La valeur des attributs naturels des terres agricoles de Côte d'Or

Jean-Sauveur Ay, Jean-Marc Brayer, Jean Cavailhes, Pierre Curmi, Mohamed Hilal, Marjorie Ubertosi

▶ **To cite this version:**

Jean-Sauveur Ay, Jean-Marc Brayer, Jean Cavailhes, Pierre Curmi, Mohamed Hilal, et al.. La valeur des attributs naturels des terres agricoles de Côte d'Or. 2012. hal-01873913

HAL Id: hal-01873913

https://institut-agro-dijon.hal.science/hal-01873913

Submitted on 13 Sep 2018

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.

UMR 1041 INRA – AGROSUP

CESAER

Centre d'Economie et Sociologie
appliquées à l'Agriculture et aux Espaces Ruraux

La valeur des attributs naturels des terres agricoles de Côte d'Or

Jean-Sauveur Ay, Jean-Marc Brayer, Jean Cavailhès
Pierre Curmi, Mohamed Hilal, Marjorie Ubertosi

Working Paper

2012/1

UMR 1041 CESAER INRA – AGROSUP
26, Bd Dr Petitjean - BP 87999 - 21079 Dijon cedex

La valeur des attributs naturels des terres agricoles de Côte-d'Or

Jean-Sauveur Ay,[1*] Jean-Marc Brayer,[2] Jean Cavailhès,[1]
Pierre Curmi,[2] Mohamed Hilal[1] et Marjorie Ubertosi.[2]

[1]INRA, UMR 1041 CESAER 21000 DIJON (FRANCE)
[2]AGROSUP, JE DIQUAS 21000 DIJON (FRANCE)

Résumé : L'importance économique des attributs naturels de la terre, sources d'avantages et de contraintes pour l'activité agricole, reste une question empirique ouverte. Elle est pourtant déterminante pour approcher la valeur productive de ce capital naturel. Cet article aborde la question par les relations hédoniques entre ces attributs et les prix observés lors des transactions foncières (Côte-d'Or, 1993–2005). Les attributs analysés sont de deux types : pédologiques (réserve en eau utile, hydromorphie, texture du sol, matière organique) et topographiques (altitude, pente, exposition). Leurs effets sur les prix sont appréhendés conjointement aux variables économiques plus classiques (taille du lot, distance au centre urbain le plus proche, mode de faire-valoir précédent). Suite à l'harmonisation et à la fusion des différentes sources d'information, nous obtenons que : (i) les attributs topographiques sont sensiblement plus valorisés que les attributs pédologiques, (ii) les relations empiriques sont non linéaires, avec d'importants effets croisés et (iii) la délimitation de paniers d'attributs évoluant de manière jointe entre les unités de terre permet d'obtenir des effets totaux non négligeables, de l'ordre de 35–60% des prix. Les pertes de précision issues de l'harmonisation des données et la multicolinéarité issue d'une grande quantité d'attributs naturels corrélés sont enfin discutées comme les principales limites de l'exercice.

Mots-clés : Prix des terres agricoles, modèle hédonique, attributs naturels.

Classification JEL : Q15, Q24.

1 Introduction

La notion de capital désigne en général un stock de matériel ou d'information disponible à un moment donné. Pour être ainsi caractérisé, ce stock doit avoir une valeur économique positive, c'est-à-dire générer – de manière autonome ou jointe – des services qui permettent

*Auteur correspondant, adresse : 26 Bd Dr Petitjean, 21000 Dijon, email : jsay@dijon.inra.fr. Cette recherche est issue d'une thèse financée par l'INRA et la région Bourgogne, elle s'inscrit également dans le cadre du programme GESSOL 3 (projet EcoSolHydro) du Ministère de l'Écologie, de l'Énergie, du Développement Durable et de la Mer.

d'accroître le bien-être humain. Par opposition au capital reproductible, l'évaluation du capital naturel (préexistant à l'homme et non reproductible) présente un regain d'intérêt ces dernières décennies (Barbier et al., 1990; Costanza et Daly, 1992; De Groot et al., 2003; Fenech et al., 2003). En particulier, cette notion est utilisée pour distinguer la durabilité faible de la forte (*weak and strong sustainability*) avec en trame de fond la question de la substituabilité entre les deux types de capitaux (Solow, 1974; Hartwick, 1977; Dietz et Neumayer, 2007; Stiglitz et al., 2009).

La valeur économique des composantes du capital naturel est alors un élément d'intérêt, sachant que la majeure partie des services qu'il rend est hors marché (Costanza et al., 1997; Bartelmusa, 2009). Le cas de la terre – ressource naturelle, immobile et hétérogène – est de ce point de vue particulier, puisque les fonctions qu'elle remplit font à la fois l'objet de valorisations privées et sociales. Lorsqu'une régulation favorise une fonction sociale non représentée dans les prix de marché, la valeur privée de la ressource ne peut pas être négligée. Que ce soit pour connaître les coûts d'opportunités (Zelek et Shively, 2003; Cameron et al., 2008) ou pour choisir la forme de l'action publique (Wu et al., 2001; Newburn et al., 2006), cette information est déterminante. Nous nous intéressons ici aux attributs naturels de la terre, sources d'hétérogénéité, et à leur valeur pour l'utilisation agricole de la terre. Questionner la dépendance de l'activité agricole aux conditions naturelles fournit une bonne illustration de la valeur privée d'un capital naturel et des relations qu'il entretient avec les autres types de capitaux (Fedoroff et al., 2010).

Nous cherchons donc à révéler les prix associés aux attributs naturels des terres agricoles pour le département de la Côte-d'Or dans la région Bourgogne (France). Nous utilisons la méthode hédonique sur des transactions foncières (achat et vente de terre) entre agents privés. Si le marché foncier est concurrentiel, le prix observé est la valeur présente nette du flux infini des rentes foncières. Ces rentes, perçues par le propriétaire de la ressource, sont également le rendement économique net du meilleur usage de la terre (Ricardo, 1817; Mendelsohn et al., 1994). Une telle approche duale possède deux avantages par rapport aux méthodes d'évaluation alternatives, principalement issues des connaissances agronomiques (Cassman, 1999; Andrews et Carrol, 2001; Wiebe, 2003). (i) Elle intègre implicitement les adaptations humaines, autant en terme d'inputs variables (fertilisant, matériel, travail) que de choix dans l'usage de la terre (grandes cultures, prairies, forêts) et (ii) elle contient les connaissances que les acteurs locaux ont sur les potentialités de la terre et sur l'avenir de sa rentabilité.

Dans notre analyse, des relations significatives apparaissent entre les attributs naturels et le prix de la terre, ce qui indique une dépendance rémanente de la production agricole aux conditions naturelles. Nous trouvons en particulier que les attributs topographiques (altitude et pente) ont des effets sensibles sur le prix des terres agricoles (25–30% pour un écart-type) alors que les attributs pédologiques (réserve en eau utile, hydromorphie, texture, épaisseur, teneur en matière organique) semblent moins déterminants (5–10% pour un écart-type). Ce résultat sur l'importance de la topographie, couplé à l'absence d'effets climatiques significatifs (températures et précipitations), est interprété comme issu de contraintes sur la combinaison du capital naturel avec le capital reproductible (matériel, compétences). D'un point de vue méthodologique, nous testons la capacité d'une variable synthétique sur les sols (la réserve en eau utile) à capter des relations complexes entre des données issues d'échelles différentes. Ce point est crucial étant données les fortes dépendances statistiques existantes entre les

attributs naturels et la nécessité d'en résumer l'information. Une approche originale est implémentée, qui consiste à utiliser les estimations pour calculer des variations de prix prédits lorsque les attributs de la terre évoluent conjointement. Ces variations sont évaluées à 35% du prix médian lorsqu'elles sont limitées aux attributs pédologiques alors qu'elles atteignent 60% pour l'ensemble des attributs naturels pour lesquels nous disposons d'information. Les modèles estimés permettent enfin de proposer une valeur agrégée (départementale) à ce capital naturel et de cartographier sa distribution entre les communes.

La suite de l'article s'organise comme suit. La section 2 présente brièvement les modèles empiriques et leur interprétation hédonique. La section 3 décrit les données utilisées ainsi que leur représentativité par rapport aux dotations foncières départementales. La section 4 discute les résultats économétriques et la section 5 présente l'approche en terme de prix prédits et de paniers d'attributs naturels. Les conclusions et perspectives que nous pouvons entrevoir à la suite de ce travail sont présentées dans la dernière section.

2 Modèles

L'analyse hédonique considère que le prix observé d'un bien est une source d'information sur les fonctions d'offre et de demande qui conduisent à l'équilibre sur son marché. Lorsque le bien échangé présente des combinaisons différentes des mêmes attributs qualitatifs, les variations de prix servent à identifier statistiquement les valeurs marginales de chacun des attributs. Dans la littérature, la méthode hédonique est l'outil privilégié pour étudier la formation du prix des terres agricoles (Miranowski et Hammes, 1984; Ervin et Mill, 1985; Palmquist et Danielson, 1989; Maddison, 2000). Son application à l'hétérogénéité de la terre (modifiable et non modifiable), en tant que facteur de production pour l'agriculture, repose sur les bases théoriques de Palmquist (1989). Les trois points importants ici sont que : (i) face à des attributs non reproductibles d'une ressource fixe, les choix optimaux en autres inputs n'ont pas à être explicités (avec des prix uniques, ce sont des fonctions implicites de l'hétérogénéité), (ii) les valeurs marginales obtenues doivent être contextualisées (par les technologies disponibles, les prix, les politiques agricoles) et (iii) cette théorie économique n'est pas contrainte par la forme des relations statistiques entre les attributs et leurs prix. Sur ce dernier point, Ekeland et al. (2004) présentent des arguments théoriques en faveur de fonctions de régression flexibles.

Notre travail repose sur des modèles semi-logarithmiques estimés par Moindres Carrés Ordinaires (MCO). Le prix reporté lors des transaction, P, est l'unique variable endogène. Sur les mêmes données économiques, Cavailhès et al. (2009) testent une transformation Box-Cox – plus générale – et trouvent une forme fonctionnelle similaire (techniquement, le λ est proche de zéro). La structure de base se note donc :

$$\ln P_i = \alpha + \mathbf{x}_i'\boldsymbol{\beta} + \mathbf{s}_i'\boldsymbol{\delta} + \varepsilon_i \,, \tag{1}$$

avec $\mathbf{x}_i$ et $\mathbf{s}_i$ respectivement les vecteurs des variables de contrôle et des attributs naturels pour une transaction i de terre. Les erreurs ε_i sont supposées sphériques pour une estimation MCO efficiente. Lorsqu'elles sont associées à une variable continue, les valeurs de $\boldsymbol{\beta}$ et $\boldsymbol{\delta}$ représentent des semi-élasticités, c'est-à-dire des variations en pourcentage du prix

consécutives à des variations en volumes de la variable explicative. La valeur marginale d'un attribut naturel continu k pour la transaction i est alors obtenue par différenciation partielle de (1), soit $\delta_k \times P_i$.

Pour une variation discrète quelconque d'une autre variable k (un changement de valeur d'une *dummy* ou un saut de $z \in \mathbb{R}^*$ unités d'une variable continue), les effets marginaux en pourcentage du prix s'écrivent :

$$\Delta_P^{\%}(z) = 100 \times \big(\exp(\widehat{\delta}_k \cdot z - \widehat{\sigma}_k^2/2) - 1 \big), \tag{2}$$

avec $\widehat{\sigma}_k^2$ l'estimateur MCO de la variance du coefficient associé à la variable k. Pour les variables discrètes, l'interprétation s'établit en posant $z = 1$. Cette écriture est issue de la critique de Halvorsen et Palmquist (1980) et de la solution proposée par Kennedy (1981). Les détails du calcul et des remarques complémentaires sont présentés en Annexe A, sachant que pour des coefficients estimés proches de zéro ou pour des variations discrètes faibles, les écarts avec une interprétation directe des coefficients sont souvent peu importants. Ces effets discrets sont simples à calculer une fois les coefficients et les écart-types estimés, bien que leurs distributions ne soient pas symétriques et les tests de Student plus appropriés pour évaluer la significativité de la variation discrète (van Garderen et Shah, 2002). Face à cette incertitude sur la significativité des effets économiques réels et lorsque les modèles estimés sont plus complexes (avec des effets polynomiaux et des interactions entre variables explicatives) cette approche en terme d'effets individuels va présenter des limites. Cela nous amène à considérer une méthode alternative d'interprétation.

À la suite des recommandations de King et al. (2000), Fox (2003) a développé une méthodologie (et les fonctions sous le logiciel R qui permettent de l'implémenter) pour interpréter des modèles linéaires généralisés – dont le modèle semi-logarithmique est un cas particulier – sur la base de prédictions et de représentations graphiques. La méthode consiste à simuler des matrices alternatives $\mathbf{X}^*$ et $\mathbf{S}^*$ qui servent à établir des vecteurs prédits de la variable endogène. Par exemple, pour le modèle semi-logarithmique (1), nous savons qu'un vecteur de prédictions non biaisées s'écrit (van Garderen, 2001) :

$$\widehat{\mathbf{P}}^* := \mathbb{E}\big(\mathbf{P} \,\big|\, \mathbf{X}^*, \mathbf{S}^*\big) = \exp\Big\{\widehat{\alpha} + \mathbf{X}^*\widehat{\boldsymbol{\beta}} + \mathbf{S}^*\widehat{\boldsymbol{\delta}} + \widehat{\sigma}_\varepsilon^2 \,/\, 2\Big\}, \tag{3}$$

avec $\mathbb{E}$ l'opérateur d'espérance conditionnelle et $\widehat{\sigma}_\varepsilon^2$ la variance estimée des erreurs. Les deux matrices simulées peuvent contenir les valeurs moyennes ou médianes des explicatives (donc des constantes) excepté pour une ou plusieurs colonnes qui vont décrire de manière incrémentale les supports des variables qu'elles représentent. Les matrices simulées $\mathbf{X}^*$ et $\mathbf{S}^*$ ont le même nombre de colonnes que les matrices empiriques $\mathbf{X}$ et $\mathbf{S}$. Les colonnes variables des matrices simulées vont entraîner des variations dans les prédictions de prix qui vont illustrer, toutes choses égales par ailleurs, les effets des explicatives correspondantes. Lorsque des effets polynomiaux ou d'interaction sont présents, l'ensemble des colonnes de $\mathbf{X}^*$ ou $\mathbf{S}^*$ inclues dans les polynômes ou les interactions est variable. Le programme de Fox (2003) permet également de calculer des intervalles de confiance locaux aux prédictions, sur la base de la matrice de variance-covariance estimée et de la matrice empirique.

