

BULETINUL INSTITUTULUI POLITEHNIC DIN IAȘI  
Publicat de  
Universitatea Tehnică „Gheorghe Asachi” din Iași  
Volumul 64 (68), Numărul 2, 2018  
Secția  
CONSTRUCȚII. ARHITECTURĂ

## UNE ETUDE SUR L'UTILISATION DU RAISONNEMENT A PARTIR DE CAS ET DES SYSTEMES DE REGLES FLOUES DANS L'EVALUATION DU PRIX IMMOBILIER

PAR

MARCO AURELIO STUMPF GONZÁLEZ<sup>1,\*</sup> and  
CARLOS TORRES FORMOSO<sup>2</sup>

<sup>1</sup>Universidade do Vale do Rio dos Sinos

<sup>2</sup>Universidade Federal do Rio Grande do Sul

Reçu le 9 mai 2018

Accepté pour publication le 26 June 2018

**Resumé.** Les biens immobiliers sont des biens différenciés et il faut tenir compte des principales caractéristiques pour l'évaluation immobilière. Les effets simultanés de plusieurs caractéristiques provoquent des difficultés dans la comparaison directe entre les cas et, par conséquent, il est nécessaire d'obtenir un modèle numérique pour évaluer ces différences. L'évaluation immobilière peut être améliorée en séparant un échantillon plus homogène d'une grande base de données contenant les cas de plus grande similitude. L'étude présente la proposition d'un modèle hybride, combinant le raisonnement à partir de cas, systèmes de règles floues et analyse de régression, avec une application pour le marché de Porto Alegre, a ville située au Brésil. Les résultats indiquent la faisabilité du système.

**Mots-clés:** marché immobilier; évaluation de la propriété; raisonnement à partir de cas; systèmes de règles floues.

### 1. Introduction

Les biens immobiliers sont des biens hétérogènes dans le sens où le prix est une composition d'effets de diverses caractéristiques. La diversité des

---

\*Corresponding author: *e-mail*: mgonzalez@unisinos.br

caractéristiques de l'immobilier nécessite la recherche d'un ensemble de données similaire. Cette technique est connue comme la méthode de comparaison des prix pratiqués sur le marché pour des biens semblables. Habituellement, les évaluations utilisent des modèles de prix hédoniques, estimés par l'analyse de régression. La qualité de l'évaluation immobilière dépend de la sélection appropriée de l'échantillon (Appraisal Institute, 2013; Rosen, 1974).

La disponibilité progressive des données provenant directement du format numérique a permis la constitution de bases de données avec des milliers de cas, par exemple dans les administrations publiques (taxes de vente) et les banques (financement immobilier). Dans cette situation, il existe une disponibilité des données, mais la sélection des cas pertinents n'est pas une tâche triviale. La sélection manuelle n'est pas pratique, en raison de la possibilité d'erreur dans la sélection et par la dépense de temps de l'évaluateur et le coût qui en découle.

De plus, la disponibilité des données permet l'utilisation d'outils d'intelligence artificielle, qui peuvent servir d'alternative aux modèles de régression multiple. Certaines études présentent des applications de ces techniques, telles que Abidoye & Chan (2017), Amri & Tularam (2012), Bagnoli & Smith (1998), D'Amato & Kauko (2008), Gerek (2014), González & Formoso (2006), Kauko & D'Amato (2017), Król *et al.* (2008), Melanda, Hunter & Barry (2016), Musa *et al.* (2013), O'Roarty *et al.* (1997), Pacharavanich *et al.* (2000), Taffese (2007), Xie *et al.* (2010), et Zurada *et al.* (2011).

En termes spécifiques, certains travaux présentant des applications dans le domaine de l'évaluation des biens peuvent être cités (Cruz & Morais, 2000; Gasparini & Escudero, 2003; Kauko, 2004, 2008; Rezayan *et al.*, 2005; Król *et al.*, 2007a; Król *et al.*, 2007b; D'Amato, 2007; Król *et al.*, 2008; Lasota *et al.*, 2008; Lasota *et al.*, 2008; Graczyk, Lasota & Trawinski, 2009; Lasota *et al.*, 2009; Lasota *et al.*, 2011; Hefferan & Boyd, 2010; Lasota & Mazurkiewicz, 2010; Tomić, 2010; Furtado, 2011; Guan, Zurada & Levitan, 2011; Lughofer *et al.*, 2011; Amri & Tularam, 2012; Muggenhuber *et al.*, 2013; Telec *et al.*, 2013; Aguiar *et al.*, 2014; Vada, 2014; Braga & Alves, 2015; Abidoye & Chan, 2016; Horváth *et al.*, 2016; Alkhatib *et al.*, 2016; Abidoye & Chan, 2017; Alcantud *et al.*, 2017; Yakubovskiy *et al.*, 2017; Chen *et al.*, 2017).

