- les différents type de zonage POS/PLU,
- les routes principales, secondaires et chemins.

2 Modélisation

Le modèle URBANSIMUL, dont la figure ?? donne une représentation schématique, est décomposé en deux sous modèles selon le type de zones POS/PLU sur lequel il s'applique. Dans les zones constructibles des plans d'urbanisme ou l'habitat nouveau se fait largement au sein de parcelles déjà constructibles⁸, nous avons privilégié un modèle déterministe a-spatial qui s'intéresse essentiellement au nombre de logements nouveaux globaux par entités spatiales. Dans les zones non constructibles des plans d'urbanisme nous nous avons développés un modèle probabiliste spatialisé qui estime pour chaque parcelle une probabilité de conversion vers des usages résidentiels.

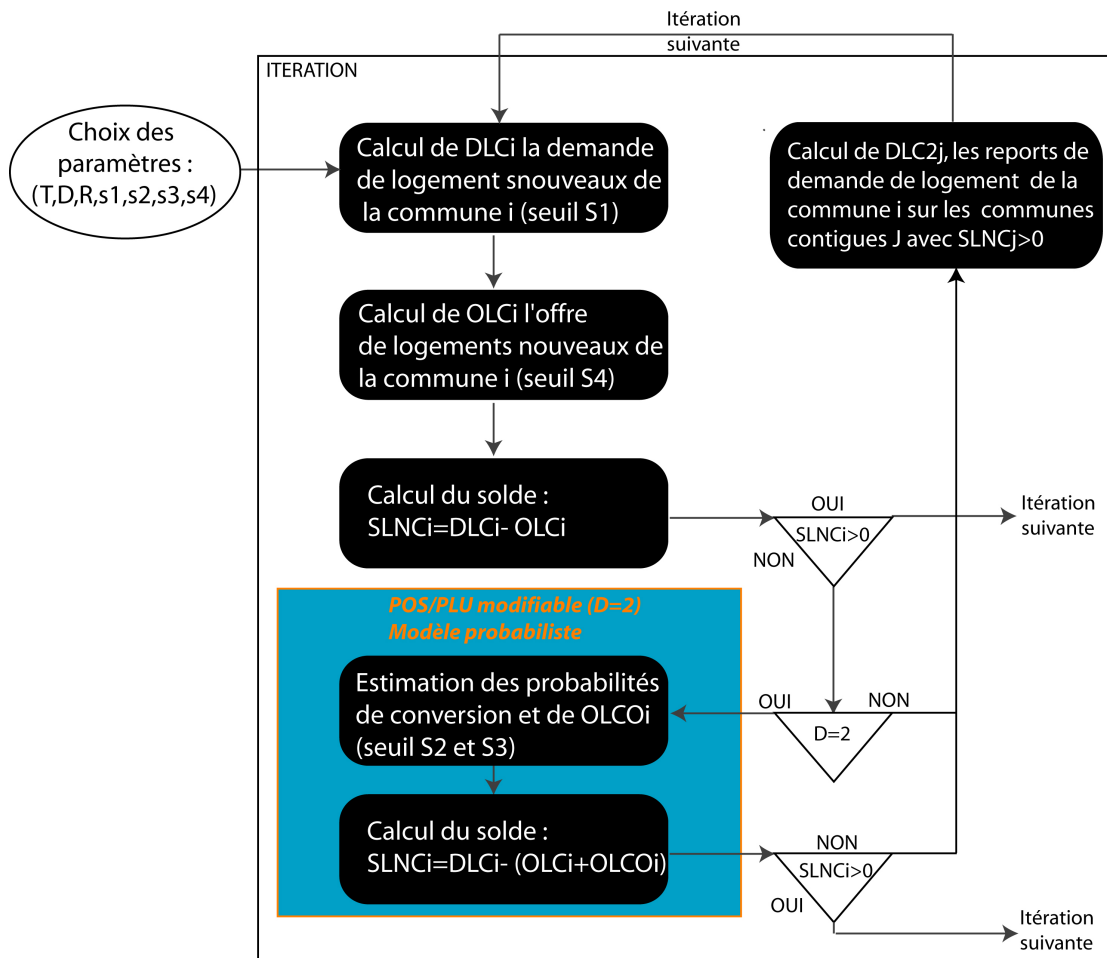


Figure 1: URBANSIMUL 1, schéma global

⁸Vimal et al. (2011), estime à partir des mêmes données que dans les zones constructibles des plans d'urbanisme entre 30 et 70 % des nouveaux logements entre 1990 et 2006 se sont faits sur des parcelles ayant déjà au moins un élément bâti.

Ce choix s'explique par différentes raisons. Pour ce qui concerne les zonages déjà constructibles, les tentatives de modélisation de la localisation de l'habitat nouveau en zone constructible que nous avons préalablement menées ont débouché sur des modèles affichant des qualités d'ajustement et de prédiction très faibles. Ensuite, si l'on se place dans une problématique de consommation d'espace par l'urbanisation et de préservation des espaces agricoles et naturelles, la localisation précise de cet habitat nouveau présente moins d'enjeux en zone déjà constructible. On peut noter, qu'à l'instar de la dynamique des prix fonciers observés en zone non constructible (Geniaux and Napoléone, 2005; Geniaux et al., 2011), les probabilités localisées de conversion constituent des indicateurs très intéressants pour repérer les zones non constructibles où des anticipations fortes des propriétaires peuvent apparaître.

