

PRINCIPES ET APPLICATIONS DES RÉSEAUX DE NEURONES :

Deux illustrations sur l'habitat au Maroc

Denis THUILLIER*

***Résumé** - Les réseaux de neurones, dont les utilisations se multiplient, n'ont pas fait jusqu'à présent l'objet de nombreuses applications en analyse économique. L'article expose les principes qui régissent l'architecture et le fonctionnement des réseaux non supervisés et supervisés, et propose ensuite deux illustrations empiriques sur une même base de logements ruraux marocains. La projection de cette base sur une carte de Kohonen met en évidence les caractéristiques analytiques de la carte auto-organisée qui constitue une alternative conviviale aux techniques classiques et devrait à l'avenir susciter un plus grand intérêt. Il apparaît, par ailleurs, que, sur la base de données utilisée, les capacités de prédiction des réseaux supervisés simples sont significativement supérieures à celles des régressions multiples. La capacité explicative des réseaux supervisés reste toutefois limitée et leur estimation ne va pas sans difficulté.*

Mots-clés - RÉSEAUX DE NEURONES, KOHONEN, HABITAT, TYPOLOGIE, PRIX HÉDONIQUE, MAROC.

L'auteur remercie Cl. Besner (UQAM) pour le traitement informatique des bases et l'estimation empirique des modèles. Il remercie également E. Chéron (UQAM) et un rapporteur de la Revue Région et Développement pour leurs remarques et suggestions mais reste seul responsable des erreurs qui peuvent subsister.

Cette recherche a bénéficié de subventions octroyées par le CRSH (Conseil de Recherches en Sciences Humaines) du Canada.

* Professeur à l'École de Sciences de la gestion de l'Université du Québec à Montréal, Département des Sciences Administratives, 315 rue Sainte-Catherine Est, Montréal, Québec H2X 3X2.

INTRODUCTION

Les outils classiques de l'analyse de données en sciences humaines sont connus. Ils appartiennent à une même famille : analyse de variance, analyse typologique, analyse en composantes principales, etc. L'analyse révèle des causalités, des typologies, des caractéristiques qui facilitent la spécification de modèles explicatifs. L'approche classique –et fructueuse– consiste à formuler explicitement un modèle et à l'estimer. L'explication d'une réalité complexe peut toutefois buter sur les difficultés d'estimation des modèles complexes (non linéaires notamment) ou sur les limites (et les biais) des modèles plus facilement estimables mais exagérément simplifiés par rapport à la réalité.

Une bonne prédiction sur données inter-temporelles résulte généralement d'un bon modèle explicatif mais on sait qu'il est possible de prévoir avec succès (modèle auto-régressif par exemple) sans pour autant expliquer. Une bonne précision sur "cross-sections" peut aussi être obtenue en travaillant l'ajustement statistique du modèle par modification des formes fonctionnelles lors de l'estimation. Le choix de la forme fonctionnelle se fait alors sans référence à un cadre conceptuel a priori cohérent avec le phénomène analysé. Dans ce contexte, une tentative d'interprétation des paramètres n'est pas sans danger. Enfin, quelles que soient les qualités de l'estimateur, il existe toujours un *meilleur* ajustement statistique que celui que l'on a obtenu notamment lorsque, en coupe transversale, la variable dépendante est fonction d'un cortège important de variables explicatives. Le modèle idéal est en effet contenu dans les données de panel ; s'en approcher suppose l'absence de toute spécification de forme fonctionnelle et l'existence d'un mécanisme d'estimation dont le résultat *est* le modèle optimal.

Les réseaux de neurones constituent, tant pour analyser des données que pour prédire une variable dépendante, une nouvelle famille d'instruments qui, bien que conçus pour d'autres usages, offrent quelque intérêt pour l'économiste. Une analyse de données socio-économiques peut en effet être réalisée à l'aide d'un réseau non supervisé. Le résultat se présente sous forme d'une carte "auto-organisée" de la base de données dont la lecture révèle la typologie et les caractéristiques essentielles. Les réseaux supervisés permettent de leur côté de travailler sans spécifier un modèle a priori. Le modèle optimal (non explicite cependant) résulte de l'estimation elle-même et doit permettre, en principe, d'obtenir des capacités de prédiction supérieures à celles que permettent les formes fonctionnelles classiques.

C'est dans ce contexte que le présent travail a été mené. Il a pour but d'illustrer d'une part les capacités d'analyse de données des réseaux non supervisés et de montrer d'autre part qu'un réseau supervisé simple permet d'améliorer sensiblement les capacités prédictives d'un modèle linéaire de régression multiple. On expose en première partie les caractéristiques essentielles et le fonctionnement des réseaux non supervisés et supervisés. La deuxième partie est consacrée aux illustrations empiriques réalisées sur une base de données qui décrit l'habitat rural marocain. Une conclusion permet de dégager les principaux attraits de tels outils pour l'économiste.

1. LES RÉSEAUX DE NEURONES

Le lecteur pourra se référer à la littérature déjà importante consacrée aux réseaux de neurones. Une introduction peut être trouvée dans Davalo et Naïm (1990) et une présentation plus complète dans Jodouin (1992). Un exposé plus large de la théorie sous-jacente et des applications suppose la lecture de quelques auteurs désormais "classiques" (Rummelhart et McClelland, 1986 ; Wasserman, 1989 ; Beale et Jackson, 1990 ; Dayhoff, 1990 ; Hecht-Nielsen, 1990 ; Gallant, 1993 ; Haikin, 1994). On se limite dans ce qui suit à un descriptif rapide de l'outil et de son fonctionnement.

Les réseaux de neurones artificiels sont construits sur une architecture semblable, en première approximation, à celle du cerveau humain. Le réseau reçoit les informations sur une couche réceptrice de "neurones", traite ces informations avec ou sans l'aide d'une ou plusieurs couches "cachées" contenant un ou plusieurs neurones et produit un signal (ou plusieurs) de sortie. Chaque neurone, qu'il appartienne à la première couche (réceptrice), aux couches cachées ou à la couche de sortie, est lié aux autres neurones par des connexions (similaires aux synapses du cerveau) auxquelles sont affectés des poids (eux-mêmes assimilables aux potentiels synaptiques). On conçoit que deux grands types d'apprentissage des réseaux puissent exister selon que l'on spécifie ou non une variable dépendante à valeurs connues. Dans l'apprentissage non supervisé, le réseau combine l'information d'entrée et l'analyse sans référence à une valeur de sortie. De tels réseaux permettent des identifications de caractéristiques spécifiques, des tris, des reconnaissances de formes, de typologies, etc. L'apprentissage supervisé suppose par contre un "output" connu. Chaque information d'entrée est couplée à une valeur de sortie et le réseau cherche à approcher, en minimisant une distance, la fonction implicite contenue dans les informations d'entrée et les valeurs de sortie. Ce type d'apprentissage correspond à des applications où l'on cherche à prédire une variable dépendante, que ce soit sous forme continue ou binaire (prix, taux d'intérêt, probabilité d'occurrence d'un événement, etc.).

On décrit dans ce qui suit les deux types d'apprentissage qui s'effectuent sur deux types de réseaux différents. Il convient de souligner que nous limitons les développements ci-après au minimum permettant au lecteur d'en comprendre les principes et de suivre l'exposé des deux applications proposées en seconde partie de l'article.

1.1. Les réseaux non supervisés : la carte de Kohonen

Le concept de réseau non supervisé est issu des observations des neurobiologistes sur la cartographie cérébrale qui s'auto-structure dans les premières années d'existence de l'être humain. Le réseau artificiel, qui ne comprend pas de couche cachée, est globalement basé sur le même principe : lorsque l'on présente la base de données au réseau lors de l'apprentissage, un seul neurone de la couche de sortie est activé par une ou plusieurs informations d'entrée (pour autant que celles-ci soient suffisamment voisines). Les neurones de sortie se spécialisent pendant l'apprentissage en reconnaissant certaines caractéristiques qui leur deviennent spécifiques et des neurones de sortie voisins reconnaissent des objets d'entrée dont les caractéristiques sont voisines. La couche de sortie une fois stabilisée après de multiples présentations des données d'entrée constitue donc une carte ordonnancée unie ou bidimensionnelle de l'ensemble des informations multidimensionnelles de la base étudiée.