Cette méthode basée sur la prédiction sera importante sur deux points principaux (en plus de faciliter l'interprétation de modèles complexes) : (i) en présence de multicolinéarité,

il est connu que les effets marginaux perdent de la précision mais les prédictions restent non biaisées (Belsley, 1991) et (ii) en simulant la matrice $\mathbf{S}^*$ sur la base des variations jointes observées en attributs naturels, nous allons reproduire des paniers d'attributs tels qu'ils sont présents dans les dotations en terre pour ainsi prédire leurs effets globaux sur les prix.

3 Données

3.1 Transactions foncières

Les données sur les transactions foncières sont fournies par *Terres d'Europe – SCAFR*.[1] Ces informations sont issues des notifications notariales précédant de manière obligatoire la vente d'un terrain dès lors que certaines conditions sont remplies, principalement une surface minimale.[2] La période considérée est 1993–2005, avec une couverture de l'ensemble du département de la Côte-d'Or (8 000 km^2 au total dont environ 4 500 sont en usage agricole). Les prix sont observés au niveau des transactions, qui contiennent souvent plusieurs parcelles. Ce sont pourtant les parcelles qui sont identifées grâce au cadastre numérique pour être reliées aux autres sources d'information que nous utilisons. La base compte 19 500 parcelles à destination agricole certaine (selon la nomenclature SCAFR) pouvant être localisées.

Nous retirons les parcelles de vignes et les parcelles situées dans les communes dont l'orientation agricole dominante est viticole (2 241 parcelles). Cette réduction de l'échantillon se justifie par les particularités de ce segment de marché, tout en sachant que les observations retirées se concentrent sur une petite zone : la côte dijonnaise. L'absence d'information sur l'âge, la qualité et le cépage des vignes nous empêche de capter de manière satisfaisante les déterminants des prix, dans un échantillon qui cherche à représenter l'usage agricole dans son ensemble. Sur les 17 259 parcelles restantes, nous obtenons 7 453 transactions soit une moyenne de 2,3 parcelles par transaction. La suppression des observations pour lesquelles toutes les variables ne sont pas renseignées ou qui contiennent des valeurs extrêmes nous conduit à un échantillon final de 4 254 transactions réparties sur 619 communes (88% des 707 communes du département).

Toutes les informations de la base SCAFR sont disponibles au niveau de la transaction foncière. Les tableaux 4 et 5 de l'Annexe C en présentent respectivement les intitulés et les statistiques descriptives. Le prix moyen d'une transaction est de 6 947 euros pour une superficie moyenne de 3,14 hectares (ha). Cette relativement faible superficie moyenne indique que le marché est surtout celui de l'agrandissement des exploitations. Moins de 3% des transactions portent sur plus de 15 ha de terre (la transaction qui regroupe le plus de surfaces en compte 36,3 ha). Les prix moyen et médian à l'hectare s'établissent respectivement à 2 223 euros/ha et 1 743 euros/ha. Toujours au niveau des transactions, la base contient une série de variables que nous considérons comme discrètes. L'année de vente, avec une

1. Sociéte Centrale d'Aménagement du Foncier Rural.

2. Selon les régions, voir les départements, et selon l'usage de la terre, la notification est obligatoire dès les premiers m^2 (notamment lorsque la vente concerne des zones agricoles ou naturelles d'un plan local d'urbanisme) ou à partir de quelques ares. Par exemple, en Bourgogne, la notification est obligatoire à partir de quatre ares pour une zone viticole en Apellation d'Origine Contrôlée, 20 ares pour des terres de cultures maraîchères et 50 ares pour des terres en polyculture-élevage.

ventilation annuelle de 246–424 transactions, la nature cadastrale[3] et le mode de faire-valoir sont utilisés comme contrôles dans les régressions. Pour 47% des transactions, l'acheteur est l'ancien fermier alors que pour 38% les terres n'étaient pas louées avant leur vente (elles sont dites libres). Seulement 25% des terres louées sont rachetées par une personne extérieure au contrat préalable, ce qui peut être expliqué par le droit de préemption du fermier sortant en cas de vente de la terre (une spécificité française).

Des variables de contrôle externes à la base SCAFER sont également utilisées. Nous calculons trois séries de distances. La distance à la mairie de la préfecture du département ($DIJON$) et la distance par rapport au pôle de bassin de vie le plus proche[4] sont des distances kilométriques par les voies de circulation (Hilal, 2010) alors que la distance de la parcelle au chef-lieu de la commune d'appartenance (qui est souvent au centre du noyau bâti) est évaluée à vol d'oiseau. Ces variables de distance sont utilisées pour capter les anticipations formulées sur la probabilité de conversion vers un usage non agricole (Chicoine, 1981; Plantinga et al., 2002; Cavailhès et Wavresky, 2003; Geniaux et al., in press) ou l'influence de la ville sur la profitabilité agricole (Livanis et al., 2006). Les transactions sont en moyenne à 32 km de Dijon, 12 km d'un pôle secondaire et 1,5 km du chef-lieu communal. Pour éviter que la distance à Dijon ne capte d'autres effets spatiaux, nous utilisons une *dummy* ($DUMDIJ$) qui prend la valeur "1" lorsque la transaction est à plus de 35 km de la ville, ce qui représente 67% de l'échantillon. Des données communales issues du recensement agricole de 2000 sont utilisées pour contrôler la présence d'un drainage ou d'un système d'irrigation, des informations que nous n'observons pas au niveau de la parcelle. Pour chaque transaction, les variables $DRAIN$ et $IRRIG$ représentent la part de la surface agricole utile communale qui est drainée (moyenne de 11%) ou irriguée (6%).

3.2 Attributs naturels

L'information sur les sols est issue du Référentiel Pédologique de Bourgogne (RPB, Chrétien, 2000). Elle est agrégée par un Système d'Information Géographique (SIG) sur la base de données géoréférencées où sont délimités 101 pédopaysages (ou Unités Cartographiques de Sols, UCS). Une UCS délimite une zone géorgraphique pour laquelle le nombre de types de sols et leur répartition dans le paysage sont les mêmes. Les données géométriques sont complétées de données sémantiques qui décrivent les caractéristiques des UCS. Toute UCS contient entre un et 10 types de sols différents (de trois à cinq le plus fréquemment) chacun constitué d'une à cinq strates (entre deux et quatre le plus fréquemment) de différentes épaisseurs. La distribution spatiale des différents sols au sein d'une UCS n'est pas connue mais nous savons quelles proportions surfaciques ils représentent. La dernière source d'information utilisée est issue d'un Modèle Numérique de Terrain (MNT) au pas de 50 mètres. Ces données sont sous forme *raster* avec une précision suffisante pour obtenir les principaux

3. Cette variable détermine le régime d'imposition. Elle est approximative, les erreurs de mesures sont nombreuses et principalement attribuables à sa faible actualisation. L'item $TERRE$ (58%) représente les terres labourées, $PRES$ (20%) les prairies temporaires ou permanentes, $TPRE$ (10%) pour une transaction qui porte à la fois sur des terres et des prés, $PANA$ (8%) lorsqu'aucune nature cadastrale ne se dégage et $BOIS$ (5%) pour les transactions de surfaces forestières à destination agricole.

4. Selon l'INSEE, il y a 18 bourgs ou petites villes animant un bassin de vie dans le département. Ils offrent des commerces et services intermédiaires, publics et privés, que les habitants fréquentent régulièrement.

indices orographiques que nous utilisons pour caractériser la topographie des parcelles : l'altitude, la pente et l'exposition.

Ces deux sources sur les attributs naturels sont superposées aux parcelles du cadastre pour être ensuite agrégées au niveau des transactions. Nous opérons pour cela une pondération à deux étages. Le SIG assigne à chaque parcelle qui a fait l'objet d'une transaction sur la période les calculs issus du MNT et une ou plusieurs UCS. Toujours par SIG, nous obtenons les surfaces (et donc les proportions) des jointures entre les deux types de polygones : les parcelles et les UCS. Ces proportions servent de coefficients de pondération pour ventiler les valeurs pédologiques au niveau des parcelles. Les parcelles d'une même transaction sont ensuite agrégées. La moyenne pondérée (par les parts de surface qu'elles représentent) s'est révélée une méthode opérationnelle. Nous obtenons ainsi une correspondance entre un prix (celui de la transaction) et des caractéristiques (celles des parcelles qui la composent). L'imprécision issue de ces manipulations est source d'erreurs de mesure sur les variables pédologiques, des erreurs qui engendrent des biais indéterminés dans leurs directions (Wooldridge, 2002; Dormont, 2007).

Les données pédologiques contiennent la texture des sols (la transaction moyenne contient 33,5% d'argile, 45,6% de limon et 20,9% de sable), le taux d'éléments grossiers (TEG : graviers, cailloux, 11% en moyenne), le taux de matières organiques (TMO : 21,3‰) et l'épaisseur ($EPAI$: 53 cm). En plus de ces informations standards des référentiels pédologiques, la Réserve en eau Utile (RU : 86 mm en moyenne) et des classes d'hydromorphie (HYD) sont disponibles.

La variable RU correspond au volume d'eau que le sol est susceptible de stocker afin de le rendre utilisable par les végétaux. Cette variable est calculée sur la base des informations standards précédentes et de valeurs calibrées par l'expérimentation, disponibles dans Bruand et al. (2004). L'Annexe B présente plus précisément le calcul. RU est une variable synthétique *a priori* valorisable par l'agriculture et qui a déjà été utilisée pour approcher la qualité de la terre (Caswell et Zilberman, 1986; Lichtenberg, 1989; Wu et Segerson, 1995). Son information est cependant (faiblement) redondante aux données classiques du référentiel pédologique. L'hydromorphie indique la présence de marques physiques de périodes d'anoxie du sol plus ou moins prolongées. Elle signifie en général la stagnation de masses d'eau dans le sol, ce qui est *a priori* préjudiciable à un usage agricole. La variable HYD prend des valeurs discrètes de 0 à 3 selon la profondeur d'apparition et l'intensité des traces d'hydromorphie (un sol est jugé d'autant plus hydromorphe que les traces sont superficielles). Pour conserver la nature discrète de la variable (qui a été considérée comme continue dans l'agrégation) nous séparons ses valeurs en trois catégories : sols Non Hydromorphes (NH : 36,6%), sols Peu Hydromorphes (PH : 33,2%) et sols Très Hydromorphes (TH : 30,2%) selon que HYD soit égale à 0, strictement entre 0 et 3 ou égale à 3.

L'altitude des transactions est comprise entre 112 et 658 mètres, pour une moyenne de 320 m. Les pentes se distribuent de 0° à 55,8° avec une moyenne à 5,9°. Toutes choses égales par ailleurs, la faible accessibilité et les contraintes pour la mécanisation sont les deux implications agricoles principales de la topographie. Elles plaident donc pour des effets négatifs *a priori*. Ces attributs naturels présentent entre eux des dépendances statistiques fortes, comme le montre le Tableau 1 des corrélations de Pearson. La plupart des coefficients sont significatifs à 1%, seules les dépendances entre l'hydromorphie et la pente et entre la

pente et le taux de matières organiques ne le sont pas à ce seuil. Il apparaît en particulier que *RU* présente de fortes corrélations (positives et négatives) avec le reste des attributs naturels, alors que les autres corrélations sont globalement plus faibles.

Table 1 – Corrélations entre les attributs naturels du foncier, *N = 4 254.*

	RU	**HYD**	**ALTI**	**PENTE**	**EPAI**	**TMO**
RU	1,0000					
HYD	+ 0,73**	1,0000				
ALTI	− 0,44**	− 0,20**	1,0000			
PENTE	− 0,35**	− 0,03*	+0,44**	1,0000		
EPAI	+ 0,87**	+ 0,68**	− 0,28**	− 0,20**	1,0000	
TMO	− 0,24**	− 0,21**	+ 0,15**	+ 0,04*	− 0,19**	1,0000

*** : Significativement différent de zéro à 1%* et ** pour 5%.*

3.3 Représentativité

Les observations de la base SCAFR sont en principe exhaustives pour les transactions mais proviennent néanmoins d'un tirage implicite lié à la nécessité d'une vente pour observer un prix. Ce phénomène peut être à l'origine d'un biais de sélection, si la probabilité d'une vente est corrélée avec des variables d'intérêt. À l'inverse, les données naturelles sont disponibles pour l'ensemble du département. Cette différence de structure nous permet de questionner la représentativité des attributs de l'échantillon SCAFR par rapport aux dotations départementales. Nous nous limitons cependant à une approche descriptive, sachant qu'une analyse rigoureuse nécessiterait l'estimation d'un modèle de sélection (Heckman, 1979). Une telle alternative nécessiterait tout de même l'estimation d'une première étape sur plus de 900 000 parcelles, avec la question en suspend du calcul des inverses de ratio de Mills au niveau des transactions.