Le fonctionnement global du modèle que reprend la figure ?? peut être résumé ainsi : Etant donné un horizon temporel T , à partir d'une projection démographique sur l'ensemble de la zone basée sur une simple prolongation de tendance des évolutions intercensitaires communales⁹, on répartie dans un premier temps, le nombre de nouveaux logements entre communes en fonction de la règle de répartition R et on détermine une demande logement communal pour chaque commune i noté $DLC_i(R, T)$. On estime pour chaque commune l'offre potentielle de nouveaux logements au sein des zonages déjà constructibles en fonction de leur surface disponible pour l'urbanisation nouvelle et des niveaux de contraintes de chaque parcelle, noté OLC_i dont les modalités d'estimation sont détaillées plus loin. Si l'offre possible en zones constructibles n'est pas suffisante pour répondre à cette demande, et si les POS/PLU sont non modifiables¹⁰, alors la demande non soldée SLC_i est reportée vers les commune J contiguës ayant un solde SLC_j positif. Si les POS/PLU sont modifiables, on estime alors des probabilités de conversion pour les parcelles situées dans les zonages non constructibles de la communes, et on estime $OLCO_i$ l'offre de logement ouvrable en zone non constructibles de la commune i , dont les modalités d'estimation sont détaillées plus loin. Si l'offre foncière en zone constructibles et non constructibles de la commune n'est pas suffisante, le solde de la demande de logement est alors reporté sur les communes contiguës non saturées.

⁹Vimal et al. (2011) propose un modèle de projection démographique communal à partir d'un modèle de régression stratifiés s'appuyant sur une procédure de minimisation des erreurs relatives et non les erreurs absolues (*percentile least square regression*) qui permet d'obtenir des projections plus fiables. Les prochaines versions d'URBANSIMUL reprendront ce type de modélisation.

¹⁰Cette possibilité est fixé par l'utilisateur.

2.1 Calcul de la demande de logement à l'horizon T

Les simulations proposées ici s'appuient sur une projection démographique DG pour l'ensemble de la zone (3SCOT). Cette projection DG peut s'appuyer sur une projection de tendance, sur les projections faites par l'INSEE ou être fixé arbitrairement par l'utilisateur. A partir de cette projection, on estime une demande de logement communal DLC_i , à partir d'une répartition de DG selon la règle de répartition $R1$, $R2$ ou $R3$. Avec la règle $R1$, DLC_i est issu d'une répartition en fonction de la part de population de la commune i au sein des 3 SCOT :

$$DLC_i = \frac{1}{pop_{it}} \times DG$$

Où pop_{it} est la population de la commune i à la date t .

Avec la règle $R2$, DLC_i est issu d'une répartition pondérée linéairement en fonction de la part de population de la commune i et de la surface de terrain disponible à l'urbanisation nouvelle au sein des zonages déjà constructibles:

$$DLC_i = \left(\alpha \frac{pop_{it}}{\sum_i pop_{it}} + (1 - \alpha) \frac{S_{kit}}{\sum_i S_{kit}} \right) \times DG \quad \alpha \in [0, 1], k \notin (GNC, GND)$$

Où S_{kit} est la surface disponible à l'urbanisation nouvelle de la commune i au sein des zonages déjà constructibles k ¹¹.

La règle $R3$ consiste à fixer arbitrairement les demandes communales. La règle $R2$ permet de faire jouer un critère de cohérence des politiques publiques qui voudrait que les communes avec beaucoup de zonages constructibles disponibles pour une urbanisation nouvelle ont planifiée un accroissement important de leur demande de logement.

DLC_i est soumis à un seuil maximal. Pour éviter des croissances de population communale irréalistes, on considère que pour chaque commune, il existe un niveau de croissance annuel maximal de la population. Ce niveau de croissance maximal est estimé en s'appuyant sur la distribution des taux de croissance de la population intercommunale (1968-1999) et exprimé en taux annuel maximum. Le niveau maximal retenu, noté $S1$, est, par défaut, le percentile 95¹² pour la région PACA par classes de population communale (0-500, 500-2000, 2000-5000, 5000-25000, >25000).

2.2 Calcul de l'offre de logement maximal en zone constructible

Chaque parcelle l est caractérisée par un type de zonage $k \in \{GHAB, GE, GNA, GNAE, GNB, GNC, GND\}$, un niveau de contrainte $z = 0, 1, 2, 3, 4$ et appartient à une classe de commune $c \in (C1, C2, \dots, C15)$.

¹¹Pour les zonages constructibles on a $k \in (GHAB, GE, GNA, GNAE, GZAC, GNB)$ et pour les inconstructibles $k \in (GNC, GND)$.