On expose ci-après le principe de l'algorithme qui sous-tend la "Self-Organizing Feature Map" (SOFM) ou carte de Kohonen (1982a, 1982b, 1990). La figure 1 illustre la projection vectorielle effectuée par l'algorithme. La base de données d'entrée est représentée à gauche de la figure par n vecteurs x_k (population d'entreprises, de logements, de ménages, etc. décrits par un vecteur de k caractéristiques). La carte de Kohonen, bidimensionnelle dans cette illustration, est située à droite. Les sites (neurones) de la carte, de dimension 4×5 dans la figure mais qui peut être de dimension quelconque, sont affectés de jeux de poids w_{ij} initiaux également de dimension k , aléatoires (quoique petits) à la première itération.

On définit d_{ij} comme étant la distance entre le site ij (représenté par ses k composantes initiales w_{ij}) et le vecteur x_k . En choisissant une distance euclidienne

on a, pour chacun des objets de la base de données, une distance d_{ij} telle que :

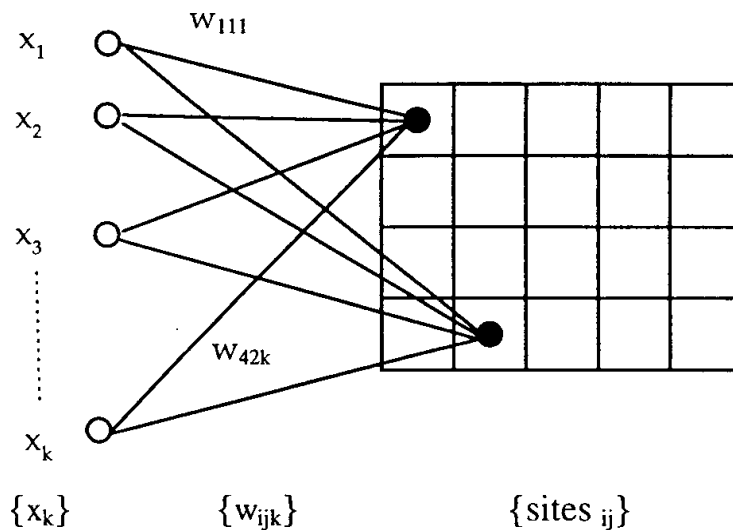
$$d_{ij} = |x_k - w_{ijk}|$$

soit

$$d_{ij} = \sum (x_k - w_{ijk})^{1/2}$$

Si le nombre d'objets x_k présentés à chaque site ij est n , il existe pour chacun des n objets un site ij pour lequel la distance d_{ij} est minimale. Chaque objet de la population se trouve ainsi affecté au site ij en question. On notera que suivant les dimensions de n et de ij , plusieurs objets peuvent se retrouver dans le même site ij de la carte et que, par là-même, certains sites peuvent très bien ne recevoir aucun objet. On comprend le parallèle avec le fonctionnement visuel du cerveau, comme si le site ij voyait (était excité par) un objet ou quelques objets de la population et ne voyait pas les autres, ces derniers étant vus par d'autres sites ; certaines zones de la carte de Kohonen pouvant par ailleurs rester aveugles.

Figure 1 : Illustration de la carte de Kohonen



Cette première projection de la base d'entrée sur un espace à deux dimensions étant effectuée, on définit pour chaque site qui a capté un ou plusieurs

terminée lorsque la grille est stabilisée. De fait, l'équation ci-dessus tend à rapprocher (au plan euclidien) les composantes w_{ij} d'un site de celles d'un vecteur d'entrée x_k et réalise en même temps un ordonnancement topologique, les sites adjacents de la grille tendant à avoir des composantes w_{ij} proches et contenant donc des objets voisins sinon similaires. La sélection des paramètres de pas d'apprentissage et de dimension du voisinage ne se fonde pour le moment que sur l'expérience. Kohonen (1982a, 1982b, 1990) distingue deux grandes phases lors de la réalisation de la grille : une phase d'ordonnancement topologique (positionnement relatif des sites) suivie d'une phase de convergence (regroupement de plus en plus "fin" des données d'entrées dans les sites). Nous reportons le lecteur à la littérature scientifique spécialisée sur les questions de convergence de l'algorithme et de stabilité de la solution et notamment à Haykin (1994) et Gallant (1993) qui recensent les travaux récents sur ces sujets.

La carte de Kohonen (SOFM : Self-Organizing Feature Map) ne se distingue pas beaucoup sur le fond des techniques classiques d'analyse de données telles que l'analyse en composantes principales, l'analyse typologique (K-means)³ ou les techniques de projection vectorielle (compression de données). Elle a cependant le mérite de visualiser la projection vectorielle de façon très lisible en conservant sur deux dimensions l'ordonnancement intrinsèque de la population multidimensionnelle et se prête par-là même à tous les traitements et interprétations auxquels nous sommes habitués sur une image en deux dimensions. La lecture d'une carte reste en effet un exercice simple, instructif et sollicite l'imagination.

1.2. Les réseaux supervisés et l'algorithme de rétro-propagation

Dans un réseau supervisé, chaque neurone de couche cachée ou de couche de sortie somme l'information provenant de l'amont, pondérée par le poids de chaque connexion entre l'amont et lui-même, et produit une information (activation) qui est la transformée par une fonction d'activation le plus souvent sigmoïde sur]0, 1[de la somme pondérée des entrées du neurone.

A chaque vecteur d'entrée (à une composante du vecteur d'entrée correspond un neurone de la couche d'entrée) correspond donc une valeur du vecteur de sortie (unidimensionnel ou multidimensionnel) qui est la réponse donnée par le réseau, pour un jeu de connexions données, au stimulus constitué par le vecteur d'entrée. L'information se propage en conséquence de gauche à droite suivant un modèle

³ "K-means" conduit en théorie à des résultats équivalents à ceux de l'algorithme de la carte de Kohonen pour un voisinage d'un site (voir Haykin, 1994).

simple qui s'écrit, dans le cas le plus utilisé d'une fonction d'activation sigmoïde :

$$s_j = \left(1 + e^{-x_j}\right)^{-1} \quad s_j \in]0,1[$$

avec $x_j = \sum p_{ij} \cdot s_i$

expressions dans lesquelles : i et j correspondent respectivement aux neurones de la couche de gauche et de droite ; s_j est la valeur de sortie du neurone j ; x_j est la valeur d'entrée du neurone j ; p_{ij} est le poids de la connexion ij .

Comme dans l'apprentissage non supervisé, les poids des connexions sont fixés initialement de façon aléatoire. Les vecteurs d'entrée sont ensuite présentés au réseau (de façon déterminée ou stochastique) et les valeurs du neurone de sortie (qu'il soit uni ou multidimensionnel) sont comparées à la valeur réelle contenue dans la base. Une correction, fonction de l'erreur, est alors effectuée sur les poids jusqu'à ce que par itération on arrive à une erreur (moyenne ou maximale) entre valeur prédite en sortie de réseau et valeur réelle (observée dans la réalité sur la base) qui satisfasse les exigences de prédiction. A chaque jeu de poids synaptiques correspond en effet une valeur d'erreur et l'apprentissage s'effectue donc en cherchant, par descente de gradient sur la surface d'erreur un minimum –que l'on espère absolu– de cette surface (le minimum n'est pas forcément unique pour toutes les fonctions d'activation). La descente de gradient se fait par un algorithme de "rétro-propagation" qui, après avoir calculé le taux de variation de l'erreur autour d'un point sur la surface d'erreur, modifie le jeu de poids dans le sens de la descente de façon proportionnelle à la pente. L'apprentissage est considéré comme terminé lorsque, conformément à ce qui est évoqué plus haut, on est arrivé, après un nombre plus ou moins grand d'itérations à un gradient d'erreur nul, ou quasi nul. Le jeu de poids final entre les couches constitue le modèle.