L'information sur les textures des horizons de surface est retravaillée par la définition de classes de valeurs pour les pourcentages d'argile, de limon et de sable. Ces trois variables se somment à l'unité pour chaque transaction et peuvent donc être représentées dans un triangle de texture, avec la classification de l'Aisne en tant que typologie de référence (Jamagne et al., 1977). Pour disposer de classes de textures relativement homogènes avec suffisamment d'observations, nous regroupons les 15 classes de textures initialement présentes en cinq classes agrégées, voir la Figure 5 de l'Annexe D. Les classes AL (argile limoneuse : 1 403 obs.), LA (limon argileux : 802 obs.) et ALO (argile lourde : 679 obs.) sont maintenues. Une première classe agrégée dite *équilibrée* ($EQUI$) regroupe les limons sablo-argileux (LSA), les limons argilo-sableux (LAS), les limons sableux (LS) et les limons moyens sableux (LMS) pour un total de 664 observations. L'autre classe que nous définissons, dite *sable et argile* (SSA, 706 obs.), regroupe les classes génériques argile (A), sable (S), sable argileux (SA), limon sablo-argileux (LSA) et sable limoneux (SL). Trois classes de textures du triangle de l'Aisne ne sont pas présentes dans l'échantillon (comme en Côte-d'Or) et ne sont donc pas considérées dans ce découpage. Il s'agit du limon moyen (LM), du limon léger (LL) et du limon léger sableux (LLS).

Questionner la représentativité de l'échantillon nécessite également la connaissance de l'usage de la terre au niveau du département, afin de limiter la comparaison aux terres agricoles similaires à la nomenclature *SCAFR*. Nous utilisons pour cela Corine Land Cover (CLC en 2000) où sont considérés comme terres à destination agricole : les *terres arables hors périmètres d'irrigation*, les *périmètres irrigués en permanence*, les *prairies*, les *cultures annuelles associées aux cultures permanentes* et les *systèmes culturaux et parcellaires complexes*. Le Tableau 2 permet de caractériser les dotations en attributs pédologiques et topographiques des terres agricoles du département et de les comparer à celles présentes dans l'échantillon.

Table 2 – Représentativité des transactions SCAFR : pour chaque classe de texture, les valeurs pour l'ensemble du département (DEP.) et pour les transactions (TRA.) sont reportées, ainsi que le taux de sondage implicite (REP.) et les écarts entre les deux populations (ECART).

	Surface (milliers d'ha)			Réserve Utile (mm)		
Textures[a]	DEP.	TRA.	REP.	DEP.	TRA.	ECART
AL	205,4	5,66	2,7%	53,56	57,25	+ 6,9%
LA	90,06	2,30	2,6%	141,5	137,8	− 2,6%
ALO	97,90	1,83	1,9%	60,10	73,77	+ 22,7%
EQUI	51,93	1,97	3,8%	122,4	114,3	− 6,6%
SSA	52,59	1,61	3,1%	51,83	57,33	+ 10,6%
TOTAL	497,9	13,4	2,7%	77,76	81,81	+ 5,2%

[a] **AL** : *argile limoneuse* ; **LA** : *limon argileux* ; **ALO** : *argile lourde* ; **EQUI** : *limon sablo-argileux (LSA), limon argilo-sableux (LAS), limon sableux (LS) et limon moyen sableux (LMS)* ; **SSA** : *argile (A), limon sablo-argileux (LSA), sable (S), sable argileux (SA) et sable limoneux (SL)*. Voir Figure 5, Annexe D.

Les données SCAFR contiennent au total 13 400 ha, soient 2,7% des 497 900 ha de terres agricoles du département.[5] Pour chaque classe de texture, les trois premières colonnes du tableau présentent les taux de sondage implicites. Le département compte 205 400 ha d'argiles limoneuses et l'échantillon en reporte 5 600, soit un taux de sondage identique à celui de l'ensemble de l'échantillon (2,7%). Les argiles lourdes sont sous-représentées dans l'échantillon (1,9% des *ALO* du département) et les textures dites *équilibrées* et *sableuses et argileuses* sont sur-représentées (respectivement 3,8% et 3,1%). Les trois dernières colonnes du tableau présentent les valeurs moyennes de *RU* pour le département et l'échantillon. Malgré une représentativité surfacique, la classe de texture qui compte le plus d'observations dans l'échantillon (*AL*) présente une moyenne de *RU* supérieure de près de 7% à la moyenne du département. La probabilité d'une vente sur les textures *ALO* et *SSA* est également positivement corrélée avec *RU*. L'inverse se retrouve pour les autres classes *LA* et *EQUI* alors qu'au niveau départemental la corrélation est positive. L'échantillonnage s'éloigne donc, pour

5. Le manque de précision des données *land cover* entraîne une sur-estimation des grands espaces contigus tels que les champs agricoles. Ce biais est issu de la sous-estimation des espaces fragmentés, voir Irwin et Bockstael (2007). Le recensement agricole de 2000 présente une quantité de terre agricole d'environ 450 000 ha. La différence des nomenclatures, moins détaillée pour CLC, est également une source de décalage qui va dans le même sens.

certaines variables, d'un tirage représentatif dans la population de base. Dans l'optique de quantifier la valeur agrégée de ce capital naturel, ces résultats plaident pour la nécessité de contrôler le biais de sélection, ce qui est laissé pour une recherche future.

4 Résultats

Cette section présente les estimations de trois modèles semi-logarithmique (M1, M2 et M3). Les modèles se différencient par les relations spécifiées entre les attributs naturels (linéaires, polynômiales et en interactions) et la variable endogène. Les différentes spécifications se retrouvent dans le Tableau 3 avec une complexité croissante des fonctions de régression. Pour l'interprétation, nous distinguons les variables de contrôle des variables naturelles d'intérêt.

4.1 Variables de contrôle

Le haut du Tableau 3 présente les coefficients et les écarts-types associés aux variables de contrôle. Ils présentent une certaine stabilité entre les spécifications. Le pouvoir explicatif des différents modèles est important, avec des R^2 supérieurs à 80%. L'élasticité prix-surface est unitaire, ce qui signifie un prix à l'hectare décroissant avec la taille des transactions. La distance au chef-lieu de la commune ($DISTMA$) présente un effet négatif, avec une diminution de plus de 5% du prix pour un éloignement d'un kilomètre supplémentaire. La distance au pôle du bassin de vie (DBV) n'est pas significative alors que le prix diminue de 2,4% par kilomètre d'éloignement de Dijon ($DIJON$). Au delà de 35 km, les prix sont en moyenne inférieurs de 11% ($DUMDIJ$). Lorsque l'ancien fermier est l'acquéreur des terres, il les achète en général à un prix 4,3% inférieur à celui d'une terre libre, la modalité de référence. À l'inverse, le prix est supérieur d'environ 5,8% lorsque des terres en fermage sont rachetées par un tiers. Le pourcentage de la surface agricole communale irriguée en 2000 ($IRRIG$) a un effet significatif et positif (+ 3,1% pour une augmentation de 10 points de pourcentage), alors que celui de terres drainées ($DRAIN$) a un effet négatif (− 1,9% pour 10 points de plus). Bien que stable en ordre de grandeur, cette dernière variable voit son coefficient augmenter avec la complexité des modèles.

Les coefficients associés à la nature cadastrale (NC) indiquent que seuls les terrains classés en bois ($NC : BOIS$) sont significativement moins coûteux que les terres labourables (la catégorie de référence) pour les trois spécifications : − 49% en moyenne. Ces estimations multivariées n'attribuent pas de différences de prix significatives entre les natures $TERRE$ et $PRES$ alors que dans l'échantillon, les terres labourables ont un prix moyen à l'hectare de 200 euros supérieur aux prés. Cette différence des prix moyens est donc issue des autres variables du modèle. Les coefficients des indicatrices annuelles ne sont pas reportés mais nous obtenons des prix significativement supérieurs à l'année 1993 de référence en 2000 et sur la période 2002–2005. En cumulé, les prix en euros courants s'établissent comme supérieurs de 49% en 2005, par rapport à 12 ans auparavant.

4.2 Variables naturelles

Étant données les corrélations internes fortes entre les variables naturelles (Tableau 1), des diagnostics de multicolinéarité sont effectués systématiquement. La multicolinéarité n'est pas toujours reliée aux corrélations entre explicatives, en particulier parce que notre échantillon est d'une grande taille (Fox et Monette, 1992). Nous présentons dans cette sous-section les résultats des spécifications qui ne sont pas multicolinéaires (c'est-à-dire avec des facteurs d'inflation de la variance inférieurs à 10 pour toutes les variables). Les calculs des *condition index*, des décompositions factorielles et une analyse par perturbation (Belsley, 1991; Hendrickx et Pelzer, 2004) ont également été effectués, sans changement majeur dans les diagnostics.

Le modèle M1 est le plus restrictif sur les relations car chaque variable est supposée avoir un effet linéaire et additif sur le logarithme du prix. Les coefficients obtenus sont significatifs, avec des attributs naturels qui influencent sensiblement le prix de la terre. Les ordres de grandeurs sont cependant différents entre les variables pédologiques et les variables topographiques dont les directions sont conformes aux intuitions formulées *a priori*. L'effet de RU est positif et significatif (probabilité limite de 3,9%) mais d'une faible amplitude : une augmentation d'un écart-type de la variable (46 mm) provoque une hausse de 2,8% du prix de la terre. L'effet de l'hydromorphie est négatif, bien que non monotone entre les classes. Les sols peu hydromorphes (PH) sont 9,4% moins coûteux que les non hydromorphes (modalité de référence). La baisse s'établit à un niveau moindre ($-5{,}7\%$) pour les sols très hydromorphes (TH). Les variables topographiques ont des effets plus importants. Augmenter d'un écart-type (112 m) l'altitude moyenne d'une transaction revient à diminuer son prix de 20%. Pour la pente, une augmentation de 6,1° entraîne une diminution de 6% du prix. Deux des trois variables *dummies* représentant l'exposition sont significatives à 5%. Nous trouvons que des expositions Ouest ou Nord tirent les prix vers le bas de respectivement 4,6% et 6% par rapport à une exposition Sud (référence). Être exposé à l'Est n'admet pas de différence significative de prix par rapport à la référence.

Des formes polynomiales[6] sont ensuite introduites pour les variables continues RU, $ALTI$ et $PENTE$ (M2). L'ordre des polynômes est déterminé par leurs significativités individuelles : nous arrêtons le développement lorsque l'ordre à venir n'est plus significatif. L'effet de RU apparaît en polynôme de degré 3, celui de $ALTI$ de degré 5 et celui de $PENTE$ de degré 2. Pour interpréter les résultats plus aisément, la Figure 1 suivante présente graphiquement ces effets par des prédictions simulées (Fox, 2003). Les prédictions sont obtenues en fixant la variable $SURF$ à l'unité (pour avoir des prix à l'hectare) et les autres explicatives à leurs médianes. Pour chacun des trois graphiques, les données simulées n'admettent des variations que pour respectivement les colonnes RU, $ALTI$, $PENTE$ et leurs ordres supérieurs, comme cela est expliqué dans la section 2.

Le logarithme du prix apparaît négativement relié à RU pour des valeurs relativement basses (moins de 50 mm : 27% des transactions) et des valeurs relativement hautes (plus de 130 mm : 23%). Entre 50 et 130 mm, la relation est positive et cette hausse de 80 mm

6. Pour éviter les problèmes numériques liés à l'estimation de polynômes d'ordres importants, les régressions utilisent la méthode des polynômes orthogonaux (Kennedy Jr et Gentle, 1980) présente dans la fonction `poly()` du logiciel R.

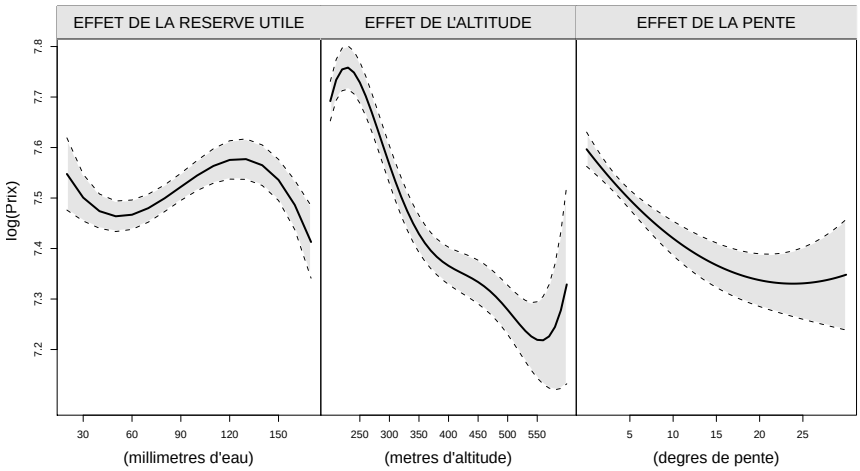
Table 3 – Les coefficients (Coeff.) et écarts-types (E.T.) estimés par MCO, modèles M1, M2 et M3 : le logarithme naturel du prix en euros courants est la variable expliquée, la transaction foncière l'unité d'observation. Les *dummies* annuelles ne sont pas reportées.