¹²Ce paramètre est modifiable par l'utilisateur

Pour chaque parcelle l , on estime la surface libre S_{lkz} en fonction de z . On estime ensuite pour chaque commune i la surface libre pour l'urbanisation nouvelle dans chaque classe de contraintes z et de type de zonage k :

$$S_{ikz} = \sum_{l \in i} S_{lkz} \quad \forall z \neq 1$$

Pour estimer l'offre de logement nouveaux qu'il est possible d'implanter sur les surfaces libres S_{ikz} de la commune i , on s'appuie sur les distributions des densités de logement observés sur l'ensemble de PACA en fonction du type de zones POS et des 20 classes de communes possibles C . Le tableau 4 présente le percentile 75 de la distribution des densités de *logement/ha* pour chaque type de zones POS en fonction de cette typologie.

Table 4: Percentile 75 de la distribution des densités de *logement/ha* pour chaque type de zones POS.

c	zone	cpop	$S4_{GHAB,c}$	$S4_{GE,c}$	$S4_{GNA,c}$	$S4_{GNAE,c}$	$S4_{GNB,c}$	$S4_{GNC,c}$	$S4_{GND,c}$
1	1	(500,2e+03]	12.09	-	-	-	-	-	-
2	1	(2e+03,5e+03]	42.93	8.30	18.42	-	8.35	2.76	6.02
3	1	(5e+03,2.5e+04]	28.26	3.94	7.18	1.50	5.95	0.99	4.09
4	1	(2.5e+04,3e+06]	51.21	4.13	9.37	2.19	6.12	3.09	2.66
5	2	(500,2e+03]	13.82	-	-	-	4.58	0.91	3.06
6	2	(2e+03,5e+03]	20.81	6.47	6.70	0.97	6.30	1.90	3.19
7	2	(5e+03,2.5e+04]	24.59	1.83	10.19	1.32	6.40	1.91	3.09
8	2	(2.5e+04,3e+06]	47.83	5.46	2.39	1.99	7.39	1.63	6.69
9	3	(0,500]	19.34	-	7.22	1.45	7.28	1.31	10.23
10	3	(500,2e+03]	22.53	9.91	6.03	1.75	5.65	1.16	5.16
11	3	(2e+03,5e+03]	21.18	2.52	7.47	1.75	6.07	1.16	2.63
12	3	(5e+03,2.5e+04]	31.79	11.47	6.62	1.14	6.12	1.05	4.58
13	4	(0,500]	19.86	5.85	4.99	0.84	5.79	0.74	1.14
14	4	(500,2e+03]	22.52	6.54	8.44	5.33	5.62	1.05	5.38
15	4	(2e+03,5e+03]	18.93	-	4.60	1.61	5.17	1.39	0.04

Le seuil de référence pour le calcul de OLC_i , noté $S4_{ck}$ s'appuie sur le percentile 75 de la distribution des densités de logement observé dans les classes (c, k) :

$$OLC_i = \sum_{k,z} S_{ikz} \times S4_{ck}(\text{logements/ha}) \text{ avec } i \in c \in C \text{ et } k \notin (GNC, GND)$$

OLC_i représente le nombre de logements maximal qu'il est possible d'implanter sans modifier les zonages POS/PLU de la commune.

2.3 Estimation des probabilités de conversion en zones non constructibles

Le modèle d'estimation des probabilités de conversion s'appuie sur un modèle logit spatial additif généralisé. C'est un modèle logit univarié car l'endogène que l'on veut modéliser concerne le caractère bâti ou non bâti de la parcelle. Le modèle estimé est le suivant :

$$p_i(bati) = \alpha + \beta_1 CLC_{111} + \beta_2 k + \beta_3 z + s(dr) + s(dv) + s(nbbatha) + s(nblogha) + s(x, y) + \epsilon \quad (1)$$

A la partie paramétrique du modèle s'ajoute des fonctions non paramétriques de type spline cubique $s(.)$ sur les variables de coordonnées spatiales des parcelles (longitude x , latitude y) et sur les variables de distance au noyau villageois (dv), de distance à la route (dr) et de densité de logement du voisinage ($nbemprisebatha$ ¹³, $nblogha$ ¹⁴). La variable CLC_{111} correspond à la classe d'occupation du sol de niveau 3 de Corrine Land Cover.

Le tableau 5 présente les résultats de la régression (1). La qualité d'ajustement du modèle est correcte avec un pseudo R2 de et un nombre faux 0 (resp. de faux 1) inférieur à 19 % (resp. inférieur à 30%). Le modèle (1) sert de référence pour la projection des probabilité dans le scénario médian qui correspond à un simple prolongement de tendance. Pour les autres scénarios, on s'appuie sur une version linéarisée du modèle (1), avec la possibilité pour les utilisateurs de modifier les coefficients des variables dr , dv , $nbbatha$ et $nblogha$, dans des plages calibrées à partir de la variance du paramètre associé.