Le principe est séduisant mais les réseaux de neurones supervisés malgré leur grande flexibilité (absence de supposition sur la forme fonctionnelle, absorption de n'importe quelle forme de variable) n'ont pas que des atouts. Leur côté "boîte noire", par essence non explicatif (à l'inverse du réseau non supervisé qui découvre l'ordonnancement de la base) leur est en effet souvent reproché. Le choix d'une structure de réseau (nombre de couches cachées et nombre de neurones dans chaque couche) ne se base pour le moment que sur un certain nombre de règles empiriques. Il faut en effet procéder le plus souvent par tâtonnements avant d'arriver à une structure qui paraisse optimale. Une fois la (ou les) structure(s) choisie(s), l'apprentissage du réseau sur la base n'est pas sans difficulté. Un excès

d'apprentissage sur un échantillon de population conduit le plus souvent à une perte des capacités de généralisation, c'est-à-dire de précision de prédiction sur un individu de la population non contenu dans l'échantillon d'apprentissage, un peu comme si le réseau avait trouvé en convergeant parfaitement " la " solution du problème posé par l'échantillon et, étant allé trop loin dans l'exploitation de l'échantillon, avait absorbé trop de spécificité qui l'empêche de généraliser correctement sur le reste de la population. Là encore des protocoles empiriques permettent d'éviter le sur-apprentissage mais la détermination par validation sur une partie de l'échantillon d'apprentissage du nombre d'itérations correspondant à un apprentissage optimal relève plus de l'artisanat éclairé que d'une démarche scientifique. Les travaux mathématiques sur la mécanique interne des réseaux se poursuivent à grande échelle aux États-Unis, en Europe et au Japon dans les laboratoires de "connexionisme". Ils permettront probablement dans un proche avenir d'une part de rationaliser les architectures et d'autre part d'interpréter et d'utiliser les jeux de poids de connexions de façon plus satisfaisante.

Les champs d'application des réseaux non supervisés ou supervisés se multiplient en effet depuis quelques années. Sans être exhaustif, citons entre autres la reconnaissance de la voix, des signaux et des images, le diagnostic médical, la prédiction sur séries temporelles, la robotique, etc. L'apparition, depuis le début des années quatre-vingt-dix, de logiciels commerciaux gérant des algorithmes ad hoc et la croissance phénoménale des capacités de calcul des ordinateurs courants autorisent dorénavant une démocratisation de l'outil et donc une augmentation notoire du nombre de ses utilisations. On comprend par exemple que compte tenu de la nature des enjeux, la recherche appliquée sur le traitement des séries boursières foisonne. Elle est d'ailleurs largement subventionnée par les milieux financiers et tend à constituer un champ spécifique qui éclipse d'autres applications. L'utilisation des réseaux supervisés sur coupes transversales –en finance comme en économie– progresse peu (cf. section 2.2) et nous n'avons recensé à notre meilleure connaissance qu'une seule recherche mettant à profit les capacités de la carte de Kohonen dans la lecture d'une base de données économique-financière (Martin del Brio et Serrano, 1995). C'est dans ce contexte que nous proposons deux illustrations empiriques. L'une consiste en la projection sur une carte de Kohonen à deux dimensions d'une base qui décrit l'habitat rural marocain. L'autre permet, sur la même base, d'évaluer l'amélioration de performance prédictive sur les prix du logement que l'on peut obtenir avec un réseau supervisé par rapport à ce que l'on obtient avec la régression multiple.

Il convient de rappeler à ce stade que le propos principal de l'article n'est pas de dégager les enseignements économiques que l'on tire de l'examen tant de la typologie de l'habitat rural marocain que de la valorisation des composantes

structurales ou d'équipement de cet habitat. On se concentre donc plus –à travers ces deux applications– sur la mise en évidence de l'intérêt de l'outil que sur l'analyse du contenu intrinsèque de la base sur laquelle on le teste.

2. DEUX ILLUSTRATIONS EMPIRIQUES

Les deux applications sont menées sur la même base de données de panel qui comporte 1852 logements ruraux répartis de façon représentative (au 1/250^{ème}) sur les zones rurales de quatre régions marocaines : Marrakech, Settat, Côte et Meknès⁴. On trouvera en annexe 1 les composantes du vecteur descriptif des logements ainsi que la codification des composantes qui spécifient le logement dans l'application avec le réseau non supervisé.

2.1. Une carte de Kohonen de l'habitat rural marocain

2.1.1. Le protocole empirique

La base de données est réduite à 1511 logements après élimination des logements présentant des valeurs manquantes sur les variables descriptives spécifiées et chacune des douze variables est recalculée sur l'intervalle $[0,1]$ afin d'éliminer les effets d'échelle⁵. La carte est de grande dimension (20 x 20, soient 400 sites) afin de tenir compte d'une part de l'importance de la base et d'autre part de la diversité des possibilités de types de logements décrits par douze variables dont cinq ne sont pas binaires.

La base, quelle que soit sa dimension, (base complète ou sous-bases régionales) est présentée trente fois (soit dans le cas de la base complète, 45 330 calculs de distance) avec un voisinage initial de 10 sites progressivement réduit à 1 et un pas d'apprentissage initial de 1 progressivement réduit à 0,125. A l'issue de cette phase d'ordonnancement, le voisinage et le pas d'apprentissage sont maintenus pendant trente cycles complémentaires respectivement à 1 et 0,0125, ce qui permet une optimisation locale par ajustement du positionnement des vecteurs d'entrée sur des sites voisins. Cette seconde phase peut être poursuivie tant que l'on observe des mouvements locaux. L'algorithme ayant convergé, on stocke les jeux de poids (composantes) des quatre cents sites (ils constituent "le modèle") et on projette une dernière fois la base de données sur la carte.

⁴ La base est décrite plus en détail dans Thuillier (1994).

⁵ D'autres séries d'estimation ont été effectuées à partir de valeurs standardisées (centrées et réduites). Les résultats obtenus ne sont pas sensiblement différents.

2.1.2. Les résultats

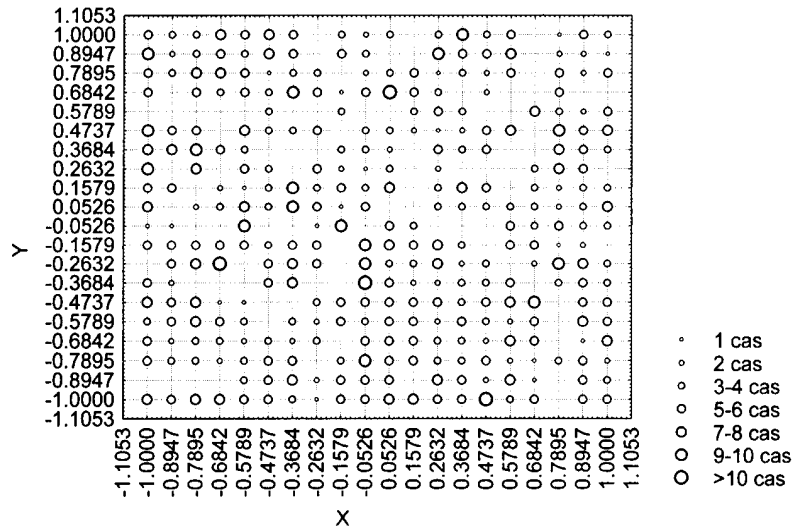
La figure 3 illustre le résultat de la projection de la base complète (quatre régions confondues) sur la carte de Kohonen. Chaque site de la carte constitue un "micro-cluster" qui regroupe un nombre (variable) de logements identiques ou quasi identiques.

Il n'apparaît pas en première lecture de "macro-clusters" importants (accumulations de "micro-clusters" nettement isolées les uns des autres par des sites vides) mais cette homogénéité cache une structure interne dont la lecture suppose l'examen des jeux de poids ou, ce qui revient au même, l'examen des valeurs des composantes des logements appartenant à des sites localisés dans différentes zones de la carte (à convergence les 12 composantes d'un site w_{ij} sont très proches des valeurs des 12 composantes des logements contenus dans le site en question). Le lecteur pourra à titre d'exemple consulter l'annexe 2 où sont regroupées les composantes des logements contenus d'une part dans les quatre micro-clusters situés en losange autour du centre (vide) de la carte, et d'autre part dans les quatre micro-clusters situés à l'extrémité (en haut à gauche) du quadrant nord-ouest. On constatera la grande homogénéité interne des micro-clusters qui ne contiennent que des logements identiques ou quasi identiques, et la très grande similitude des valeurs des composantes des logements contenus dans des sites voisins.