	M1		M2		M3	
Variables	**Coeff.**	**E.T.**	**Coeff.**	**E.T.**	**Coeff.**	**E.T.**
$CONSTANTE$	+ 8,985**	*0,084*	+ 8,318**	*0,081*	+ 8,236**	*0,084*
$log(SURF)$	+ 1,007**	*0,007*	+ 1,007**	*0,007*	+ 1,009**	*0,007*
$DISTMA$	− 0,054**	*0,010*	− 0,053**	*0,010*	− 0,052**	*0,001*
DBV	− 0,001	*0,001*	− 0,000	*0,001*	+ 0,000	*0,001*
$DIJON$	− 0,030**	*0,002*	− 0,024**	*0,002*	− 0,024**	*0,002*
$DUMDIJ$	− 0,172**	*0,028*	− 0,120**	*0,030*	− 0,127**	*0,030*
$FERM : OUI$	− 0,051*	*0,018*	− 0,044*	*0,018*	− 0,042*	*0,018*
$FERM : NON$	+ 0,056*	*0,025*	+ 0,060*	*0,024*	+ 0,059*	*0,025*
$IRRIG$	+ 0,259**	*0,057*	+ 0,326**	*0,059*	+ 0,321**	*0,059*
$DRAIN$	− 0,169*	*0,057*	− 0,178*	*0,057*	− 0,187*	*0,057*
$NC : PRES$	− 0,031	*0,022*	− 0,032	*0,023*	+ 0,036	*0,030*
$NC : TPRES$	+ 0,040	*0,030*	+ 0,052	*0,030*	+ 0,043	*0,030*
$NC : PANAT$	− 0,012	*0,034*	− 0,017	*0,034*	− 0,014	*0,034*
$NC : BOIS$	− 0,392**	*0,042*	− 0,410**	*0,042*	− 0,400**	*0,042*
RU	+ 0,001*	*0,000*	+ 0,407	*0,875*	+ 0,001*	*0,000*
RU^2	–	–	− 1,211*	*0,613*	–	–
RU^3	–	–	− 2,503**	*0,598*	–	–
PH	− 0,098**	*0,024*	− 0,115**	*0,026*	− 0,061	*0,054*
TH	− 0,058*	*0,029*	− 0,022	*0,030*	+ 0,132*	*0,061*
$ALTI$	− 0,002**	*0,000*	− 9,118**	*0,869*	− 8,555**	*0,858*
$ALTI^2$	–	–	− 0,581	*0,661*	− 1,187	*0,647*
$ALTI^3$	–	–	+ 3,458**	*0,569*	+ 3,456**	*0,570*
$ALTI^4$	–	–	− 2,403**	*0,553*	− 2,131**	*0,551*
$ALTI^5$	–	–	+ 2,566**	*0,540*	+ 2,392**	*0,539*
$PENTE$	− 0,010**	*0,002*	− 4,798**	*0,715*	− 5,712**	*0,704*
$PENTE^2$	–	–	+ 1,747**	*0,612*	+ 2,016**	*0,608*
EST	− 0,007	*0,020*	− 0,002	*0,020*	+ 0,000	*0,020*
$OUEST$	− 0,047*	*0,022*	− 0,045*	*0,022*	− 0,044*	*0,022*
$NORD$	− 0,060*	*0,026*	− 0,047	*0,027*	− 0,042	*0,026*
$RU \times PH$	–	–	–	–	− 0,001	*0,001*
$RU \times TH$	–	–	–	–	− 0,002**	*0,001*
R^2 Ajusté	0,838		0,842		0,845	
NOBS	4 254		4 254		4 254	
$\widehat{\sigma}^2_\varepsilon$	0,268		0,263		0,262	

******** *: Significativement différent de zéro à 1%* et ******* *pour 5%.*

entraîne une hausse de 13% du prix à l'hectare. L'effet négatif au-delà de 130 mm peut avoir des explications agronomiques et nous reviendrons en particulier sur sa relation avec l'hydromorphie. Les transactions à faibles RU ayant des prix relativement hauts sont plus difficiles à interpréter. Les terres les plus basses du département (34% des transactions sont à une altitude inférieure à 225 m) présentent des prix prédits à l'hectare de l'ordre de 2 290 euros, ce qui représente 131% de la médiane de l'échantillon. Entre 225 et 550 m d'altitude (65% de l'échantillon), la relation négative entre la pente et le logarithme du prix présente une rupture de pente aux environs de 350 m. L'effet de la variable $PENTE$ est plus monotone :

Figure 1 – Les effets des variables naturelles pour M2, intervalles de 5% :



passer de 3,5° à 7,5° entraîne une diminution de 8,8% du prix à l'hectare avec également une décroissance pour des valeurs plus à droite. La largeur de l'intervalle de confiance pour ces fortes valeurs s'explique par la faible proportion (1,3%) des transactions qui présentent des valeurs supérieures à 25°.

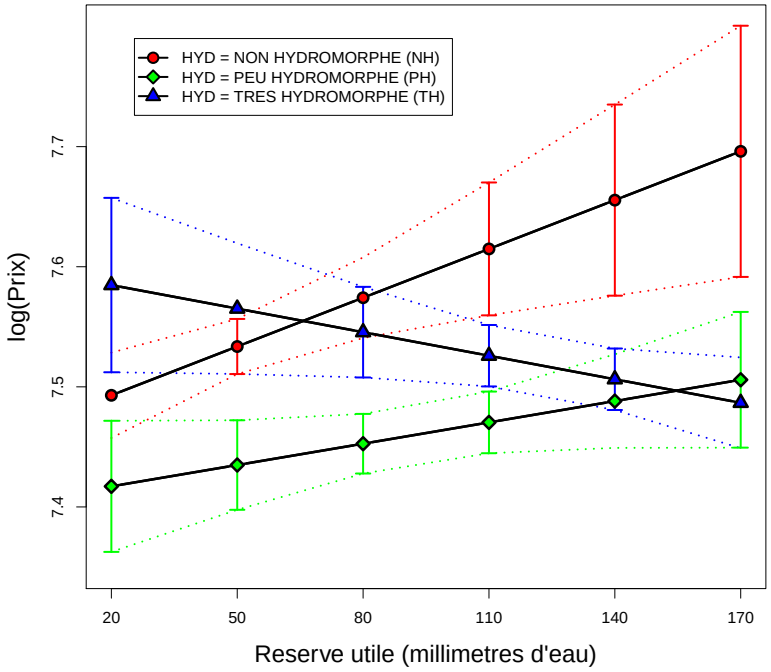
Ces effets négatifs importants de la topographie peuvent être issues de variables omises[7] mais nous pensons que les explications propres à ces variables ont du sens, en particulier par leurs rôles sur les possibilités de mécanisation ou d'accessibilité pour le travail. Bien que corrélées ($\rho = 0,44$: Tableau 1), ces deux variables topographiques présentent de faibles effets de colinéarité avec des facteurs d'inflation de la variance respectivement de 5 et 2,5. Les deux effets négatifs sont donc bien distincts et ne peuvent pas être expliqués par une seule variable omise (le climat).

Le modèle M3 contient une interaction entre RU et HYD au détriment du terme polynomial à l'ordre 3 sur RU. Les effets polynomiaux de $ALTI$ et $PENTE$ changent peu, de même que l'effet de l'exposition. Nous obtenons des effets d'interaction significatifs (par un test joint de Fisher) donc la valeur de RU est différente selon le degré d'hydromorphie des sols. L'effet de RU pour les sols NH augmente par rapport à l'effet identifié par M1. Il s'établit à environ 6,7% de hausse de prix pour un écart-type.[8] La disponibilité en eau possède donc relativement plus de valeur pour des sols non hydromorphes. Les effets propres de l'hydromorphie (lignes PH et TH) changent fortement par rapport à M1, le coefficient associé à TH devient positif, ce qui est contre-intuitif. La significativité du terme $RU \times TH$ indique que l'effet de RU est différent pour les sols très hydromorphes (TH), ce qui n'est pas le cas pour les sols peu hydromorphes (PH). Ainsi, augmenter RU d'un écart-type provoque une hausse non significative du prix des sols PH et une baisse significative de 4% du prix des sols TH. L'utilisation de l'estimation M3 sur données simulées permet d'être plus précis sur cette interaction et ce résultat contre-intuitif.

7. Nous avons intégré et testé une base communale sur le climat et les précipitations (Joly et al., 2009) sans que ses variables additionnelles entraînent une diminution des effets de la topographie.

8. Les arrondis du tableau empêchent de voir cette progression. Le coefficient de RU passe de 0,0006 à 0,0014.

Figure 2 – Interactions entre RU et HYD pour M3, intervalles de confiance à 10%.



La Figure 2 permet de comparer simultanément les effets volumes (en ordonnées) et les effets marginaux (les pentes) sous-jacents à l'interaction entre RU et HYD. Lorsque RU est à 110 mm, le modèle présente une hiérarchie statistiquement significative des catégories d'hydromorphie car les intervalles de confiance ne se chevauchent pas. Les sols NH sont les plus chers, suivis des sols TH puis des sols PH. Les sols NH sont significativement plus coûteux que les PH dès que RU dépasse 30 mm. Pour RU inférieur à 110, les sols TH sont significativement plus coûteux à l'hectare que les sols PH mais cette différence diminue avec RU. Pour de faibles valeurs de RU, la supériorité du prix moyen des sols TH relativement aux NH est étonnante mais pas significative. Cette hiérarchie se renverse aux alentours de 70 mm, les sol NH apparaissant significativement plus coûteux pour RU supérieure à 110 mm. La largeur des intervalles de confiance donne une idée de la quantité d'observations présentes localement. Ainsi l'échantillon contient peu de sols TH présentant de faibles valeurs de RU et peu de sols NH présentant de fortes RU (les intervalles sont larges). La présence d'un effet volume de l'hydromorphie sur les sols à faibles RU est probablement le résultat le plus marquant car il indique des relations de complémentarité/substituabilité non triviales entre la Réserve en eau Utile et l'hydromorphie. Il apparaît en effet une dévalorisation de RU sur les sols très hydromorphes.

La Figure 6 de l'Annexe D présente les interactions entre RU et la topographie. Les variables $ALTI$ et $PENTE$ sont discrétisées en quatre classes de valeurs pour simplifier l'interprétation. Les graphiques sont issus de deux régressions supplémentaires (M3a et M3b) dont les résultats bruts ne sont pas reportés ici mais qui sont disponibles sur demande. Des relations de substituabilité entre les effets volumes et les effets marginaux apparaissent également. Nous observons que les catégories de terre pour lesquelles RU est valorisé (c'est-à-dire qui admettent une droite de régression croissante) sont celles qui présentent des niveaux

de prix les plus bas. Des tests de nullité jointe concluent à un effet positif de RU pour les terres à plus de 300 m d'altitude ($M400$ et $P400$) et pour les terres qui présentent une pente supérieure à 8° ($FORTE$). Dans les cinq autres cas, le prix accordé à RU n'est pas significatif. La difficulté de substituer RU par de l'irrigation sur les terres en altitude et à fortes pentes (ruissellement et nappes phréatiques profondes) nous semble une explication possible de cette substituabilité entre les effets. Les variations de prix issues de RU sont toutefois faibles relativement aux effets volumes qui apparaissent entre les huit cases de la Figure 6. Cela renforce le constat d'un poids économique supérieur accordé à la topographie.

L'inclusion des autres variables disponibles sur les sols ($TARG$, $TLIM$, $TSAB$, TEG, TMO et $EPAI$) dans les modèles précédents s'effectue au prix d'une multicolinéarité et donc d'une perte en précision. Remarquons par ailleurs que le résultat principal (le faible poids relatif des caractéristiques pédologiques) est conditionné par deux hypothèses implicites : (i) les erreurs de mesures liées à l'appariement et à l'agrégation de l'information sont neutres sur les relations économiques estimées et (ii) les variables RU et HYD résument parfaitement l'information pédologique. La première hypothèse peut être testée en effectuant des régressions en sous-échantillon, sur les observations pour lesquelles ces erreurs sont minimales. Il existe en effet des degrés de pureté des UCS en fonction de la diversité des sols qu'elles contiennent. De même, le MNT contient de l'information pertinente pour ventiler les données sémantiques du RPB. Ces pistes de recherche, qui nécessitent de l'information supplémentaire, n'ont pas pu être traitées dans ce travail. Par contre, nous pouvons fournir des éléments sur la deuxième hypothèse. Nous avons en effet les variables pédologiques qui servent à calculer RU et qui peuvent représenter des attributs naturels pertinents que la variable synthétique ne capte pas. C'est ce que nous étudions dans la section suivante.

5 Prédictions

Les résultats de cette section sont issus du modèle dont les résultats sont présentés en Annexe E, Tableau 6. En plus des variables de la section précédente, les variables qui servent à calculer RU sont incluses parmi les explicatives. La matrice $\mathbf{S}$ des attributs naturels contient les classes de texture de la Figure 5 (CLT), le taux de matières organiques (TMO), l'épaisseur ($EPAI$) et le taux d'éléments grossiers (TEG). Les classes CLT sont mises en interaction simultanée avec RU et HYD, ce qui produit 12 variables explicatives supplémentaires.[9] Certains degrés supérieurs d'interaction apparaissent significatifs, en particulier pour la texture croisée à l'hydromorphie et les interactions entre les trois variables. Les informations topographiques restent polynomiales avec, toujours, une significativité à l'ordre cinq pour $ALTI$ et à l'ordre deux pour $PENTE$. Les coefficients associés aux variables de contrôle changent peu, de même que le R^2 ajusté qui ne progresse que d'environ un point.

TMO présente un effet négatif sur le prix, ce qui remet en cause sa capacité à définir la qualité agricole d'un sol. Cet effet peut s'expliquer par la faible valeur économique des sols forestiers qui présentent en général des TMO importants ou par la faible dépendance de l'agriculture à la richesse organique des sols. Cet effet s'est révélé robuste à l'ensemble des spécifications effectuées dans ce travail. Les deux autres nouvelles variables présentent

9. HYD est considérée comme continue car l'utilisation des classes NH, PH et TH en triple interaction aurait produit 26 variables supplémentaires.

des relations plus intuitives avec les prix observés, $EPAI$ possède un effet positif et TEG un effet négatif. Malgré la multicolinéarité de ce modèle sur laquelle nous reviendrons, les ordres de grandeurs de ces coefficients ne rivalisent toujours pas avec ceux de la topographie. De plus, les effets de premier ordre des variables RU et HYD ne sont plus significatifs.