2.4 Implantation du bâti et mécanisme de report sur les communes contigues

Etape 1. Pour les zones constructibles, si la demande de logement communal est inférieure à OLC_i , le modèle de simulation implante le bâti de façon itérative en fonction des contraintes z en privilégiant les parcelles sans contraintes $z = 0$ puis en faisant croître le niveau de contrainte $z = 4, 3, 2$, le niveau $z = 1$ étant considéré comme définitivement inconstructible. Si le POS/PLU est modifiable dans la simulation, et si le nombre de logements qu'il est possible d'implanter en zones constructibles de la commune n'est pas suffisant, le modèle de simulation identifie des parcelles ouvrables à l'urbanisation dans les zones non constructibles et calcule un nombre de logements fonction de ce type d'habitat en zone non constructible. L'implantation du bâti en zone non constructible s'opère en premier lieu en fonction du degré de contrainte z , puis en fonction de la

¹³ $nbemprisebatha$ représente le nombre d'emprises au sol des éléments bâtis par hectare dans un voisinage de 2.5 ha de la parcelle.

¹⁴ $nblogha$ représente le nombre de logement à l'hectare dans un voisinage de 2.5 ha de la parcelle.

probabilité estimée, en vérifiant 2 contraintes sur le taux et la surface de nouvelle zone constructible. On considère en effet que pour chaque commune, il existe des niveau maximaux d'ouverture de nouvelles zones constructibles des POS/PLU. Ces niveaux maximum sont estimés en PACA à partir de l'analyse des 280 communes ayant adopté une nouvelle version de leur POS entre 1999 et 2006. On estime deux types de seuil : un seuil exprimé en Ha et un seuil exprimé en % de surface communale en se basant sur le percentile 75. On note $S2$ et $S3$ ces seuils basés par défaut sur le percentile 75.

Etape 2 à n. Si la demande logement de la commune i n'est pas saturée, la demande non soldée SLC_i est reportée vers les commune J contiguës ayant un solde SLC_j positif. A ce stade du modèle, ce report ne peut s'opérer que dans les zones constructibles des communes j et ne peut pas y induire une pression à l'ouverture de nouvelles zones constructibles.

3 Simulation

A ce stade URBANSIMUL 1 est en version test chez nos partenaires publics (DREAL, CETE, REGION PACA), et une série de simulations a été mise en œuvre pour tester les comportements du modèle dans différents contextes de choix des paramètres urbanistiques. La figure 2 présente la fenêtre de choix des paramètres de simulation. Chaque simulation dure en moyenne un demi-heure sur ce territoire.

On présente ici quelques représentations cartographiques à différentes échelles du modèle sans commenter en détail les résultats de ces simulations. A l'échelle du territoire des 3, on peut par exemple se poser la question de savoir si la demande de logement impose des ouvertures de nouvelles zones constructibles, étant donnés les paramètres urbains choisis, ou à l'inverse, si les zones actuellement constructibles sont suffisantes pour accueillir ces nouveaux logements. La figure 3 représente le résultat d'une de ces simulations. Dans le cas où des POS/PLU sont modifiés, on peut également cartographier les parts de surface consommées en zones agricoles et en zones naturelles. La figure 4 permet d'identifier les lieux où des pressions particulière peuvent apparaître, et ainsi répondre à ce type de question qui préoccupe de nombreux élus. Enfin, et c'est tout l'intérêt du modèle, il est possible de spatialiser les parcelles identifiées pour l'urbanisation nouvelle, étant donnée les paramètres urbanistiques choisis. La concentration spatiale des parcelles "ouverte à l'urbanisation" peut constituer un indicateur précieux pour identifier les zones potentiellement à forte pression foncière. Lorsque les paramètres urbanistiques sont modifiés pour que le modèle s'écarte d'une simple prolongation de tendance et simule des choix urbanistiques marqués, le modèle permet d'identifier les

Table 5: Résultats du modèle d'estimation des probabilités de conversion (1)

Variable	Coef	Ecart-Type	t value	P value
(Intercept)	-0.359	0.069	-5.195	0
as.factor(c111)1	-0.619	0.034	-18.42	0
as.factor(genniv2)GHAB	0.767	0.036	21.597	0
as.factor(genniv2)GNA	0.319	0.038	8.298	0
as.factor(genniv2)GNAE	0.19	0.041	4.644	0
as.factor(genniv2)GNB	0.856	0.041	20.968	0
as.factor(genniv2)GNC	-0.092	0.036	-2.567	0.01
as.factor(genniv2)GND	-0.12	0.039	-3.059	0.002
as.factor(zs)1	-3.088	0.025	-123.516	0
as.factor(zs)3	-3.195	0.015	-206.214	0
as.factor(zs)4	-3.423	0.02	-170.447	0
dc	-0.012	0	-62.564	0
I(as.numeric(dr < 15000) * dr)	0	0	-11.865	0
I(as.numeric(dr > 15000) * dr)	0	0	-10.81	0
I(as.numeric(dvill < 8000) * dvill)	0	0	5.891	0
I(as.numeric(dvill > 8000) * dvill)	0	0	13.87	0
I(as.numeric(nbemprisebatha < 1) * nbemprisebatha)	1.064	0.02	54.342	0
I(as.numeric(nbemprisebatha >= 1) * nbemprisebatha)	0.296	0.005	64.072	0
I(as.numeric(nblogha < 2) * nblogha)	0.183	0.009	20.448	0
I(as.numeric(nblogha >= 2) * nblogha)	-0.005	0.001	-7.579	0
Spline Variable	EDF	Red.EDF	F value	P value
s(x,y)	28.946	29	21074.473	0
R-sq.(adj) = 0.405 Deviance explained = 34.4 % UBRE score = -0.19177 Scale est. = 1 n = 422088				