Lorsqu'on poursuit la lecture des composantes de site en site (balayage de la carte), on constate l'existence :

- au centre d'un macro-cluster, de logements qui n'ont ni eau courante ni électricité et qui sont construits en matériaux traditionnels (terre et roseaux, pisé-adobe, pierre et terre). Ce macro-cluster se prolonge au sud-est, sud et sud-ouest avec des variations mineures : c'est ainsi que le coin sud-est regroupe des logements traditionnels de grande surface (nettement supérieure à 100 m²) alors que le coin sud-ouest regroupe les maisons de petite taille (16 à 77 m²) ;
- au nord-ouest d'un macro-cluster (beaucoup plus petit), d'habitations en bon état, construites en matériaux non traditionnels (ciment, briques, etc.) et dotées de l'eau courante, caractéristique qui discrimine les premières rangées nord du reste de la carte.

Figure 3 : Carte de la base complète



Afin d'illustrer le propos ci-dessus, on a imaginé une succession de six maisons rurales fictives (qui n'appartiennent pas à la base) décrites par leurs composantes a priori et on les a projetées sur les jeux de poids issus de l'apprentissage du réseau. Les dénominations⁶ et les composantes de chaque maison sont regroupées à l'annexe 3. La figure 4 montre sur la carte de Kohonen la position des sites qui ont capté ces six archétypes, positions cohérentes avec la lecture précédente et qui permet de comprendre comment l'ordonnancement (de Taudis à Luxe) s'est "fermé" sur la carte à deux dimensions⁷.

Une cartographie régionale de la base peut aussi être réalisée. Les figures 5a et 5b montrent les résultats obtenus en sélectionnant, lors de la projection, les maisons qui appartiennent d'une part à la région de Marrakech (fig. 5a) et d'autre part à la région côtière (fig. 5b). On voit apparaître sur les cartes deux faciès régionaux d'habitat rural très différents (une évidence pour qui connaît ces deux régions) :

- La carte de Kohonen de la région rurale de Marrakech est –exception faite d'un

⁶ Sans doute trop normatives et arbitraires. Elles doivent être considérées ici simplement comme des repères.

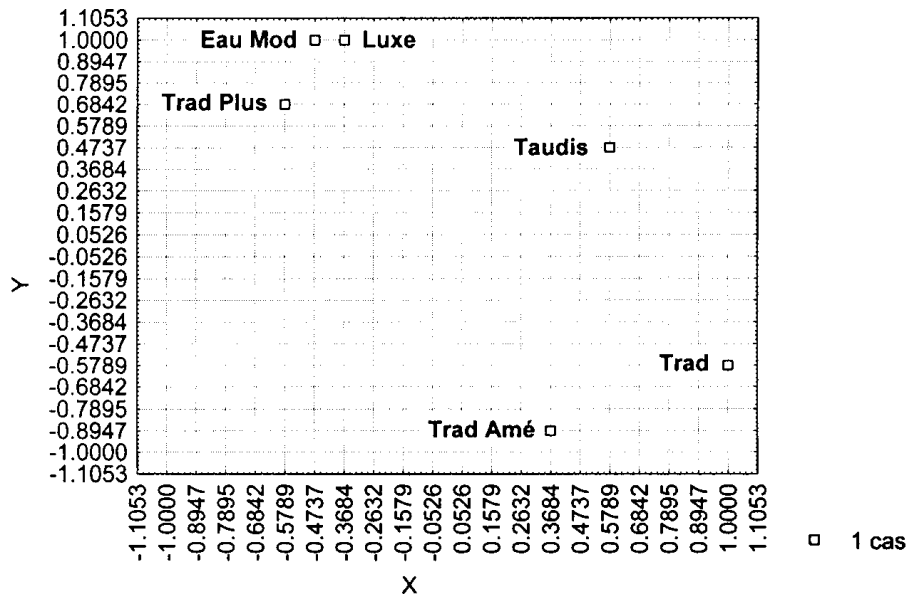
⁷ Une autre initialisation des vecteurs-poids aurait conduit à un point de départ de l'ordonnancement localisé à un autre endroit de la carte.

logement sur 436– vierge dans le coin nord-ouest (moderne avec eau courante), très clairsemée au nord de la carte mais densément peuplée sur toute la moitié sud qui comprend environ 90 % des cas et correspond au logement traditionnel plus ou moins amélioré (avec une tendance vers les grandes surfaces plutôt que les petites).

- La région côtière, qui comprend 250 logements présente un faciès presque inversé. Un cluster modeste de logements traditionnels de petite taille occupe le coin sud-est de la carte alors que toute la partie nord et nord-ouest est densément peuplée (abondance de structures en ciment, adduction d'eau plus fréquente, etc.). L'habitat rural de la région côtière qui comprend encore de l'habitat précaire a cependant clairement "muté" et ce, de façon irréversible. On en connaît les raisons : urbanisation rampante, niveau de vie plus élevé, etc.

Les régions de Settat et de Meknès (non reprises ici) ont été également traitées : elles présentent, comme on peut s'y attendre pour qui connaît les régions rurales marocaines, des faciès intermédiaires (transition progressive de l'habitat rural, du traditionnel au moderne).

Figure 4 : Projection des archétypes d'habitat rural



2.2. La capacité prédictive des réseaux et de la régression multiple

Parallèlement à sa lecture macro-économique⁸, le prix du logement peut être appréhendé à partir d'estimations micro-économétriques "sur cross-sections" de la valeur de chacun des éléments –espace et structures– qui le constituent. Le logement, bien hétérogène, fournit en effet au ménage (propriétaire ou locataire) une somme de flux de services générés par chaque élément de confort, de surface, d'équipement, ou de localisation. Il est donc possible d'expliquer et de prédire le prix d'un logement résidentiel à partir "du prix" de ses caractéristiques (proximité du marché, du métro, de l'école, etc.). L'analyse hédonique du prix du logement trouve ses fondements conceptuels dans les travaux d'Houthakker (1952), Becker (1965) et Lancaster (1966), et sa méthodologie d'application chez Rosen (1974). L'utilisation de cette technique dans les études sur l'économie du logement s'est largement développée dans les années soixante-dix et jusqu'à récemment (pour une bibliographie sur ce point voir Thuillier, 1994). Si x_k est le vecteur des caractéristiques, on écrit simplement $P = f(x_k)$, expression dans laquelle P est la valeur spécifiée par le prix de vente, la valeur locative ou le loyer effectif du logement. Le prix hédonique de la i -ème caractéristique du vecteur x_k est alors défini par la dérivée partielle du prix par rapport à la i -ème caractéristique et son estimation est effectuée par régression multiple, linéaire ou non. Le choix de la forme fonctionnelle est fait le plus souvent de façon arbitraire à partir de l'ajustement statistique et souvent sans référence à une fonction d'utilité quelconque. On peut en effet estimer des formes fonctionnelles flexibles (transformations de Box-Cox) englobant comme cas particuliers les fonctions linéaires ou log-log, et tester la vraisemblance de chaque forme avant d'en exploiter les résultats. Mais, dans la mesure où d'une part la non linéarité de la forme fonctionnelle trouve difficilement une interprétation économique dans ce type d'analyse et où d'autre part son estimation est coûteuse, il paraît opportun d'examiner l'intérêt que peuvent présenter d'autres outils d'estimation de la valeur du logement qui, même s'ils ne contribuent pas –pour le moment– à l'amélioration de la compréhension de la formation du prix, n'en permettent pas moins une prédiction qu'il est intéressant de confronter avec celle obtenue par les techniques classiques.