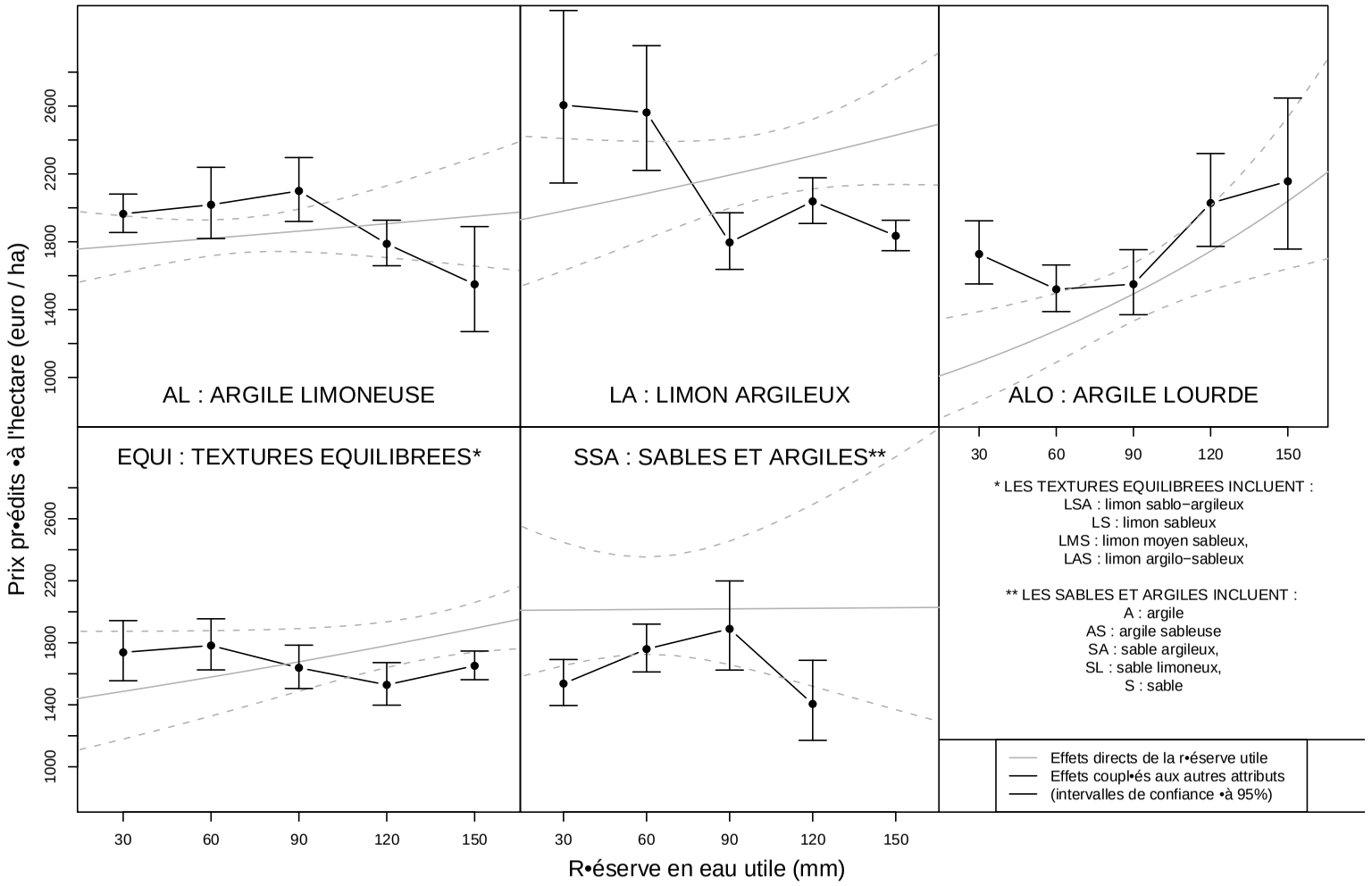
À première vue, nous obtenons que les nouvelles variables captent les effets préalablement attribués à RU et HYD. Cela amène deux interprétations non mutuellement exclusives et que nous ne pouvons pas discriminer sur la base de nos résultats. Soit (i) la texture, l'épaisseur d'un sol ou la présence d'éléments grossiers sont sources d'effets indépendants de la disponibilité en eau (par exemple, le labour peut nécessiter des tracteurs plus puissants et donc plus coûteux), soit (ii) la formule et/ou la pondération pour le calcul de RU ne correspondent pas au mieux au prix de marché de la terre et donc à la valorisation agricole. La multicolinéarité de ce modèle – et des nombreux autres qui ont été estimés sur l'ensemble des variables – est diagnostiquée uniquement pour les attributs pédologiques mais pour l'ensemble de ces variables : autant RU et HYD que TMO, $EPAI$ et TEG. Pour interpréter les effets agrégés des 32 variables naturelles de ce modèle détaillé, nous allons désormais considérer des *paquets* d'attributs.

Le Tableau 7 de l'Annexe E découpe les 4 254 transactions de la base en 25 catégories par un croisement de RU avec CLT. Pour cela, la variable RU est découpée en cinq catégories de valeurs (I–V) décrites dans la légende du tableau. Pour chacune des catégories, le nombre de transactions présentes ($NOBS$) et les médianes des variables naturelles (pédologiques et topographiques) sont reportés. L'inégale ventilation des transactions entre CLT et les classes de RU renforce le constat qui avait été fait sur la largeur des intervalles de confiance des Figures 1 et 2. La grande majorité des transactions aux textures AL et SSA présente des valeurs de RU inférieures à 80 mm. Il apparaît en particulier qu'aucune transaction SSA ne possède une valeur de RU supérieure à 130 mm. Connaissant la faible capacité du sable à stocker l'eau, ce résultat n'est pas étonnant mais fournit une illustration des combinaisons déterministes (biophysiques) entre les attributs naturels. L'inégale distribution des observations est complétée de variations jointes entre les attributs naturels, y compris topographiques. Parmi les éléments les plus marquants, les différentes classes de textures présentent des altitudes et des pentes médianes assez variables, avec un rôle important de RU bien que difficile à interpréter de manière systématique.

Ces 25 catégories sont alors utilisées pour prédire des prix et des intervalles de confiance aux parcelles médianes. Pour cela, nous simulons une matrice $\mathbf{X}^*$ de colonnes constantes qui contiennent les médianes observées sur l'échantillon (sauf la colonne $SURFHA$ qui est posée à l'unité). La matrice $\mathbf{S}^*$ est simulée de manière à présenter une colonne RU^* qui prend les valeurs 30, 60, 90, 120 et 150 pour chacune des classes de texture de la colonne CLT^*. Ces cinq valeurs correspondent aux milieux des classes I–V du Tableau 7 de manière à ce que les autres colonnes de $\mathbf{S}^*$ (y compris les polynômes et les interactions) puissent être remplies avec les valeurs médianes correspondantes. Ces deux matrices simulées nous permettent ensuite d'établir des prédictions de prix et d'intervalles de confiance en utilisant (3). Nous effectuons l'exercice deux fois, en considérant d'abord les variables $ALTI$ et $PENTE$ comme constantes (égales à leur médianes) puis en les laissant prendre les valeurs du Tableau 7. Les représentations graphiques sont disponibles dans les Figures 3 et 4.

Les droites en arrière plan de la Figure 3 représentent les effets propres de RU en in-

Figure 3 – Comparaisons entre effets propres et effets couplés à la pédologie : les droites en gris clair représentent les résultats bruts de l'interaction entre RU et CLT, sachant que la variable HYD (également en interaction dans le modèle) est fixée à sa valeur médiane, soit 1,2. Les points avec des intervalles de confiance associés (à 5%) sont des prédictions de prix avec les autres variables pédologiques (HYD, TMO, TEG et $EPAI$) comme dans le Tableau 7. Dans ces simulations, les attributs topographiques sont constants aux niveaux de leurs médianes, tout comme les variables de contrôle.



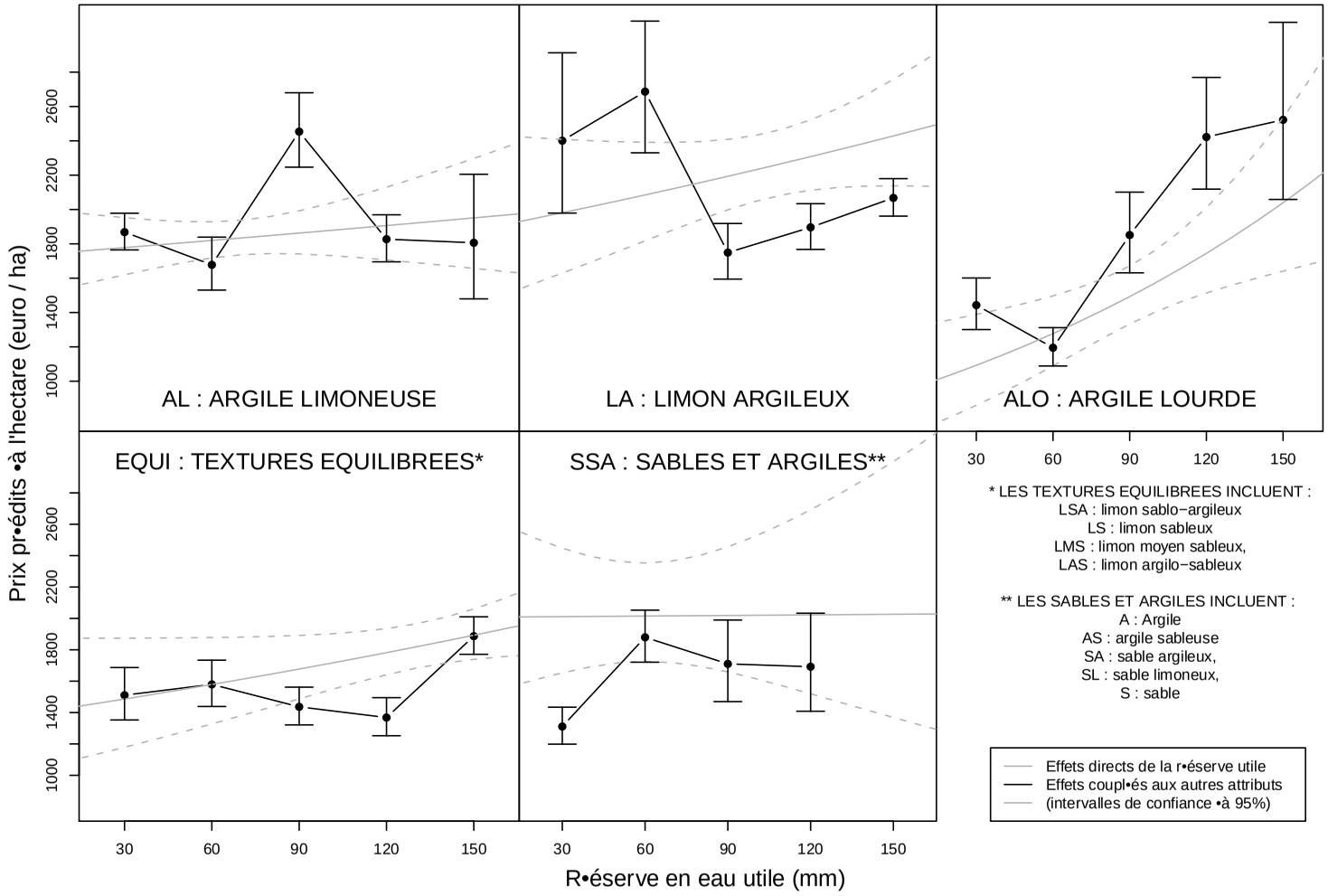
teraction avec CLT, pour des variables HYD fixées à la médiane (soit 1,2). La méthode de prédiction de ces droites est identique à celle des Figures 1 et 2 pour être interprétée comme des effets marginaux de RU. RU présente un prix positif et significatif pour les textures ALO et, à des degrés moindres, les textures LA et $EQUI$. En revanche, le prix de RU n'est pas significativement différent de zéro pour les LA et les SSA. Les 24 points reportés et leurs intervalles de confiance à 5% sont les prix prédits avec les autres variables pédologiques (y compris HYD) à leurs valeurs médianes à l'intérieur des catégories. Chaque point correspond donc à un profil médian de terre, caractérisé par un panier d'attributs. Ces variations supplémentaires sont source d'écarts avec les effets propres de RU mais rarement de manière significative. Seuls les limons argileux (LA) à fortes RU et les argiles lourdes à faibles RU s'éloignent significativement. Ces différences (négative pour LA et positive pour ALO) peuvent être expliquées par les valeurs médianes d'hydromorphies (respectivement fortes et faibles) correspondantes dans le Tableau 7.

En examinant les niveaux, les prix prédits se distribuent entre 1 400 euros/ha pour les sols SSA à forte RU et 2 600 euros/ha pour les sols LA à faible RU. Ces derniers étant peu nombreux en terme de transactions (13+44), il nous semble préférable de déduire que les variations de prix s'établissent sur l'intervalle 1 400–2 000 soit un écart (*range*) de 600 euros, représentant environ 35% du prix médian à l'hectare de l'échantillon (1 743 euros). Présentés ainsi, les effets de RU sont peu monotones et peu significatifs avec toutefois une relation globalement négative pour les sols LA et une relation globalement positive pour les sols ALO. Par contre, lorsque l'on compare les points deux à deux entre différentes classes de textures, des variations significatives apparaissent. Les argiles limoneuses (AL) avec RU inférieure à 90 mm ont un prix prédit significativement supérieur à 1 800 euros. En parallèle, les ALO avec RU entre 60 et 90, les $EQUI$ avec RU supérieure à 90 et les SSA avec RU à 30 et à 120 présentent un prix significativement inférieur à 1 800 euros.