URBANSIMUL – interface php v1

choix des paramètres de simulation

Annee de projection Ex: 2022	<input type="text"/>
Règle de projection de logements communaux	<input type="radio"/> : R1 : Le Choix_projection de logement d'une commune i est fonction de la part de population de cette commune au sein de l'ensemble des communes <input type="radio"/> : R2 : règle R1 + fonction de la surface disponible <input type="radio"/> : R31 : Fixez vous même les Choix_projections de population communale dans un fichier DBF <input type="radio"/> : R32 : Fixez vous même les Choix_projections de logements dans un fichier DBF
SI vous avez répondu R2, fixez alpha : ** Si alpha est proche de 0.8 on privilégiera la surface de terrain disponible à l'urbanisation nouvelle au sein des zonages constructibles ** Si alpha est proche de 1. on favorisera une répartition en fonction de la part de population	<input type="text"/>
Seuil S1 : niveau de croissance maximal de la population par an Il est défini par un percentile par classes de commune. Choisissez ce percentile	<input type="text"/>
Seuil S4 : estimation du nombre maximal de logements raisonnablement constructibles Il est défini par type de zone POS, par classe de commune et des densités observés dans le passé. Choisissez ce percentile	<input type="text"/>
as4 : coefficient permettant de favoriser une urbanisation nouvelle de densité proche de celles des parcelles voisines ** Si as4 est proche de 1, on prend S4 comme estimation du nombre maximal de logements raisonnablement constructibles ** Si as4 est proche de 0, l'estimation du nombre maximal de logements raisonnablement constructibles se réfère à la densité existante des parcelles voisines	<input type="text"/>
Type de scenario : Révision de POS/PLU	<input type="radio"/> : non La simulation se fera avec un scenario_type/PLU fixe <input type="radio"/> : oui La simulation se fera avec un scenario_type/PLU variable
Si vous avez choisi de faire varier les POS/PLU	
pdvill : favorise la continuité au noyau villageois	<input type="text"/>
pdr : favorise la proximité au réseau routier principal	<input type="text"/>
pdc : favorise la proximité au routes secondaires et aux chemins	<input type="text"/>
nblogha : favorise l'urbanisation dans les zones denses en logements	<input type="text"/>
pnbemprisebatha : favorise l'urbanisation dans les zones denses en éléments bâtis	<input type="text"/>
seuilNCND : il s'agit de la part maximale de zones agricoles (GNC) et naturelles (GND) pouvant être rognées par l'urbanisation nouvelle	<input type="text"/> 0%
Seuil bmin En cas de changement de POS/PLU, il sera impossible de construire moins de "bmin" logements par hectare. La valeur "2" pourrait être par défaut	<input type="text"/> 2
Seuil bmax En cas de changement de POS/PLU, il sera impossible de construire plus de "bmax" logements par hectare. La valeur "50" pourrait être par défaut	<input type="text"/> 50
<input type="button" value="Valider"/>	

L'utilisateur remplit un certain nombre de paramètres dans un formulaire

Figure 2: URBANSIMUL 1, Interface php.

zones que ces critères privilégient, et permet d'évaluer les conséquences spatiales qui en découlent. Les figures 5, 6 et 7 illustrent sur un même espace trois choix urbanistiques différents.

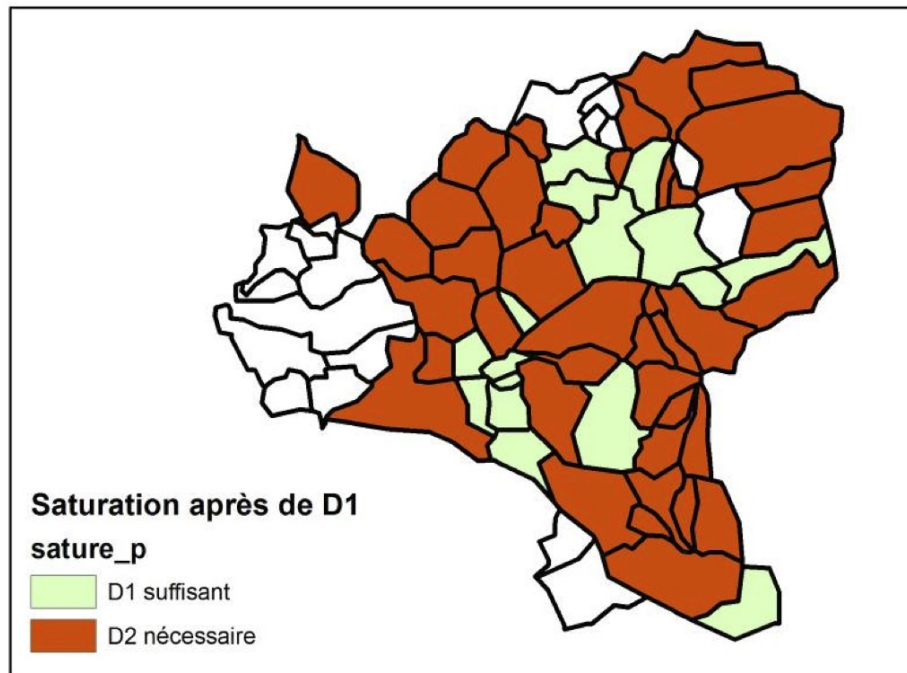


Figure 3: Simulation - ouverture POS/PLU.