Il n'existe à notre connaissance que peu de travaux consacrés à l'analyse empirique comparée des performances de la régression multiple et des réseaux de

⁸ Pour une synthèse de la littérature anglo-américaine sur la modélisation macro-économique de la demande de logement, lire entre autres Muth et Goodman (1989). Pour un examen des problèmes posés par sa modélisation dans le cas spécifique des pays en développement, le lecteur pourra consulter Thuillier (1995).

neurones. On pourra toutefois consulter Bansal, Kaufman et Weiz (1993) pour une analyse de l'influence de la qualité des données sur chaque outil dans le cadre d'une analyse des séries temporelles (l'insolvabilité des ménages relativement à leur emprunt-logement) et Curram et Mingers (1994) pour une comparaison entre réseau de neurones, arbres de décision et analyse discriminante linéaire (proche de la régression multiple) sur sept bases différentes. Dans ce dernier cas, conformément aux attentes, le réseau performe sensiblement mieux que l'analyse discriminante lorsque la base n'est pas linéairement séparable. Deux articles traitent par ailleurs spécifiquement de l'utilisation de l'intelligence artificielle et des réseaux de neurones dans l'évaluation des prix immobiliers (Tay et Ho, 1992 ; Jepson, Collins et Evans, 1994) sans pour autant procéder à une comparaison systématique avec les outils classiques.

Disposant de peu d'exemples pratiques sur lesquels se baser et ayant pour objectif de tester essentiellement les robustesses respectives des deux outils, on a élaboré un protocole qui permet d'une part de multiplier les échantillons d'apprentissage des réseaux et d'estimation de la régression multiple et d'autre part de tester les capacités prédictives de chaque modèle sur une partie de la base extérieure à l'échantillon sur lequel on effectue l'apprentissage et l'estimation. Le lecteur pourra se reporter à l'annexe 4 qui contient le détail de la méthodologie utilisée ainsi que les informations relatives à la spécification des variables de l'équation de la valeur du logement et à leur re-codification.

2.2.1. L'estimation des modèles

La régression multiple linéaire sur les quarts, moitiés et trois quarts de la base de 1776 logements ruraux n'appelle pas de commentaire spécifique. Deux types d'estimations ont été effectuées : l'une avec la procédure "stepwise"⁹ et l'autre avec "enter". Dans le second cas, toutes les variables descriptives sont conservées dans le modèle estimé, quelle que soit la significativité des variables.

Les modèles de réseaux de neurones retenus procèdent, ainsi qu'évoqué à la section 1.2, d'une analyse empirique préliminaire sur plusieurs structures a priori. Deux types de structures ont été retenues, l'apprentissage se faisant avec ou sans validation, avec la totalité des variables descriptives (pour le réseau à six neurones en couche cachée) ou avec seulement les variables retenues par la procédure "stepwise" de la régression multiple. Le tableau 1 ci-après permet de lire plus

⁹ Il convient de noter que, suivant les sous-populations de la base sur lesquelles les régressions sont estimées, la régression "stepwise" ne conserve qu'entre cinq et douze variables parmi les quinze spécifiées pour expliquer le prix du logement.

facilement le tableau 2 qui présente les résultats relatifs aux capacités prédictives de chaque modèle.

Figure 5a : Carte de la région de Marrakech

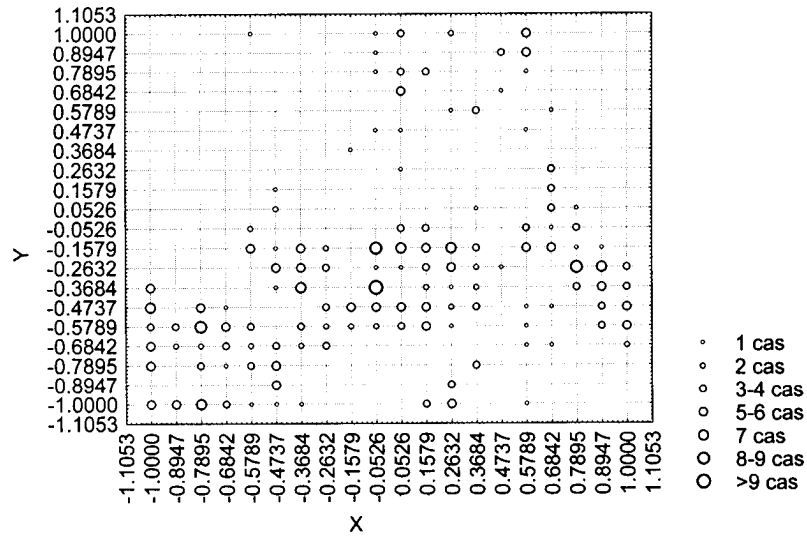


Figure 5b : Carte de la région de la Côte

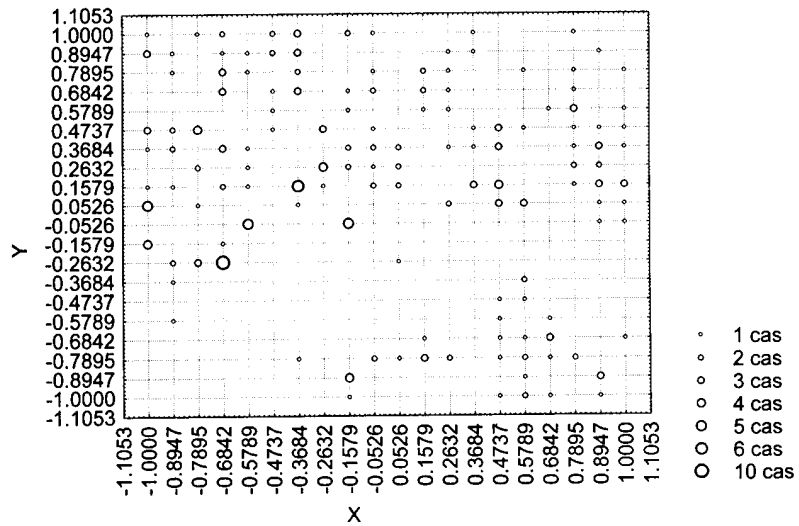


Tableau 1 : Paramètres de modélisation et nomenclature des modèles

Paramètres	
Régression multiple	Réseaux de neurones
<p>Forme fonctionnelle : linéaire</p> <p>Fonction d'erreur : distance euclidienne</p> <p>Algorithme : moindres carrés</p> <p>Procédures :</p> <p>1) stepwise in : 0,05 out : 0,10</p> <p>2) enter</p>	<p>Fonction d'activation : sigmoïde</p> <p>Fonction d'erreur : distance euclidienne</p> <p>Couches cachées : une seule à 1 ou 6 neurones</p> <p>Algorithme : rétro-propagation stochastique</p> $\Delta w_n = l \left(\frac{de}{dw} \right)_n + a \Delta w_{n-1}$ <p>où Δw_n : changement du jeu de poids de l'itération n</p> <p>l : pas d'apprentissage (0,1) ou (0,001)</p> <p>e : erreur totale</p> <p>a : coefficient d'ajustement tenant compte de la variation du jeu de poids de l'itération précédente, ici a = 0,9 quel que soit le pas d'apprentissage</p> <p>Nombre d'itérations : variable</p>
Nomenclature	
<p>RG pour régression</p> <p style="text-align: center;">+</p> <p>S pour "stepwise",</p> <p>E pour "enter".</p>	<p>NN pour réseau de neurones, x (1 ou 6) pour le nombre de neurones en couche cachée, S pour vecteur d'entrée constitué par les variables retenues par la procédure stepwise de la régression multiple, E pour vecteur constitué par les 15 variables descriptives, V s'il y a procédure de validation (sur 25% de la population d'apprentissage), Y si le modèle est non validé, (y) pour le taux d'apprentissage.</p> <p>Ex. : NN1 SY(0,001) correspond au modèle de réseau de neurones (jeu de poids) avec un neurone en couche cachée, estimé avec uniquement les variables sélectionnées par la procédure stepwise de la régression, sans validation, avec un pas d'apprentissage de 0,001.</p> <p>NN6 EV (0,1) correspond au modèle, à 6 neurones en couche cachée, de réseau de neurones estimé avec les 15 variables descriptives avec validation et un pas d'apprentissage de 0,1.</p>

Tableau 2 : Capacités prédictives des modèles selon la population d'estimation et la population-test