Dans la Figure 4 suivante, les deux variables topographiques principales ($ALTI$ et $PENTE$) ne sont plus constantes entre les 24 points mais se modifient selon les médianes des différentes catégories. Les conséquences en terme d'amplitude des variations de prédiction sont importantes alors que les droites en arrières plan ne sont pas modifiées par rapport à la figure précédente. Les écarts significatifs à ces effets propres de RU sont plus nombreux. Lorsque les variations issues de la topographie sont prises en compte, de nombreux points sortent des intervalles de confiance. Toujours en négligeant les points qui ne contiennent que peu de transactions, les prix prédits se distribuent dans l'intervalle 1 400–2 500 euros, soit un écart équivalent à 60% du prix médian de l'ensemble de l'échantillon. Parmi les profils de terre qui se distinguent, les argiles limoneuses (AL) avec RU aux environs de 90 mm (classe III) présentent un prix prédit à l'hectare de l'ordre de 2 500 euros. Ce profil concerne 143 transactions qui sont en majorité non hydromorphes et présentent une altitude et une pente médiane relativement faibles (respectivement 202 m et 0,6°, Tableau 7). Autant d'éléments qui font que leur prix prédit soit élevé.

À l'inverse, les limons argileux (LA) avec RU aux alentours de 90–120 mm (158 transactions) présentent des prix prédits faibles. Cela est expliqué à la fois par des altitudes médianes relativement fortes (317–385 m) alors que les pentes médianes ne sont pas spécialement supérieures aux autres terres de la même classe de texture. Les sols $EQUI$ de la classe IV de RU sont caractérisés à la fois par des altitudes élevées et des pentes fortes, sans que les différences relatives soient très marquées. Elles se caractérisent toutefois par des différences

Figure 4 – Comparaisons entre effets propres et effets couplés à l'ensemble des variables naturelles : les droites en gris clair représentent les résultats bruts de l'interaction entre RU et CLT, sachant que la variable HYD (également en interaction dans le modèle) est fixée à sa valeur médiane, soit 1,2. Les points avec des intervalles de confiance associés (à 5%) sont des prédictions de prix avec les autres variables naturelles : pédologiques (HYD, TMO, TEG et $EPAI$) et topographiques ($ALTI$ et $PENTE$) comme dans le Tableau 7. Dans ces graphiques, les variables de contrôle sont constantes aux niveaux de leurs médianes.



de prix significatives, en particulier comparativement au terres de la même classe de texture mais de la classe de *RU* supérieure (V). Les sols *SSA* à faible *RU* (classe I) se distinguent par de faibles prix prédits qui peuvent être attribués à la forte altitude médiane (416 m) et à la forte pente médiane (7,5°) que ces terres présentent.

Pour finir sur les prédictions, deux cartes départementales sur les variations de prix issues des attributs naturels sont disponibles dans la Figure 7 de l'Annexe D. La légende de la figure décrit la méthodologie de calcul des écart (en %) aux prix médians départementaux. La Carte 1 présente la variabilité issue uniquement des attributs pédologiques du modèle complet estimé pour cette section. Il apparaît que les sols les plus valorisés sont dans les communes à l'Est du département, avec des prix prédits de l'ordre de 20% de la médiane départementale. À l'inverse, les sols au Sud-Ouest du département présentent des contraintes issues de leurs attributs pédologiques. Leurs prix sont globalement inférieurs de 15% à la médiane. Lorsque les attributs topographiques sont ajoutés pour établir les variations de prix, les écarts entre les communes augmentent, ce qui signifie un effet de renforcement des inégalités de dotation. Les principaux effets géographiques persistent même si les terres du Nord-Ouest, relativement basses et peu pentues, apparaissent désormais comme relativement plus coûteuses. L'autocorrélation spatiale des prix prédits semble également s'accentuer.

6 Conclusion

La valeur d'une unité de terre, même limitée à sa dimension privative, ne peut pas se réduire à la composition intrinsèque du matériau. Il est en effet établi que de nombreux éléments extrinsèques (l'accessibilité, le climat ou l'option de pouvoir la dédier à un usage alternatif) font partie intégrante de cette valeur. Les attributs naturels intrinsèques, non reproductibles (à court et moyen termes) et variables dans l'espace, sont toutefois sources d'avantages et de contraintes qui influencent la production et la répartition de la richesse. En tant que valeurs différenciées, toute évaluation de leur contribution doit être effectuée relativement à une situation contrefactuelle de dotations alternatives. Autrement dit, l'évaluation économique de ce capital naturel ne s'établit par comparaison de l'absence à la présence de terre – ce qui ferait également varier les éléments extrinsèques – mais par la comparaison de quantités ou de paniers variables d'attributs.

L'analyse des relations hédoniques entre le prix des biens fonciers et leurs attributs naturels permet de faire apparaître ces derniers comme bénéficiant d'une valeur significative, affirmant ainsi leur statut de capital naturel. L'agrégation de ces valeurs pour l'ensemble des terres agricoles de Côte-d'Or[10] conduit à une évaluation économique des attributs pédologiques de 270 millions d'euros. Par la même méthode, la valeur agrégée de l'ensemble des attributs naturels est évaluée à 512,5 millions d'euros. Étant donné le biais de sélection potentiel présenté dans la section 2, les limites sur la qualité des données et des estimations, ainsi que les écarts potentiels à l'hypothèse de marchés fonciers concurrentiels, ces résultats ne sont pas définitifs et doivent être utilisés avec précaution. À titre indicatif, ces

10. L'agrégation consiste à sommer, pour l'ensemble des terres du département, les prix prédits par le modèle complet avec des variables des contrôle évaluées à leur niveau le plus défavorable : le maximum lorsqu'elles ont un effet positif et le minimum pour un effet négatif. Pour tenir compte de l'inflation, la *dummy* annuelle de 2005 est posée à "1".

chiffres représentent respectivement 1,86% et 3,5% du PIB départemental en 2005, tel qu'il est calculé par l'INSEE (14 534 millions d'euros[11]).

Cette évaluation de la valeur marchande (pour la production agricole) des attributs naturels n'est qu'une étape dans un processus de questionnement des coûts et des bénéfices sociaux issus de l'usage de la terre. Par rapport aux sociétés agraires, le poids actuel des attributs naturels est souvent présenté comme plus faible (Huston, 2005) mais nous avons montré qu'il compte encore. Les résultats obtenus sur le poids de la topographie relativement à la pédologie remettent en cause la vision d'une hétérogénéité de la terre basée sur les rendements agricoles au profit d'une hétérogénéité issue des complémentarités avec le capital reproductible (matériel, travail). Une prochaine étape consisterait à croiser ce gradient agricole d'hétérogénéité avec les autres fonctions sociale remplies par la terre (en usage urbain ou en usage naturel).

La qualité de l'appariement entre les sources de données apparaît comme un point faible du travail, qui peut potentiellement remettre en cause certains résultats. En particulier, nous obtenons que les variables les plus significatives sont celles qui sont mesurées avec le plus de précision et qui contiennent donc le moins d'erreurs. La précision des résultats reportés dans ce chapitre ne doit pas voiler ces incertitudes, qui vont pouvoir être réduites en améliorant l'information de base et la méthode d'harmonisation dans l'espace. Il est toutefois intéressant de constater les différences entre disciplines et les impacts en terme de production et d'utilisation de données. Ce travail repose sur des informations issues de disciplines différentes (économie et pédologie en particulier) avec des logiques de récolte et de mise en forme différentes. Le meilleur exemple de cet écart est la précision avec laquelle la diversité des sols est représentée dans les données du RPB (au niveau de la strate) alors que leur localisation à l'échelle départementale s'avère peu précise, avec "seulement" 101 polygones (UCS). Le coût d'échantillonnage, l'absence d'un registre de l'ensemble des observation pédologique et l'objectif d'établir une carte des sols pouvant être consultable visuellement semblent les explications principales pour cette méthodologie qui se révèle *a posteriori* un frein à une utilisation alternative. La même remarque est valable pour les données $SCAFR$ qui n'ont pas pour vocation première de servir à l'estimation de modèles statistiques et dont l'absence de représentativité, la présence de valeurs non renseignées et leur coût d'achat constituent des freins à leur utilisation.

11. `http://www.insee.fr/fr/themes/tableau.asp?reg_id=99&ref_id=t_2601D`.

Références

Andrews, S. S. et C. R. Carrol (2001). Designing a soil quality assessment tool for sustainable agroecosystem management. *Ecological Applications* 11 : 1573–1585.

Barbier, E., D. Pearce et A. Markandya (1990). Environmental sustainability and cost-benefit analysis. *Environment and Planning A* 22 : 1259–1266.

Bartelmusa, P. (2009). The cost of natural capital consumption : Accounting for a sustainable world economy. *Ecological Economics* 68 : 1850–1857.

Belsley, D. (1991). *Conditioning diagnostics, collinearity and weak data in regression*. New York : John Wiley & Sons.

Bruand, A., O. Duval et I. Cousin (2004). Estimation des propriétés de rétention en eau des sols à partir de la base de données SOLHYDRO : Une première proposition combinant le type d'horizon, sa texture et sa densité apparente. *Étude et Gestion des Sols* 11 : 323–332.

Cameron, S. E., K. J. Williams et D. K. Mitchell (2008). Efficiency and concordance of alternative methods for minimizing opportunity costs in conservation planning. *Conservation Biology* 22 : 886–896.

Cassman, K. G. (1999). Ecological intensification of cereal production systems : Yield potential, soil quality, and precision agriculture. *Proceedings of the National Academy of Science* 96 : 5952–5959.

Caswell, M. F. et D. Zilberman (1986). The effects of well depth and land quality on the choice of irrigation technology. *American Journal of Agricultural Economics* 68 : 798–811.

Cavailhès, J., M. Hilal, P. Wavresky, G. Contesti et H. el Yousfi (2009). Marché foncier et périurbanisation. Tech. rep., MUHC-INRA.

Cavailhès, J. et P. Wavresky (2003). Urban influences on periurban farmland prices. *European Review of Agricultural Economics* 30 : 333–357.

Chicoine, D. (1981). Farmland values at the urban fringe : An analysis of sales prices. *Land Economics* 57 : 353–362.

Chrétien, J. (2000). Référentiel Pédologique de Bourgogne au 1/250 000 : Régions naturelles, pédopaysages et sols de la Côte d'Or. Tech. rep., Institut National de la Recherche Agronomique.

Costanza, R. et H. E. Daly (1992). Natural capital and sustainable development. *Conservation Biology* 6 : 37–46.

Costanza, R., R. d'Arge, R. de Groot, S. Farberk, M. Grasso, B. H. K. Limburg, S. Naeem, R. V. O'Neill, J. Paruelo, R. G. Raskin, P. Suttonkk et M. van den Belt (1997). The value of the world's ecosystem services and natural capital. *Nature* 387 : 253–260.

De Groot, R., J. van der Perk, A. Chiesura et A. van Vliet (2003). Importance and threat as determining factors for criticality of natural capital. *Ecological Economics* 44 : 187–204.

Dietz, S. et E. Neumayer (2007). Weak and strong sustainability in the seea : Concepts and measurement. *Ecological Economics* 61 : 617–626.

Dormont, B. (2007). *Introduction à l'économétrie*. Montchrestien, 2e édition.

Ekeland, I., J. J. Heckman et L. Nesheim (2004). Identification and estimation of hedonic models. *Journal of Political Economy* 112 : 60–109.

Ervin, D. E. et J. W. Mill (1985). Agricultural land markets and soil erosion : Policy relevance

and conceptual issues. *American Journal of Agricultural Economics* 67 : 938–942.

Fedoroff, N. V., D. S. Battisti, R. N. Beachy, P. J. M. Cooper, D. A. Fischhoff, C. N. Hodges, V. C. Knauf, D. Lobell, B. J. Mazur, D. Molden, M. P. Reynolds, P. C. Ronald, M. W. Rosegrant, P. A. Sanchez, A. Vonshak et J.-K. Zhu (2010). Radically rethinking agriculture for the 21st century. *Science* 327 : 833.

Fenech, A., J. Foster, K. Hamilton et R. Hansell (2003). Natural capital in ecology and economics : An overview. *Environmental Monitoring and Assessment* 86 : 3–17.

Fox, J. (2003). Effect display in R for generalized linear models. *Journal of Statistical Software* 8.

Fox, J. et G. Monette (1992). Generalized collinearity diagnostics. *Journal of the American Statistical Association* 87 : 176–183.

Geniaux, G., J.-S. Ay et C. Napoléone (in press). A spatial hedonic approach on land use change anticipations. *Journal of Regional Science* .

Halvorsen, R. et R. Palmquist (1980). The interpretation of dummy variables in semilogarithmic equations. *American Economic Review* 70 : 474–475.

Hartwick, J. (1977). Intergenerational equity and the investing of rents of exhaustible resources. *American Economic Review* 67 : 972–974.

Heckman, J. J. (1979). Sample selection as a specification error. *Econometrica* 47 : 153–161.

Hendrickx, J. et B. Pelzer (2004). Collinearity involving ordered and unordered categorial variables. In *Paper presented et the RC33 conference in Amsterdam, August 17-20*.

Hilal, M. (2010). Odomatrix. Calcul de distances routiéres intercommunales. *Cahier des Techniques de l'INRA* Numéro spécial : Méthodes et outils de traitement des données en sciences sociales. Retours d'expériences : 41–63.

Huston, M. A. (2005). The three phases of land-use change : Implications for biodiversity. *Ecological Applications* 15 : 1864–1878.

Irwin, E. G. et N. E. Bockstael (2007). The evolution of urban sprawl : Evidence of spatial heterogeneity and increasing land fragmentation. *Proceedings of the National Academy of Science* 104 : 20672–20677.