4 Conclusion

Actuellement en phase de calibrage et de validation sur les trois SCOT ouest du Vaucluse, l'extension du modèle sur l'intégralité de la région PACA est en cours. Une version moins précise utilisant des mailles de 100 mètres par 100 mètres, a été développée et testée sur 2 régions PACA et LR. Cette version, dont la finalité était de projeter à 20 ans la pression foncière sur les zones ZNIEFF des deux régions a permis plusieurs développements méthodologiques (Vimal et al., 2011) qui seront repris dans les prochaines versions d'URBANSIMUL.

On prévoit plusieurs grandes évolutions du modèle. Dans la prochaine version, le modèle couvrira la région PACA, utilisera les dates de construction des parcelles pour mieux prendre en compte la dynamique récente d'urbanisation, notamment après la loi SRU. Le modèle logit d'estimation des probabilités de conversion intègrera de la dépendance spatiale¹⁵ avec une pondération en fonc-

¹⁵Version linéarisée de l'estimateur GMM pour modèle probit spatial, (Klier and McMillen, 2008)

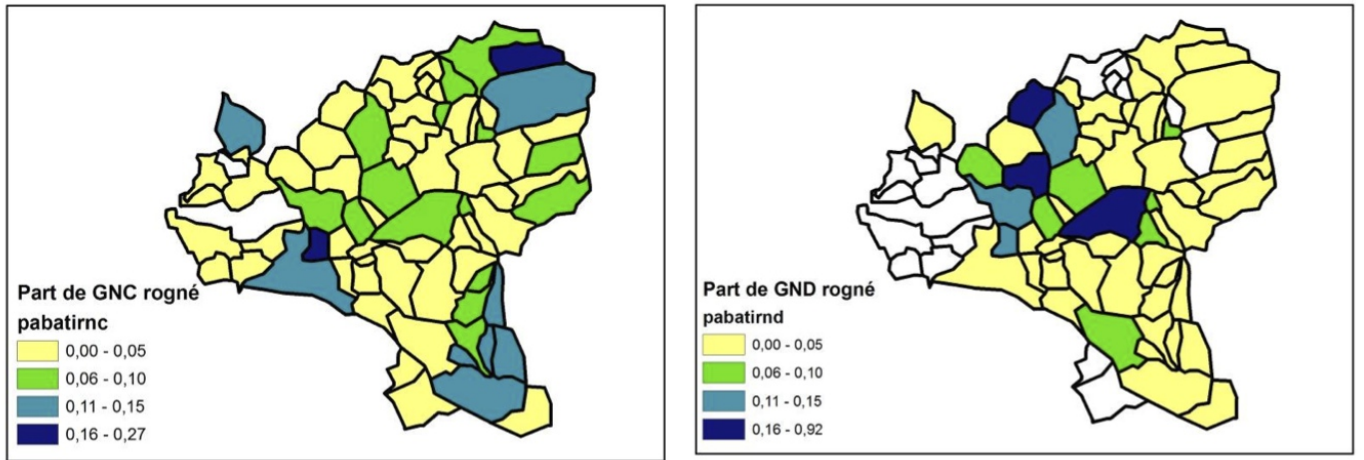


Figure 4: Simulation - NC-ND consommé.