Lors du choix des techniques, il est important de tenir compte simultanément des exigences de l'évaluation immobilière, telles que l'exactitude des estimations et la facilité d'explication des modèles utilisés. Les réseaux de neurones artificiels, par exemple, ne présentent pas de modèle explicite et impliquent un plus grand degré de complexité dans le développement des applications (González, 2002).

Le processus d'évaluation de l'immobilier a une certaine subjectivité, avec l'influence de l'évaluateur sur le processus de sélection des cas et les attributs des évaluations. Dans ce cas, le développement d'un format d'analyse de données systématique plus objectif peut aider à augmenter la précision.

Compte tenu de l'importance des évaluations et du marché immobilier, en termes économiques et sociaux, il est important de chercher des alternatives pour améliorer les évaluations. Cet article propose l'utilisation de raisonnement à partir de cas pour la sélection des cas et des systèmes de règles floues pour l'évaluation. L'étude présente la proposition et une application du modèle, en comparant l'analyse de régression traditionnelle avec les systèmes de règles floues.

## **2. Le raisonnement à partir de cas**

Le raisonnement à partir de cas (RàPC) (ou case-based reasoning, CBR) est une technique de résolution de problèmes fondée sur la connaissance basée sur la réutilisation des expériences antérieures (synthétisées dans les « cas »). Contrairement à d'autres techniques basées sur la connaissance qui résolvent les problèmes d'une base de connaissances générale, RàPC s'intéresse à des expériences spécifiques, représentées par les cas contenus dans la base de données. Les hypothèses de RàPC sont que (a) des problèmes similaires ont des solutions similaires et (b) la réutilisation des solutions des cas précédents est plus appropriée qu'une solution basée sur les règles (Aamodt & Plaza, 1994 ; Kolodner, 2014; Leake, 2003). Selon Riesbeck et Schank (2013), les avantages de RàPC incluent le fait que la méthode ne dépend pas d'un modèle explicite pour résoudre le problème, la flexibilité de travailler avec de grandes quantités de données et la possibilité d'apprendre avec de nouveaux cas, étant relativement facile à entretenir le système mis à jour.

Le raisonnement à partir de cas est un processus cyclique, composé essentiellement de quatre phases, les 4 R (récupérer, réutiliser, réviser, retenir) : (a) récupération de cas similaires, (b) réutilisation de l'information en tant que solution proposée, c) révision de la solution et (d) rétention d'expérience pour une utilisation future (Aamodt & Plaza, 1994).

## **3. Systèmes basés sur les règles floues**

Les systèmes basés sur les règles floues (SBRF) ont connu un développement théorique et pratique important, en particulier avec les systèmes hybrides. Un SBRF consiste en un ensemble de règles qui utilisent la logique floue dans les parties antécédentes ou conséquentes des règles, et l'une des caractéristiques intéressantes de ces systèmes est la possible conjugaison des

effets de plusieurs règles pour obtenir le résultat final (Cordón *et al.*, 2001; Fourati, 2017).

Un ensemble flou est une généralisation du concept classique d'ensembles, dans lequel la fonction caractéristique n'a pris que deux valeurs,  $\{0,1\}$ , et est caractérisé par une fonction d'appartenance  $\mu_A(x)$ , qui suppose une valeur dans l'intervalle continue  $[0,1]$ . La fonction d'appartenance d'un ensemble flou est une généralisation de la fonction d'indicateur dans les ensembles classiques. En logique floue, il représente le degré de vérité de l'association de  $x$  à l'ensemble  $A$  (Zadeh, 1965).