		Modèles					
Population d'estimation/apprentissage		Population-test : dernier quart (1333-1776)					
	# des quarts	RGS	RGE	NN6 EV (0,1)	NN1 Sv(0,1)	NN1 SV (0,1)	NN1 Sv(0,001)
75%	123	0,7205	0,7218	0,7111	0,7456	0,6785	0,7584
50%	12	0,6974	0,6958	0,7259	0,7287	0,7197	0,7233
50%	13	0,7153	0,7120	0,6964	0,7514	0,6844	0,7565
50%	23	0,7079	0,7016	0,7026	0,7227	0,6770	0,7400
25%	1	0,6801	0,6810	0,7320	0,7330	0,7298	0,7324
25%	2	0,6748	0,6984	0,7236	0,7305	0,6985	0,7249
25%	3	0,7050	0,6943	0,6954	0,7415	0,7144	0,7343
Population d'estimation/apprentissage		Population-test : troisième quart (899-1332)					
	# des quarts	RGS	RGE	NN6 EV (0,1)	NN1 Sv(0,1)	NN1 SV (0,1)	NN1 Sv(0,001)
75%	124	0,7226	0,7344	0,7031	0,6818	0,6739	0,7352
50%	12		0,7003	0,6991	0,6893	0,7025	0,6966
50%	14	0,7262	0,7237	0,6916	0,7177	0,7271	0,7352
50%	24	0,7136	0,7294	0,7170	0,7210	0,6522	0,7278
25%	1	0,6910	0,6990	0,6681	0,7078	0,7032	0,6939
25%	2	0,6943	0,7051	0,6970	0,6669	0,6844	0,6979
25%	4	0,6577	0,6516	0,6617	0,6572	0,6963	0,6327
Population d'estimation/apprentissage		Population-test : second quart (445-888)					
	# des quarts	RGS	RGE	NN6 EV (0,1)	NN1 Sv(0,1)	NN1 SV (0,1)	NN1 Sv(0,001)
75%	134	0,7725	0,7708	0,7772	0,7931	0,8069	0,8037
50%	13	0,7714	0,7694	0,7223	0,7901	0,7038	0,8006
50%	14	0,7533	0,7613	0,7445	0,7570	0,7716	0,7733
50%	34	0,7578	0,7452	0,7457	0,7569	0,7485	0,7642
25%	1	0,7583	0,7625	0,7802	0,7677	0,7936	0,7987
25%	3	0,7591	0,7390	0,7172	0,7827	0,7560	0,7876
25%	4	0,6934	0,6650	0,7278	0,6703	0,7376	0,6733
Population d'estimation/apprentissage		Population-test : premier quart (001-444)					
	# des quarts	RGS	RGE	NN6 EV (0,1)	NN1 Sv(0,1)	NN1 SV (0,1)	NN1 Sv(0,001)
75%	234	0,7256	0,7372	0,7426	0,7583	0,7553	0,7627
50%	23	0,7277	0,7382	0,6969	0,7690	0,6791	0,7789
50%	24	0,7233	0,7359	0,7739	0,7703	0,6838	0,7750
50%	34	0,7199	0,7106	0,7394	0,7523	0,7498	0,7537
25%	2	0,7225	0,7403	0,7871	0,7632	0,7647	0,7776
25%	3	0,6929	0,7059	0,7047	0,7616	0,7259	0,7634
25%	4	0,6659	0,6482	0,6690	0,6630	0,7231	0,6690
Moyenne		0,7164	0,7178	0,7198	0,7340	0,7193	0,7418
Écart type		0,031	0,032	0,033	0,040	0,039	0,042

On rappelle que quel que soit le modèle estimé sur un échantillon de un, deux, ou trois quarts de base, sa capacité prédictive est donnée par le coefficient de régression entre la série des valeurs de logements prédites et celle observée sur le ou les quarts non contenus dans l'échantillon d'estimation (ou d'apprentissage). La capacité prédictive d'un modèle donné varie, à l'évidence, avec l'échantillon sur lequel il est testé.

2.2.2. Les résultats empiriques

Le tableau 2 résume les résultats obtenus. Les deux modèles de gauche correspondent à la régression multiple linéaire, les quatre autres montrent la performance de réseaux répertoriés suivant la nomenclature du tableau 1. Les capacités prédictives moyennes et les écarts-types sont notés en dessous des colonnes correspondantes à chaque modèle. Un survol de ce tableau permet d'avoir une vue d'ensemble des performances des modèles. Les meilleures prédictions (corrélations la plus élevée entre série prédite et réelle) figurent en caractères gras. Il apparaît en première lecture que les réseaux à un neurone (notamment sans validation et à apprentissage lent) montrent une nette supériorité tant par rapport aux autres réseaux que par rapport à la régression multiple.

L'observation rapide du tableau 2 ne permet toutefois pas de conclure sans un examen statistique des séries. On a donc effectué des tests de Student sur les différences de capacités prédictives moyennes de chaque modèle, ceux-ci étant comparés deux à deux. On obtient quinze tests de significativité qui permettent d'affirmer :

- qu'un réseau à un neurone à fonction d'activation sigmoïde sans validation permet une meilleure prédiction du prix du logement que les régressions multiples quelle que soit la procédure (stepwise ou enter).
- que la diminution du pas d'apprentissage dans le cas d'un réseau à un neurone améliore la capacité de généralisation du modèle (mais augmente la durée de "l'apprentissage"...).)
- que l'augmentation du nombre de neurones en couche cachée (dans ce cas la validation sur une partie de la base d'apprentissage est obligatoire) semble altérer la capacité prédictive du réseau (alors que, on a pu le vérifier, elle améliore la performance sur la base d'apprentissage...). Il faut en déduire que les non linéarités "découvertes" par le réseau lors de l'apprentissage appartiennent aux échantillons et, n'étant pas toutes identiques, ne sont pas généralisables sur la base.

CONCLUSION

Les résultats ci-dessus ne sont pas exhaustifs. Une estimation systématique des réseaux intermédiaires (à deux neurones, trois neurones, etc.) n'a pas été effectuée pas plus que n'ont été testées de façon méthodique les possibilités qu'offre l'adjonction d'une (ou plusieurs) couche cachée supplémentaire. De meilleures performances, comme de moins bonnes, auraient pu être obtenues. Il convient donc, pour le moment, de considérer les réseaux supervisés comme un outil complémentaire à ceux dont dispose le chercheur dans le traitement des données sur coupe transversale (comme d'ailleurs dans celui des séries longitudinales). Leur utilisation se justifie lorsque la précision de la prédiction – plus que la qualité de l'explication – est en jeu. C'est le cas par exemple en matière d'évaluation municipale, où la justesse (et la facilité de révision) de la base de la taxation foncière ou immobilière importe plus que sa compréhension. Les enjeux, on le comprend, concernent sur ce sujet plus l'urbain que le rural et –pour le moment– les pays avancés plus que les pays en développement. Mais les réseaux peuvent trouver en pays en voie de développement d'autres applications à plus court terme (prédiction de faillite, octroi des prêts aux particuliers ou aux entreprises, etc.) d'un intérêt opérationnel manifeste quel que soit le niveau de développement d'une économie nationale.

Les réseaux non supervisés qui n'ont pas fait de leur côté l'objet d'une grande attention en économie offrent pourtant des possibilités que nous n'avons fait qu'effleurer dans l'illustration proposée. La carte de Kohonen peut en effet être rendue dynamique. Si l'on dispose par exemple d'une suite (quelques années ou quelque périodes) de coupes transversales, il est possible, sur la série des cartes, de suivre un individu donné de la base de période en période, et d'analyser son cheminement. Pensons ici à une cohorte d'entreprises jeunes (PME, entreprises de haute technologie, etc.) définies par quelques variables (effort de recherche et développement, exportations, localisation, productivité, etc.) Les cartes pourront le cas échéant faire apparaître des tendances (regroupements systématiques, cheminements similaires) qui même si elles ne sont pas significatives au sens statistique du terme, n'en constituent pas moins autant de thèmes de réflexion qui peuvent s'avérer fructueux. La carte de Kohonen peut aussi être couplée avec un réseau supervisé qui relie chaque site à une variable dépendante continue ou binaire. Le réseau non supervisé agit alors comme un filtre qui à de nombreuses données d'entrée en substitue une seule (le micro-cluster défini par ses composantes à convergence) et élimine ainsi le bruit contenu dans les bases sans perte importante d'information. Le couplage d'un réseau non supervisé avec un réseau supervisé constitue un ensemble intégré d'analyse et de prédiction qui, sans remplacer les outils existants, propose une alternative intéressante au chercheur.