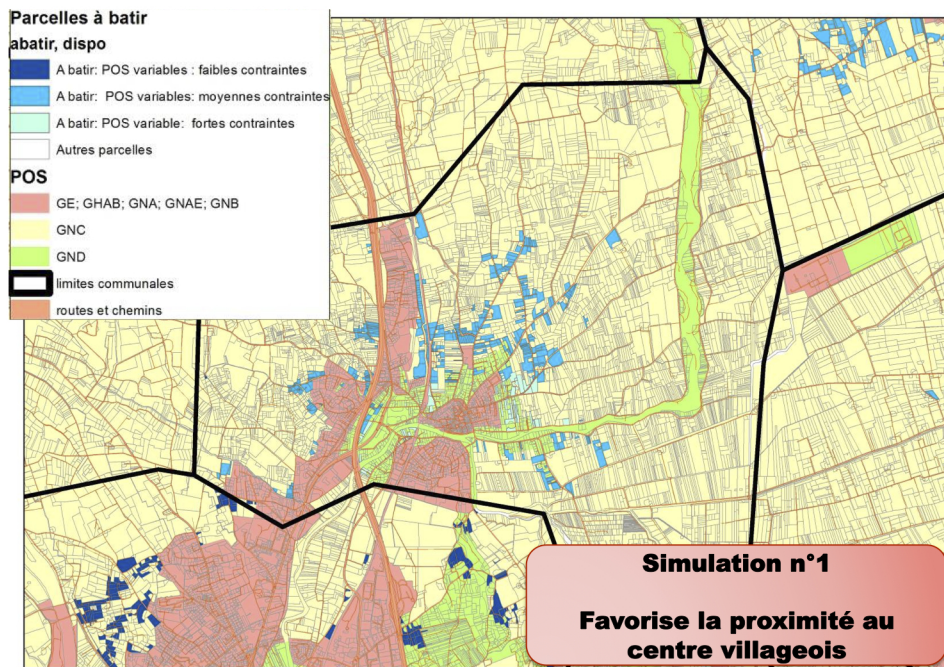


Figure 5: Simulation 1 - proximité au noyau villageois privilégiée

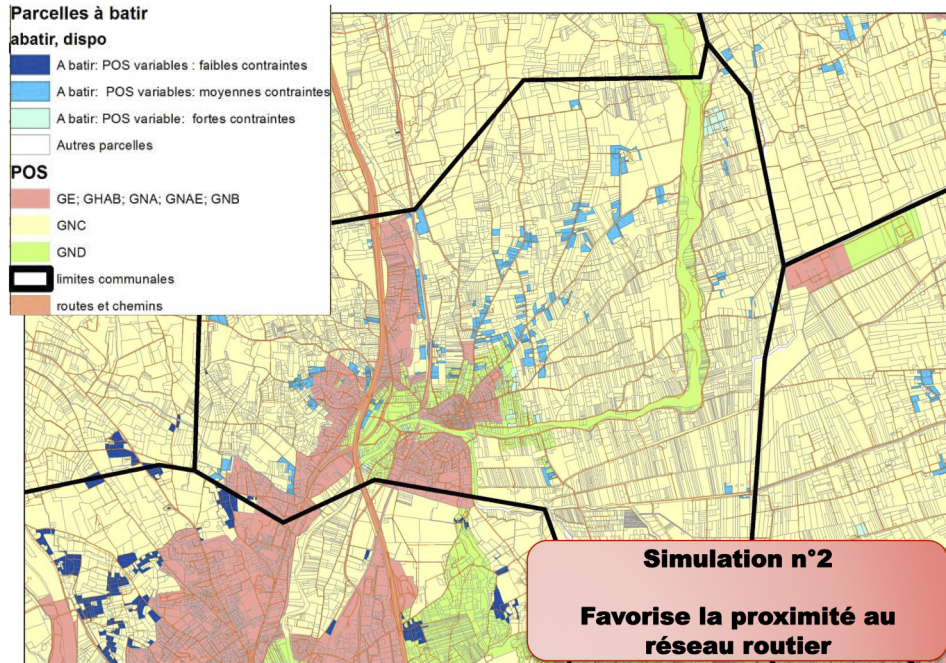


Figure 6: Simulation 2 - proximité au réseau routier privilégiée

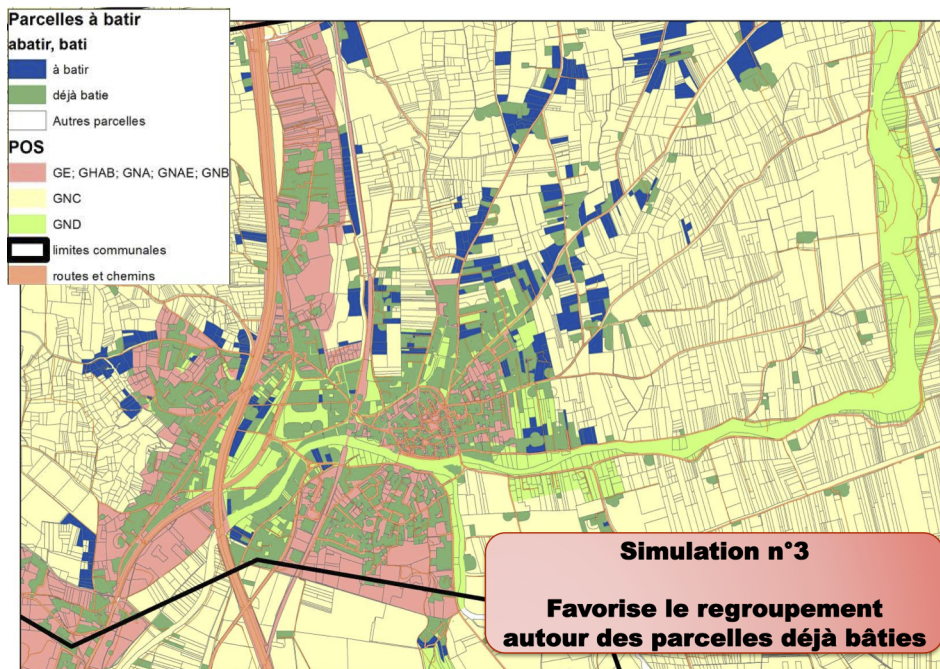


Figure 7: Simulation 3 - proximité aux parcelles déjà bâties privilégiée

tion de la date de conversion des parcelles. Les projections démographiques reprendront le modèle PLSR¹⁶ développé dans Vimal et al. (2011). La version 3 s'appuiera sur un modèle de durée avec de la dépendance spatiale visant à estimer les dates futures de construction plutôt qu'une probabilité qui demande à s'appuyer sur des informations externes au modèle. Une thèse en statistique spatiale débutée en 2010 à l'INRA d'Ecodéveloppement devrait contribuer à développer un cadre analytique adapté aux modèles de durée multi-états (avec état absorbant) avec dépendance spatiale.

Remerciements : *Je tiens à remercier Pascal Pluvinet, géomaticien de talent, pour avoir développé l'interface de ce logiciel et amélioré le code R et postgres/postgis du modèle de simulation. Je remercie également Johan Jouve pour avoir transposé sur MAPINFO une partie de la méthodologie de construction des données d'entrée d'URBANSIMUL 1 et ainsi rendu accessible à un plus large public la construction de l'indice de constructibilité théorique.*

References

Cho, S.-H and D. H. Newman. 2005. "Spatial analysis of rural land development", *Forest Policy and Economics*, 7 (5), 732 – 744, economic and Policy Perspectives of the Wildland-Urban Interface.

URL <http://www.sciencedirect.com/science/article/B6VT4-4G27SM5-1/2/c460aa05fe5064d7bb892>

Evans, A. 1999. "The land market and government intervention", *Handbook of Regional and Urban Economics*, 3, 1637–1669.

Geniaux, G, E. Dumas, C. Bartoli, P. Cezanne-Bert, and N. Napoléone. 2005. "Identification qualitative des espaces disponibles pour l'urbanisation nouvelle". Tech. Rep. Lot 7, rapport pour le Conseil régional Provence Alpes Côte d'Azur / Association CESSA.

Geniaux, G and N. Napoléone. 2005. "Rente foncière et anticipations dans le périurbain", *Economie et Prévision*, 168 (2), 77–95.

Geniaux, G, N. Napoléone, and J. Ay. 2011. "A spatial hedonic approach on land use change anticipations", *Journal of Regional Science* (En révision,).