Un SBRF a deux composantes: (a) le système d'inférence, qui effectue le processus d'inférence nécessaire pour générer une sortie lorsqu'une entrée donnée est spécifiée, et (b) la base de connaissances, représentant les connaissances sur le problème à résoudre, consistant en une collection de règles. Une règle floue au format de l'équation 1 (Cordón *et al.*, 2001).

$$R_i: \text{SI } X \text{ est } A_i \text{ ALORS } Y \text{ est } B_i. \quad (1)$$

La partie « Si  $X$  est  $A_i$  » est appelée antécédent, et la partie « Alors  $Y$  est  $B_i$  » est appelée une conséquente de la règle  $R_i$ . Le règle est composé d'antécédent représenté par des variables linguistiques, étant conséquente une fonction des variables d'entrée, généralement une fonction linéaire, comme celle de l'équation 2 (Cordón *et al.*, 2001) :

$$R_i: \text{SI } X_1 \text{ est } A_1 \text{ et } \dots \text{ et } X_n \text{ est } A_n \text{ ALORS } Y_i = f(Z_1, \dots, Z_n). \quad (2)$$

Où  $X$  sont les variables d'entrée,  $A$  sont des ensembles flous spécifiant leur signification,  $Y_i$  est le résultat de la règle  $R_i$ , et les  $Z$  sont les variables de la fonction  $f$ . L'ensemble  $A$  peut être composée des variables linguistiques associées à des ensembles flous ou directement des variables flous. Les fonctions ne sont pas obtenues directement par le système de règles, mais sont généralement estimées par régression ou algorithmes génétiques, en utilisant une base de connaissances (Cordón *et al.*, 2001; Fourati, 2017).

Le résultat de calcul pour une situation étudiée en considérant un ensemble de règles est obtenu par une moyenne pondérée des règles du système, en considérant l'appartenance du cas étudié pour chacune des règles.

L'utilisation d'un SBRF est intéressante pour permettre la continuité des modèles. Il est courant de diviser le marché total en sous-marchés, en développant un modèle pour chaque partie. Cependant, les cas proches des limites de ces segments peuvent avoir des valeurs différentes dans chaque modèle, ce qui n'est pas souhaitable. Le SBRF permet le développement d'un système d'évaluation continu.

#### 4. Méthode de recherche

Nous utilisons une combinaison de techniques, être raisonnement à partir de cas (RàPC) pour sélectionner les cas de la base de données, et puis les régressions et les systèmes de règles floues.

Pour ce format d'évaluation, un cycle d'évaluation distinct et plus agile basé sur les concepts de RàPC a été proposé, consistant en la sélection d'un échantillon de cas semblable au cas donné (le bien sur l'évaluation) dans la phase de récupération avec la génération de modèles avec techniques auxiliaires dans la phase d'adaptation des cas. Les modèles généraux du marché peuvent contribuer en générant des estimations de valeur accélérées, importantes pour guider l'analyste. La vérification de la solution est difficile sur le marché immobilier car il existe un espace de libre négociation entre les parties. Cependant, la valeur peut être vérifiée à partir des intervalles de confiance de l'estimation. Par exemple, deux écarts-types peuvent être considérés pour chaque côté de l'estimation (ce qui équivaut à environ 95% de la loi normale). La solution proposée peut être confirmée lorsque le prix est compris dans cette fourchette. Dans ce cas, la rétention de cas (apprentissage) se produit.

Dans le marché immobilier, l'emplacement et de nombreux autres attributs sont importants. Le site est un attribut qui ne peut pas être mesuré directement, mais on sait que deux propriétés à proximité tendent à avoir la même accessibilité et la même qualité de voisinage, diminuant la similarité à mesure que la distance augmente. Ainsi, la similitude de localisation peut être mesurée en utilisant la distance entre eux. Pour les autres attributs, les différences entre les propriétés peuvent être déterminées par les différences trouvées pour chaque attribut pertinent, pondérées par l'importance relative de ces attributs. Le format de l'indice de similarité adopté consiste en la relation suivante:

$$SIM_{a,b} = 1 / [(1+k_0*d_{a,b})*(1+k_1*\sum_i (a_i*|X_{ib}-X_{ia}|^{c_i}))] *100\%. \quad (3)$$

Où  $SIM_{a,b}$  est la mesure de la similitude entre la propriété évaluée (a) et le cas de la base de données (b);  $d_{a,b}$  est la distance entre eux (en kilomètres) et la seconde partie est la mesure des différences entre les deux, identifié à travers leurs ensembles d'attributs  $X_{ia}$  et  $X_{ib}$ , respectivement, et pondérés à travers les prix hédoniques  $a_i$ . En outre, des renforts de pondération peuvent être utilisés en utilisant des exposants distincts pour chaque attribut ( $c_i$ ). Deux constantes ( $k_0$ ,  $k_1$ ) ont été utilisées pour étalonner la mesure, en fonction de la dimension des termes.