Jamagne, M., R. Betremieux, J. Begon et A. Mori (1977). Quelques données sur la variabilité dans le milieu naturel de la réserve en eau des sols. *Bulletin Technique d'Information* 324-325 : 627–641.

Joly, D., T. Brossard, H. Cardot, J. Cavailhès, M. Hilal et P. Wavresky (2009). Interpolation par régressions locales : Application aux précipitations en France. *L'Espace Géographique* 38 : 157–170.

Kennedy, P. E. (1981). Estimation with correctly interpreted dummy variables in semilogarithmic equations. *American Economic Review* 71 : 801.

Kennedy Jr, W. et J. Gentle (1980). *Statistical computing*. New York : Marcel Dekker, Inc.

King, G., M. Tomz et J. Wittenberg (2000). Making the most of statistical analyses : Improving interpretation and presentation. *American Journal of Political Science* 44 : 347–36.

Lichtenberg, E. (1989). Land quality, irrigation development, and cropping patterns in the nothern high plains. *American Journal of Agricultural Economics* 71 : 187–194.

Livanis, G., C. B. Moss, V. E. Breneman et R. Nehring (2006). Urban sprawl and farmland prices. *American Journal of Agricultural Economics* 88 : 915–929.

Maddison, D. (2000). A hedonic analysis of agricultural land prices in England and Wales. *European Review of Agricultural Economics* 27 : 519–532.

Mendelsohn, R., W. D. Nordhaus et D. Shaw (1994). The impact of global warming on agriculture : A Ricardian analysis. *American Economic Review* 84 : 753–771.

Miranowski, J. A. et B. D. Hammes (1984). Implicit prices of soil characteristics for farmland in Iowa. *American Journal of Agricultural Economics* 66 : 745–749, proceeding Issue.

Newburn, D. A., P. Berck et A. M. Merenlender (2006). Habitat and open space at risk of land use conversion : Targeting strategies for land conservation. *American Journal of Agricultural Economics* 88 : 28–42.

Palmquist, R. B. (1989). Land as a differentited factor of production : A hedonic model and its implications for welfare measurement. *Land Economics* 65 : 23–28.

Palmquist, R. B. et L. E. Danielson (1989). A hedonic study of the effects of erosion control and drainage on farmland values. *American Journal of Agricultural Economics* 71 : 55–62.

Plantinga, A. J., R. N. Lubowski et R. N. Stavins (2002). The effects of potential land development on agricultural land prices. *Journal of Urban Economics* 52 : 561–581.

Ricardo, D. (1817). *Principles of political economy and taxation*. Great minds series, London.

Solow, R. (1974). Intergenerational equity and exhaustible resources. *Review of Economic Studies* 41 : 29–46.

Stiglitz, J., A. Sen, J.-P. Fitoussi, B. Agarwal, K. Arrow, A. Atkinson, F. Bourguignon, J.-P. Cotis, A. Deaton, K. Dervis, M. Fleurbaey, N. Folbre, J. Gadrey, E. Giovannini, R. Guesnerie, J. Heckman, G. Heal, C. Henry, D. Kahneman, A. Kruger, A. Oswald, R. Putnam, N. Stern, C. Sunstein, P. Weil, J.-E. Chapron, D. Blanchet, J. Le Cacheux, M. D'Ercole, P. Pionnier, L. Rioux, P. Schreyer, X. Timbeau et V. Marcus (2009). Rapport de la Commission sur la mesure des performances économiques et du progrès social. Tech. rep., Paris.

van Garderen, K. J. (2001). Optimal prediction in log-linear models. *Journal of Econometrics* 104 : 119–140.

van Garderen, K. J. et C. Shah (2002). Exact interpretation of dummy variables in semilogarithmic equations. *Econometrics Journal* 5 : 149–159.

Wiebe, K. (2003). *Land quality, Agricultural productivity, and food security : Biophysical processes and economic choices at local, regional, and global level*. Edward Elgar.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. The MIT press, Cambridge.

Wu, J. et K. Segerson (1995). The impact of policies and site characteristics on potential grounwater pollution in Wisconsin. *American Journal of Agricultural Economics* 77 : 1033–1047.

Wu, J., D. Zilberman et B. A. Babcock (2001). Environmental and distributional impacts of conservation targeting strategies. *Journal of Environmental Economics and Mangement* 41 : 333–350.

Zelek, C. A. et G. E. Shively (2003). Measuring the opportunity cost of carbon sequestration in tropical agriculture. *Land Economics* 79 : 342–354.

Annexes

A Interprétations du modèle semi-logarithmique.

Bien qu'estimé par MCO, le modèle semi-logarithmique n'est pas un modèle linéaire en terme d'effets en volumes, ce qui peut être gênant pour obtenir des prix marginaux. Halvorsen et Palmquist (1980) ainsi que Kennedy (1981) discutent de la mesure appropriée pour interpréter l'effet d'une variable *dummy* $s_k = 1$ par rapport à la modalité alternative $s_k = 0$. Nous étendons ici les résultats à toute variation discrète, y compris pour une variable continue. Dans le cas *dummy*, la formule théorique du vrai paramètre g d'intérêt est :

$$\mathbb{E}\big(P \,\big|\, s_k = 1\big) = (1 + g) \times \mathbb{E}\big(P \,\big|\, s_k = 0\big). \tag{4}$$

Cette équation montre tout d'abord que si des estimations pour les espérances conditionnelles (c'est-à-dire des prédictions) sont disponibles, il suffit de faire le rapport des deux et de retrancher 1 pour obtenir le paramètre d'intérêt. En économétrie, on préfère en général interpréter des coefficients MCO d'où la recherche d'une équivalence entre g et $\widehat{\delta}_k$. En prenant le logarithme de l'équation (sous l'hypothèse $g > -1$) nous remarquons que la différence des *logarithmes* des espérances est le coefficient théorique δ_k, identifié par le modèle semi-logarithmique. La relation entre le paramètre d'intérêt et le paramètre identifié est donc :

$$g = \exp\big(\delta_k\big) - 1. \tag{5}$$

Par l'hypothèse sur les erreurs, δ_k est une variable aléatoire normalement distribuée de moyenne $\widehat{\delta}_k$ donnée par les MCO. Sa variance estimée est, par abus de notations sur l'appariement des matrices d'explicatives : $\widehat{\sigma}_k^2 = \widehat{\sigma}_\varepsilon^2 \cdot \big[(\mathbf{XS}'\mathbf{XS})^{-1}\big]_{kk}$. Le paramètre d'intérêt g est également une variable aléatoire avec pour espérance $\widehat{g} := \mathbb{E}(g) = \mathbb{E}\big(\exp(\delta_k)\big) - 1$. Nous avons donc que g suit une loi log-normale (non symétrique). L'espérance d'une variable log-normale est telle que $\exp(\widehat{\delta}_k) = \exp\big(\delta_k + \widehat{\sigma}_k^2/2\big)$, ce qui nous donne ici :

$$\widehat{g} = \exp\big(\widehat{\delta}_k - \widehat{\sigma}_k^2/2\big) - 1. \tag{6}$$

L'équation (2), qui sera utilisée pour interpréter les résultats marginaux dans cet article, est une généralisation de l'équation précédente qui peut admettre des saut discrets de variables continues : typiquement le passage de s_k^0 quelconque à s_k^1 avec $s_k^1 := s_k^0 + \widehat{\sigma}_{s_k}$.

B Calcul de la Réserve en eau Utile

À partir des caractéristiques de strates de sol contenues dans le RPB, RU a été calculée à l'aide d'une fonction de pédotransfert présentée dans Bruand et al. (2004). La valeur de RU pour chaque strate (indicée c) est déduite à partir de l'ensemble des variables pédologiques observées. Les valeurs par défaut de la fonction de pédotransfert ont été utilisées, ce qui est sans effet sur les résultats. La formule finale est la suivante :

$$RU_c(mm) = RU\big(CLT_c, TMO_c\big) \times EPAI_c(cm) \times \big(1 - TEG_c(\%)\big). \tag{7}$$

La fonction $RU(\cdot)$ est calibrée par l'expérimentation et les valeurs sont disponibles dans Bruand et al. (2004). Les RU par type de sol sont les sommes des strates et les valeurs par UCS sont leurs moyennes pondérées par les surfaces. Nous devons également mentionner des hypothèses qui ont été admises pour ce calcul : (i) la profondeur maximale de sol prise en compte est 1,20 m et (ii) les remontés capillaires ont été prises en compte uniquement au sein des types de sol où elles sont susceptibles d'apporter une contribution significative. Lorsque qu'un substrat du type alluvions sableuses ou grèzes calcaires est présent à moins de 1,20 m de profondeur, 10–15 cm de ce matériau ont été inclus dans le calcul de RU et 20-25 cm pour la craie pour laquelle les remontées sont très importantes.

C Description des données

Table 4 – Sources et intitulés des variables utilisées

NOM	SOURCE	INTITULÉ
PRIX	*SCAFR*	Prix en euros courants
SURF	*SCAFR*	Surface en hectares
PXHA	*SCAFR*	Prix en euro par hectare
AN	*SCAFR*	Année de la vente
NC	*SCAFR*	Nature cadastrale
FERM	*SCAFR*	"OUI" si l'acheteur est le dernier fermier "NON" si l'acheteur n'est pas le dernier fermier "LIBRE" si pas de fermage avant la vente
DIJON	ODO	Distance, en km, à la mairie de Dijon
DUMDIJ	ODO	"1" pour plus de 35 km de Dijon et "0" sinon
DISTBV	ODO	Distance, en km, au chef-lieu du pôle de bassin le plus proche
DISTMA	ODO	Distance, en km, au chef-lieu de la commune d'appartenance
DRAIN	SSP	Part *communale* de surface agricole utile drainée
IRRIG	SSP	Part *communale* de surface agricole utile irriguée
RU	RPB	Réserve en eau Utile en millimètres
HYD	RPB	Indice d'hydromorphie dominant
NH	RPB	"1" si *HYD* est 0 sur toutes les parcelles, "0" sinon
PH	RPB	"1" si *HYD* est strictement entre 0 et 3, "0" sinon
TH	RPB	"1" si *HYD* est 3 sur toutes les parcelles, , "0" sinon
TARG	RPB	Taux d'argile en %
TLIM	RPB	Taux de limon en %
TSAB	RPB	Taux de sable en %
TEG	RPB	Taux d'éléments grossiers en %
TMO	RPB	Taux de matière organique en ‰
EPAI	RPB	Epaisseur du sol
CLT	RPB	Classe de texture calculée (Figure 5)
ALTI	MNT	Altitude en mètres
PENTE	MNT	Pente en degrés
EXPO	MNT	Exposition : *SUD*, *NORD*, *EST* ou *OUEST*

SCAFR : *Société Centrale d'Aménagement du Foncier Rural.*
RPB : *Référentiel Pédologique de Bourgogne.* **MNT** : *Modèle Numérique de Terrain.*
ODO : *distancier Odomatrix.* **SSP** : *Service de la Statistique et de la Prospective.*

Table 5 – Statistiques descriptives : *NOBS = 4 254*

CONTINUES	**Moyenne**	**E.T.**	**Min.**	**Max.**
PRIX	6 947	12 430	59,46	211 000
SURF	3,14	4,51	0,02	36,3
PXHA	2 223	1 583	289	9 995
DIJON	31,56	6,72	0	35
DISTBV	11,56	7,71	0	44,64
DISTMA	1,35	0,86	0,03	5,8
DRAIN	0,11	0,25	0	0,98
IRRIG	0,06	0,17	0	1
RU	85,83	46	12	171
HYD	1,38	1,28	0	3
TARG	33,54	13,17	7	67
TLIM	45,57	14,64	12	70
TSAB	20,88	20	1	81
TEG	10,68	15,82	0	80
TMO	21,3	20	1	81
EPAI	53,33	20,57	5	93
ALTI	319,8	112,4	175	658
PENTE	5,93	6,5	0	55,79
DISCRÈTES		**Effectif**		**Pourcentage**
DUMDIJ		2 844		66,85%
FERM : OUI		1 982		46,59%
FERM : NON		652		15,32%
FERM : LIBRE		1 620		38,08%
Nature Cadastrale				
TERRE	2 464	(57,92%)	*PANA*	331 (7,78%)
PRES	847	(19,91%)	*BOIS*	182 (4,28%)
TPRE	430	(10,11%)		
Année de vente				
1993	367	(8,63%)	2000	319 (7,50%)
1994	274	(6,44%)	2001	303 (7,12%)
1995	299	(7,03%)	2002	385 (9,05%)
1996	275	(6,46%)	2003	389 (9,15%)
1997	246	(5,78%)	2004	419 (9,85%)
1998	289	(6,79%)	2005	424 (9,97%)
1999	265	(6,22%)		
Classes de texture				
AL	1 403	(33,0%)	*EQUI*	664 (15,6%)
LA	802	(18,9%)	*SSA*	706 (16,6%)
ALO	679	(16,0%)		
Exposition				
SUD	1 325	(31,15%)	*OUEST*	948 (22,28%)
EST	1 336	(31,41%)	*NORD*	645 (15,16%)

D Figures additionnelles

Figure 5 – Les transactions $SCAFR$ dans le triangle des textures : chaque croix blanche représente une transaction ($NOBS = 4\,254$) en fonction de sa combinaison de textures. Cette figure a été réalisée grâce au package de Julien Moeys (2009) : The soil texture wizard : R functions for plotting, classifying, transforming and exploring soil texture data, `http://julienmoeys.free.fr/?Soil-textures-triangle`.