Geniaux, G, P. Pluvinet, P. Cezanne-Bert, and J. Jouve. 2009a. "Identification qualitative

¹⁶Procédure de régression visant à minimiser l'erreur relative plutôt que l'erreur absolue.

- des espaces disponibles pour une urbanisation nouvelle. test sur 3 scot : Guide d'étape méthodologique". Tech. rep., Direction Régionale de l'Équipement Provence-Alpes-Côte d'Azur.
- Geniaux, G, P. Pluvinet, P. Cezanne-Bert, and J. Jouve. 2009b. "Identification qualitative des espaces disponibles pour une urbanisation nouvelle.test sur 3 scot : Simulations à moyen-long terme : méthodologie". Tech. rep., Direction Régionale de l'Équipement Provence-Alpes-Côte d'Azur.
- Irwin, E. 2010. "NEW DIRECTIONS FOR URBAN ECONOMIC MODELS OF LAND USE CHANGE: INCORPORATING SPATIAL DYNAMICS AND HETEROGENEITY*", *Journal of Regional Science*, 50 (1), 65–91.
- Irwin, E, K. Bell, and J. Geoghegan. 2003. "Modeling and managing urban growth at the rural-urban fringe: a parcel-level model of residential land use change", *Agricultural and Resource Economics Review*, 32 (1), 83–102.
- Kammann, E. E and M. P. Wand. 2003. "Geoadditive models", *Journal of the Royal Statistical Society C*, 52, 1–18.
- Klier, T and D. P. McMillen. 2008. "Clustering of auto supplier plants in the united states: Generalized method of moments spatial logit for large samples", *Journal of Business & Economic Statistics*, 26 (4), 460–471.
- McDonald, J. F and D. P. McMillen. 2004. "Determinants of Suburban Development Controls: A Fischel Expedition", *Urban Stud*, 41 (2), 341–361.
URL <http://usj.sagepub.com/cgi/content/abstract/41/2/341>
- Munneke, H. J. 2005. "Dynamics of the urban zoning structure: An empirical investigation of zoning change", *Journal of Urban Economics*, 58 (3), 455 – 473.
URL <http://www.sciencedirect.com/science/article/B6WMG-4H45GSX-2/2/3736a46f522819689851a>
- Vimal, R, G. Geniaux, P. Pluvinet, J. Leppart, and C. Napoléone. 2011. "Current and future urbanization impact on biodiversity in a mediterranean region" (In Prep. for Biological Conservation).
- Wood, S. 2006. *Generalized Additive Models: An Introduction with R*, ., Chapman and Hall/CRC.
- Wu, J and S.-H. Cho. 2007. "The effect of local land use regulations on urban development in the western united states", *Regional Science and Urban Economics*, 37 (1), 69–86.