On comprend par ailleurs l'intérêt qu'un tel couplage peut présenter pour des usages plus commerciaux et prévoir une croissance rapide de ses utilisations pratiques.

ANNEXE 1
VARIABLES DESCRIPTIVES DE L'HABITAT RURAL

Nomenclature		Codification	Statistiques descriptives (n = 1 511)		
			Moyenne	Écart-type	Médiane
AGELOG	Age du logement	Continue	24,56	19,68	20,0
CUIS	Existence d'une cuisine intérieure	(0, 1)	0,78	0,42	1,0
DEPEND	Existence de dépendances	(0, 1)	0,73	0,44	1,0
EAU	Eau courante (publique ou privée)	(0,1)	0,35	0,48	0,0
ECLAIR	Éclairage - bougie - huile - génératrice - éclairage public	(1, 2, 3, 4) 1 2 3 4	2,04	0,90	2,0
ETATLOG	État du logement	(0, 1)	0,71	0,45	1,0
MURS	Matériaux des murs - terre et roseaux - pisé adobe - pierre et terre - pierre et ciment - briques ou ciment aggloméré	(1, 2, 3, 4, 5) 1 2 3 4 5	2,94	1,05	3,0
PIECE	Nombre de pièces	Continue	3,03	1,31	3,0
SANIT	Évacuation sanitaire - aucune toilette - latrine - fosse sceptique - tout à l'égout	(0, 1, 2, 3) 0 1 2 3	0,41	0,77	0,0
SOL	Nature du sol - terre battue - ciment - carrelage ou mosaïque	(1, 2, 3) 1 2 3	1,55	0,56	2,0
SUPCOUV	Superficie couverte de la maison	Continue	67,12	51,03	55,0
TOIT	Nature des matériaux de couverture - terre et roseaux - tôle de zinc - terre cuite, béton	(1, 2, 3) 1 2 3	1,83	0,79	2,0

ANNEXE 2

Caractéristiques des logements contenus dans les micro-clusters du centre
(base complète de la figure 3)Coordonnées des
micro-clusters sur la
carte de Kohonen

AGELOG	CUIS	DEPEND	EAU	ECLAIR	ETATLOG	MURS	PIECE	SANIT	SOL	SUPCOUV	TOIT	Coordonnées des micro-clusters sur la carte de Kohonen	
												X	Y
20	1	1	0	1	1	2	2	0	1	25	1	-0.052631	-0.157895
5	1	1	0	1	1	2	2	0	1	50	1	-0.052631	-0.157895
6	1	1	0	1	1	2	1	0	1	22	1	-0.052631	-0.157895
15	1	1	0	1	1	2	2	0	1	40	1	-0.052631	-0.157895
20	1	1	0	1	1	2	2	0	1	40	1	-0.052631	-0.157895
14	1	1	0	1	1	2	2	0	1	40	1	-0.052631	-0.157895
3	1	1	0	1	1	3	2	0	1	25	1	-0.052631	-0.157895
4	1	1	0	1	1	2	1	0	1	40	1	-0.052631	-0.157895
7	1	1	0	1	1	2	2	0	1	50	1	-0.052631	-0.157895
7	1	0	0	2	1	2	2	0	1	33	2	-0.157895	-0.052631
4	1	0	0	2	1	2	2	0	1	27	2	-0.157895	-0.052631
20	1	0	0	2	1	2	2	0	1	30	2	-0.157895	-0.052631
8	1	0	0	2	1	2	2	0	1	30	2	-0.157895	-0.052631
30	1	0	0	2	1	2	1	0	1	25	2	-0.157895	-0.052631
25	1	0	0	2	1	2	2	0	1	55	2	-0.157895	-0.052631
7	1	0	0	2	1	2	2	0	1	36	2	-0.157895	-0.052631
5	1	0	0	1	1	1	2	0	1	32	2	-0.157895	-0.052631
15	1	0	0	2	1	2	1	0	1	24	2	-0.157895	-0.052631
5	1	0	0	1	1	3	3	0	1	40	2	-0.052631	0.052632
4	1	0	0	2	1	2	2	1	1	26	2	-0.052631	0.052632
10	1	0	0	2	1	3	1	0	1	16	2	-0.052631	0.052632
35	1	0	0	2	1	3	2	1	1	39	2	-0.052631	0.052632
3	1	0	0	1	1	3	2	0	1	40	2	-0.052631	0.052632
7	1	0	0	2	1	3	3	1	1	48	2	-0.052631	0.052632
40	1	1	0	1	1	2	2	0	1	40	1	0.052632	-0.052631
30	1	1	0	1	1	2	2	0	1	43	1	0.052632	-0.052631
30	1	1	0	1	1	2	2	0	1	70	1	0.052632	-0.052631
50	1	1	0	1	1	2	2	0	1	60	1	0.052632	-0.052631
35	1	1	0	1	1	3	1	0	1	23	1	0.052632	-0.052631
6	1	1	1	1	1	5	3	0	2	100	2	-1	1
15	1	1	1	1	1	5	2	0	2	38	2	-1	1
20	1	1	1	1	1	4	2	0	2	30	2	-1	1
1	1	1	1	2	1	5	1	0	2	24	2	-1	1
5	1	1	1	1	1	5	1	0	2	15	2	-1	1
15	1	1	1	1	1	3	2	0	2	40	3	-1	0.894737
10	1	1	1	1	1	3	2	0	2	31	3	-1	0.894737
10	1	1	1	1	1	3	2	0	2	34	3	-1	0.894737
15	1	1	1	1	1	3	3	0	2	60	3	-1	0.894737
20	1	1	1	1	1	3	3	0	2	33	3	-1	0.894737
10	1	1	1	1	1	3	3	0	2	45	3	-1	0.894737
5	1	1	1	1	1	3	2	0	2	60	3	-1	0.894737
20	1	1	1	1	1	3	2	0	2	27	3	-1	0.894737
20	1	1	1	2	1	3	2	0	2	37	3	-1	0.894737
33	1	1	1	1	1	5	5	0	3	95	3	-0.894737	0.894737
20	1	1	1	1	1	5	4	0	2	84	3	-0.894737	0.894737
5	1	1	1	1	1	5	1	0	2	24	3	-0.894737	1
5	1	1	1	2	1	5	2	0	2	40	3	-0.894737	1
30	1	1	1	1	1	4	1	0	3	20	3	-0.894737	1
5	1	1	1	1	1	5	1	0	2	21	3	-0.894737	1

ANNEXE 3
ARCHÉTYPES D'HABITAT RURAL

	Abbré- viation (carte)	Valeurs des variables descriptives												Coordonnées de la projection sur la carte de Kohonen	
		AGEL OG	CUIS	DEPE ND	EAU	ÉCLAI R	ÉTATL OG	MURS	PIECE	ANIT	SOL	SUPC OUV	TOIT	X	Y
Taudis ou habitat insalubre	Taudis	25	0	0	0	1	0	1	1	0	1	10	1	0,5789 48	0,4736 84
Tradition-nel	Trad	25	1	1	0	1	0	2	3	0	1	67	1	1 - 0,5789 47	
Tradition-nel amélioré	Trad Ame	25	1	1	0	1	1	3	3	0	2	67	2	0,3684 21	- 0,89473 7
Tradition-nel amélioré, eau, éclairage	Trad Plus	25	1	1	1	4	1	3	3	1	2	67	2	- 0,5789 47	0,6842 11
Moderne avec eau courante	Eau Mod	10	1	1	1	3	1	4	3	1	3	67	3	- 0,4736 84	1
Luxeux	Luxe	5	1	1	1	4	1	5	6	3	3	100	3	- 0,3684 21	1