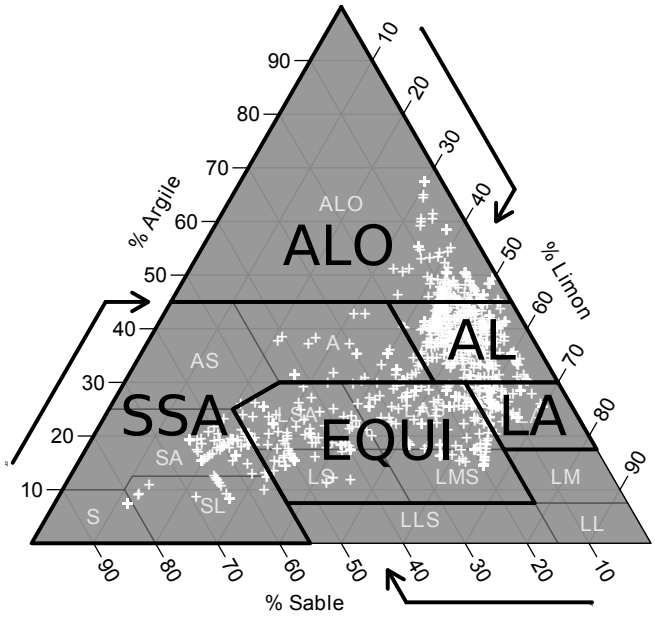


Figure 6 – Les interactions de la réserve utile avec l'altitude et la pente : le codage de l'altitude est le suivant : $M200$ pour les terres à moins de 200 m, $M300$ entre 200 et 300 m, $M400$ entre 300 et 400 m et $P400$ plus de 400 m. Le codage de la pente est : $NULLE$ pour les terres de moins de 2 degrés, $FAIBLE$ entre 2 et 5 degrés, $MOYENNE$ entre 5 et 8 et $FORTE$ pour plus de 8 degrés.

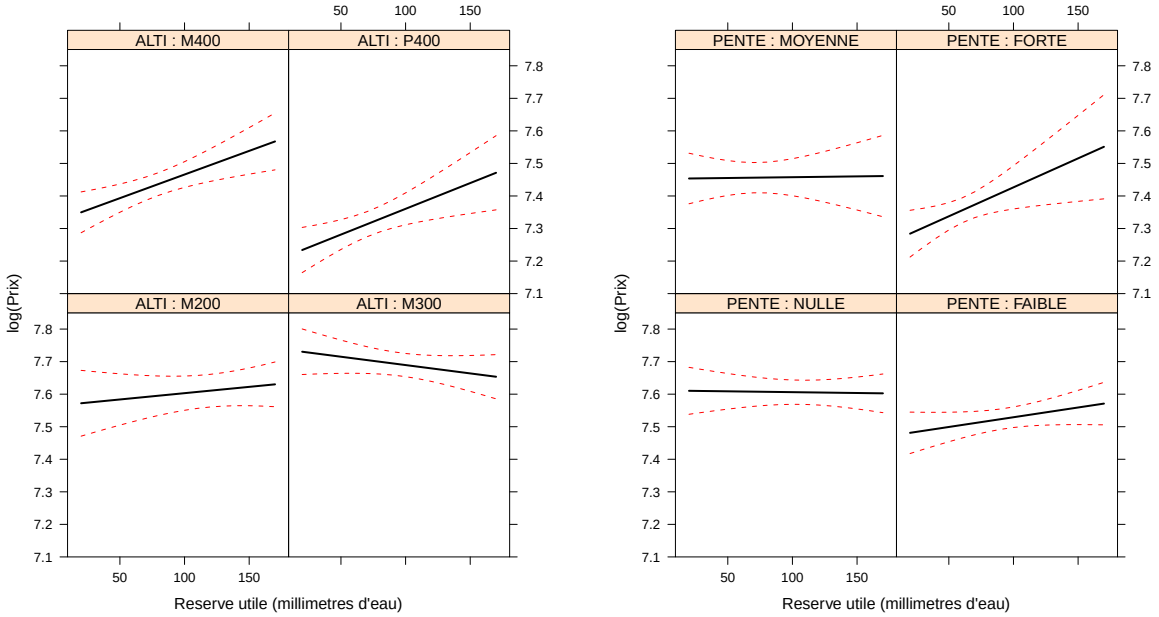
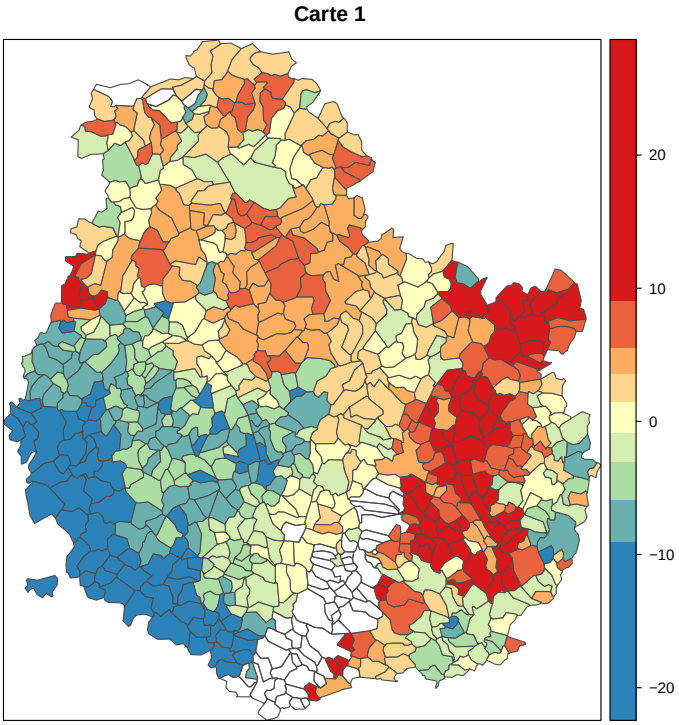
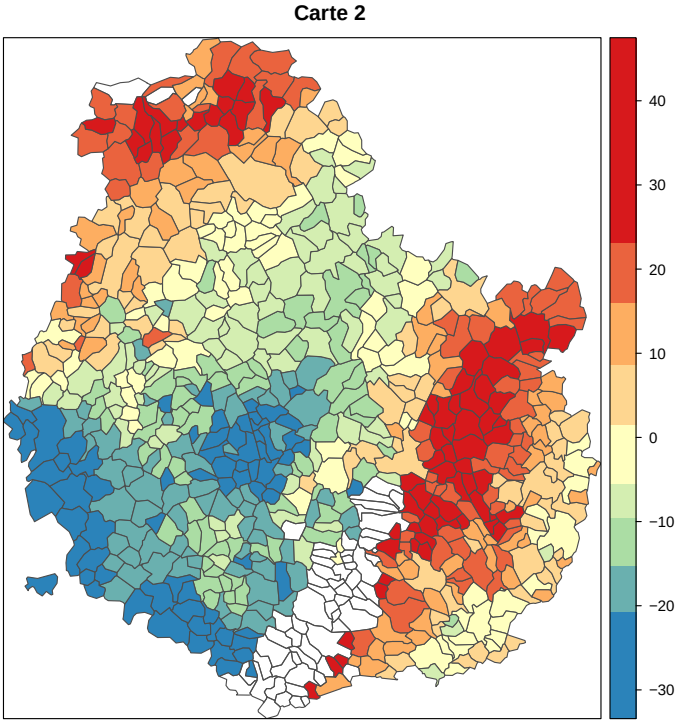


Figure 7 – Les variations de prix issues des attributs pédologiques et de l'ensemble des attributs naturels : ces cartes sont obtenues sur la base de prédictions semblables aux Figures 3 et 4. Nous établissons des prix prédits en ne faisant varier que les attributs pédologiques (Carte 1) et l'ensemble des attributs naturels (Carte 2). Les prix prédits sont des moyennes communales exprimées en pourcentage de la valeur médiane des prédictions, soient 2 200 euros pour la Carte 1 et 2 135 pour la 2. Les communes de la Côte viticole sont en blanc.





E Tableaux de la section 5

Table 6 – Les résultats du modèle complet de la section 5

Variables	Coeff.	E.T.	Variables	Coeff.	E.T.
$CONSTANTE$	+ 8,456**	*0,094*	$log(SURFHA)$	+ 0,999**	*0,007*
$DISTMA$	− 0,040**	*0,001*	DBV	− 0,002	*0,001*
$DIJON$	− 0,023**	*0,002*	$DUMDIJ$	− 0,127**	*0,030*
$FERM : OUI$	− 0,041*	*0,018*	$FERM : NON$	+ 0,051*	*0,024*
$IRRIG$	+ 0,293**	*0,059*	$DRAIN$	− 0,101*	*0,057*
$NC : PRES$	− 0,001	*0,030*	$NC : TPRES$	+ 0,074	*0,029*
$NC : PANAT$	− 0,002	*0,034*	$NC : BOIS$	− 0,411**	*0,042*
$AN : 1994$	− 0,021	*0,041*	$AN : 1995$	− 0,009	*0,040*
$AN : 1996$	− 0,026	*0,041*	$AN : 1997$	− 0,003	*0,042*
$AN : 1998$	+ 0,004	*0,040*	$AN : 1999$	+ 0,037	*0,041*
$AN : 2000$	+ 0,121**	*0,039*	$AN : 2001$	+ 0,070	*0,040*
$AN : 2002$	+ 0,385**	*0,039*	$AN : 2003$	+ 0,379**	*0,038*
$AN : 2004$	+ 0,303**	*0,038*	$AN : 2005$	+ 0,370**	*0,037*
RU	+ 0,001	*0,001*	HYD	+ 0,082	*0,055*
$ALTI$	− 7,096**	*0,941*	$PENTE$	− 6,071**	*0,785*
$ALTI^2$	+ 0,666	*0,714*	$PENTE^2$	+ 1,768**	*0,607*
$ALTI^3$	+ 2,546**	*0,609*	TMO	− 0,006**	*0,001*
$ALTI^4$	− 1,376*	*0,573*	$EPAI$	+ 0,003*	*0,001*
$ALTI^5$	+ 1,902**	*0,546*	TEG	− 0,001*	*0,000*
Exposition (ref = SUD) :					
EST	− 0,002	*0,020*	$NORD$	− 0,052	*0,026*
$OUEST$	− 0,052*	*0,022*			
Textures (ref = AL) :					
LA	+ 0,175	*0,133*	$EQUI$	+ 0,098	*0,117*
ALO	− 0,034	*0,092*	SSA	+ 0,082	*0,123*
Interactions					
$RU \times HYD$	− 0,001	*0,331*			
Interactions CLT (ref = AL)					
$LA \times HYD$	− 0,308**	*0,089*	$EQUI \times HYD$	− 0,162*	*0,080*
$ALO \times HYD$	− 0,801**	*0,203*	$SSA \times HYD$	+ 0,468*	*0,183*
Interactions RU (ref = AL)					
$\times LA$	+ 0,001	*0,001*	$\times LA \times HYD$	+ 0,002*	*0,001*
$\times ALO$	− 0,002	*0,002*	$\times ALO \times HYD$	+ 0,008**	*0,002*
$\times EQUI$	+ 0,000	*0,002*	$\times EQUI \times HYD$	+ 0,001	*0,001*
$\times SSA$	+ 0,000	*0,002*	$\times SSA \times HYD$	− 0,006*	*0,002*
R^2 Aj.	0,8589		$\widehat{\sigma}^2_\varepsilon$	0,2590	
$NOBS$	4 254				

*** : Significativement différent de zéro à 1%* et ** pour 5%.*

Table 7 – Médianes intra-catégories des attributs naturels : le codage de la Réserve en eau Utile est le suivant : I pour *RU* entre 0 et 50, II pour *RU* entre 50 et 80, III pour *RU* entre 80 et 100, IV pour *RU* entre 100 et 130, et V pour *RU* entre 130 et 170.

	RU	**AL**	**LA**	**ALO**	**EQUI**	**SSA**
	I	671,0	13,00	284,0	155,0	375,0
	II	617,0	44,00	25,00	55,00	308,0
NOBS	III	143,0	31,00	75,00	80,00	51,00
	IV	160,0	127,0	346,0	40,00	34,00
	V	18,00	643,0	2,000	381,0	0,000
	I	0,000	0,000	0,000	0,089	0,000
	II	3,000	0,000	0,371	0,000	1,000
HYD	III	0,000	2,000	3,000	2,170	3,000
	IV	1,000	2,000	1,980	2,000	1,000
	V	2,463	3,000	1,461	3,000	–NR–
	I	350,7	360,5	473,9	371,4	416,2
	II	358,7	223,0	486,3	398,6	257,2
ALTI	III	202,8	316,7	183,7	437,1	411,5
	IV	298,9	385,3	185,0	401,5	182,0
	V	192,1	221,2	187,6	205,9	–NR–
	I	4,563	6,127	4,790	9,555	7,550
	II	13,73	6,989	8,260	6,120	3,091
PENTE	III	0,631	5,375	0,452	4,041	3,629
	IV	3,200	3,585	0,367	5,175	0,000
	V	1,588	1,893	1,694	2,036	–NR–
	I	7,126	10,00	10,00	50,00	65,00
	II	15,00	7,000	10,58	41,00	64,29
TMO	III	5,000	10,00	3,000	26,00	25,00
	IV	10,00	13,00	8,000	35,88	42,00
	V	12,47	15,00	11,15	20,00	–NR–
	I	33,00	20,00	51,00	33,02	38,00
	II	59,00	28,00	51,98	40,00	47,00
EPAI	III	35,00	61,00	45,00	50,32	52,00
	IV	74,00	44,00	67,00	65,99	44,00
	V	79,50	80,00	90,00	80,00	–NR–
	I	20,00	50,00	5,000	45,00	40,00
	II	10,00	5,000	5,614	17,00	1,000
TEG	III	0,000	10,00	0,000	0,000	0,000
	IV	0,000	0,000	0,000	0,105	0,000
	V	0,000	0,000	0,000	0,000	–NR–