ANNEXE 4
PROTOCOLE EMPIRIQUE DE L'ANALYSE COMPARATIVE

- a) Mélange aléatoire de la base complète (1776 logements ont été cette fois conservés).
- b) Tirage aléatoire de quatre quarts (segmentation) de base, trois fois.
- c) Analyse de la représentativité de chacun des trois tirages par rapport à la base totale sur moyenne médiane et variance du prix du logement, et sur l'appartenance régionale du logement.
- d) Choix de la base segmentée la plus représentative sur chacun de ces quarts.
- e) Établissement des régressions hédoniques "stepwise" sur chaque quart, sur les six "moitiés" de base obtenues combinant les quarts deux à deux et sur les quatre groupes de "trois quarts de base" possibles.
- f) Avec chaque modèle hédonique obtenu sur chacun des quatorze échantillons décrits en (e), on établit la valeur prédite du logement pour chaque logement du ou des quart(s) non contenu(s) dans l'échantillon qui a servi à établir le modèle.
- g) Sur chacun des quatorze échantillons ayant servi à établir les modèles hédoniques (échantillons d'apprentissage), on fait converger deux types de réseaux :
- un réseau à une couche cachée avec un neurone (avec ou sans validation interne sur l'échantillon d'apprentissage),
 - un réseau à une couche cachée avec plusieurs neurones (six neurones dans le cas présent) avec validation interne sur l'échantillon d'apprentissage.
- h) Chaque modèle (jeu de poids synaptiques à convergence) est stocké et testé sur le ou les quarts non contenus dans l'échantillon d'apprentissage.
- i) Chaque série de valeur réelle de logement sur chaque quart est confrontée à la série de valeurs prédites sur ce même quart :
- par le modèle hédonique de régression multiple,
 - par chaque réseau de neurones.
- Le niveau de précision de la capacité prédictive de chaque modèle est défini par une corrélation (Pearson) entre séries réelles et prédites.
- j) Un test sur les différences de coefficient de corrélation est effectué pour statuer significativement sur les différences de capacités prédictives.

Variables spécifiées dans l'équation du prix de l'habitat rural (VALHAB). Outre les variables régionales, on a ajouté aux mêmes variables explicatives :

- *SUPTER* : superficie du terrain (continue)
- *SDB* : existence d'une salle de bain équipée (0,1)
- *TOIL* : existence de WC intérieurs à la maison (0,1).

La variable dépendante VALHAB est donc régressée sur quinze variables. Les variables catégorielles (telles qu'elles se présentent dans le questionnaire initial) ont été recodées en variables binaires.

RÉFÉRENCES

- Bansal A., Kauffman R.J. et Weitz R.R., 1993, "Comparing the Modeling Performance of Regression and Neural Networks as Data Quality Varies: a Business Value Approach", *Journal of Management Information Systems*, n° 10, (1), p. 11-32.
- Beale R. et Jackson T., 1990, "*Neural Computing, an Introduction*", Adam Hilger ed., Bristol.
- Becker G.S., 1965, "A Theory of the Allocation of Time", *Economic Journal*, n° 75, p. 493-517.
- Collins A. et Evans A.H., 1994, "Aircraft Noise and Residential Property Values: an Artificial Neural Network Approach", *Journal of Transport Economics and Policy*, p. 175-197.
- Curram S., et Mingers J., 1994, "Neural Networks, Decision Tree Induction and Discriminant Analysis; an Empirical Comparison", *Journal of the Operational Research Soc.*, Vol. 45, n° 4, p. 440-450.
- Davalo E., et Naïm P., 1990, "*Des réseaux de neurones*", Eyrolles, Paris.
- Dayhoff J.E., 1990, "*Neural Network Architecture: An Introduction*", Van Nostrand Reinhold, New York.
- Gallant S.I., 1993, "*Neural Network Learning and Expert Systems*", MIT Press, Cambridge, Mass.
- Haykin S., 1994, "*Neural Network: a Comprehensive Foundation*", Macmillan College Publishing Company, New York.
- Hecht-Nielsen R., 1990, "*Neuro Computing*", Addison Wasley Publishers, Reading, Mass.
- Houthakker H.S., 1952, "Compensated Changes in Quantities and Qualities Consumed", *Review of Economic Studies*, n° 19, p. 155-164.
- Jepson B., Collins A., et Evans A.H., 1993, "Post Neural Network Procedure to Determine Expected Prediction Values and their Confidence Limits", *Neural Computing and Applications*, n° 1 (3), p. 224-228.

- Jodouin J.F., 1994, "*Les réseaux de neurones, principes et définition*", Hermes éd., Paris.
- Kohonen T., 1990, "The Self-Organizing Map", *Proceedings of IEEE*, n° 78, p. 1464-1480.
- Kohonen T., 1982a, "Self-Organized Formation of Topologically Correct Feature Maps", *Biological Cybernetics*, n° 43, p. 53-69.
- Kohonen T., 1982b, "Clustering, Taxonomy, and Topological Maps of Patterns", *Proceedings of the 6th International Conference on Patterns Recognition*, Munich, Germany, p. 114-128.
- Lancaster K., 1966, "A New Approach to Consumer Theory", *Journal of Political Economy*, n° 74, p. 132-157.
- Martin del Brio B. et Serrano C., 1995, "*Self-Organizing Neural Networks: the Financial State of Spanish Companies*", in *Neural Network in Capital Markets*, John Wiley & Sons Ltd, p. 341-357.
- Muth R.F. et Goodman A.C., 1989, "*The Economics of Housing Markets*", Harwood Academic Publishers, New York.
- Rosen S., 1974, "Hedonic Prices and Implicits Markets: Products Differentiation in Pure Competition", *Journal of Political Economy*, n° 82, p. 34-55.
- Rummelhart D.E., et McClelland J.L., 1986, "*Parallel Distributed Processing*", Vol. 1., Bradford Press, M.I.T.
- Tam K.Y., et Kiand M.Y., 1992, "Managerial Applications of Neural Networks: the Case of Bank Failure Prediction", *Management Science*, n° 38, (7), p. 925-947.
- Tay D.P.H. et Ho D.K.H., 1992, "Artificial Intelligence and the Mass Appraisal of Residential Apartments", *Journal of Property Valuation and Investment*, n° 10, (4), p. 525-540.
- Thuillier D., 1994, "La modélisation micro-économétrique de la demande de logement. Apports d'une analyse appliquée au Maroc", *Revue d'économie du développement*, n° 2, p. 65-96.

Thuillier D., 1995, "Les modèles de logement en pays en voie de développement", *Région et Développement*, n° 2, p. 45-71.

Wasserman P.D., 1989, "*Neural Computing*", Van Nostrand Reinhold, New York.

Abstract

Although the utilization of neural networks has become widespread, they have not as yet been the object of diverse applications in economic analysis. The paper expounds the principles which govern the architecture and functioning of unsupervised and supervised networks, and goes on to propose two empirical illustrations using the same rural Moroccan housing base. The projection of this base on a Kohonen map shows clearly the analytical characteristics of the self-organizing map which, besides representing a comprehensive alternative to traditional techniques, should also arouse greater interest in the future. Moreover, it appears that taking into account the data base used, the prediction capacities of the simple supervised networks are significantly superior to those of multiple regressions. The explanatory capacity of the supervised networks remains limited however and their assessment is not readily practicable.

Resumen

Las redes de neuronas, cuyas utilizaciones se van multiplicando, no han sido hasta ahora el objeto de numerosas aplicaciones en análisis económico. El artículo expone los principios que dirigen la arquitectura y el funcionamiento de las redes no-supervisadas y supervisadas, y después propone dos ilustraciones empíricas en una misma base de viviendas rurales marroquíes. La proyección de esta base sobre un mapa de Kohonen evidencia las características analíticas del mapa auto-organizado que constituye una alternativa operacional frente a las técnicas clásicas y tendrían en adelante que suscitar un más importante interés. Por otra parte, aparece que, en la base de datos utilizados, las capacidades de predicción de las redes supervisadas sencillas son significativamente superiores y las de las regresiones múltiples. La capacidad explicativa de las redes supervisadas queda sin embargo limitada y su estimación resulta algo difícil.