La similitude maximale est atteinte lorsque les deux composants sont annulés simultanément, c'est-à-dire lorsque le cas de la base est une propriété

ayant des caractéristiques égales et située dans le même bâtiment de l'évaluation. Ayant obtenu l'échantillon de cas similaires (cas pertinents), des compléments peuvent être réalisés avec des données de terrain et un modèle prédictif spécifique est déterminé, générant enfin l'estimation.

L'étape suivante consiste à déterminer les modèles d'évaluation pour le calcul subséquent de la valeur immobilière. L'évaluation traditionnelle est basée sur des modèles de prix hédoniques, avec une fonction de base similaire à celle présentée dans:

$$\text{Valeur} = a_0 + a_1 * \text{Surface} + a_2 * \text{Chambres} + a_3 * \text{Parking} + a_4 * \hat{\text{Age}} + a_5 * \text{Qualité du quartier}. \quad (4)$$

Dans certains cas, la fonction peut avoir un certain degré de non-linéarité. Par exemple, la valeur peut apparaître comme un logarithme de la valeur. Les modèles hédoniques sont estimés par analyse de régression (en général utilisant MCO) et sont évalués avec les paramètres traditionnels d'analyse de l'ajustement des modèles aux données ( $R^2$ , tests t et F, et analyse des résidus), avec le niveau de signification adopté dans 5%. Dans ce cas, la procédure d'ajustement avec un échantillon et de test avec un autre a été adoptée.

Enfin, les règles ont été élaborées en fonction de la taille de l'immobilier, à travers la surface habitable. À partir des données sélectionnées, trois équations de régression ont été obtenues, suivant le format de base de l'équation (4), qui ont été utilisées pour composer un système. La règle de base était la suivante:

$R_i : \text{SI Surface est } A_i \text{ ALORS Valeur} = (a_{i0} + a_{i1} * \text{Surface} + a_{i2} * \text{Chambres} + a_{i3} * \text{Parking} + a_{i4} * \hat{\text{Age}} + a_{i5} * \text{Qualité du quartier}).$	(5)
--	-----

Où les  $a_{ij}$  sont les coefficients  $a_j$  calculés pour chaque règle  $R_i$ . Comme les systèmes de règles floues représentent un ensemble d'équations, le système est testé comme un modèle unique. Pour la comparaison avec le modèle de régression, deux mesures d'erreur ont été utilisées, l'erreur absolue moyenne (MAE, pour *Mean Absolute Error*) et l'erreur quadratique moyenne (MSE, pour *Mean Square Error*), ont été utilisées pour comparer les modèles.

## 5. Estimations et résultats

### 5.1. Zone d'étude et données collectées

Pour démontrer les techniques, nous avons utilisé des données de Porto Alegre, une ville dans le sud du Brésil, qui comprend les données de taxe de vente de l'appartement vendu entre août 1998 et décembre 2017, avec plus de 165 000 cas initiaux. La partie empirique de cette étude a été réalisée en ajustant

plusieurs modèles avec les mêmes données pour permettre la comparaison. Le tableau 1 présente quelques informations sur la base de données.

L'information disponible comprend les variables suivantes : prix de vente (Prix), la taille de la propriété en mètres carrés (Surface), nombres de chambres et des places de parking, la taille du terrain en mètres carrés, la qualité de la construction et l'âge en année (Âge). Les appartements duplex avec terrasse ont été identifiés. Le moment de la vente a été inclus en utilisant une échelle continue de temps, à partir du mois de l'ancienne à vendre dans l'échantillon (Mois de la vente = 1 à août 1998).

Une variable d'emplacement a été considérée en utilisant une mesure fondée sur le revenu du ménage, le motif du logement et de la qualité générale (Qualité du quartier). Le positionnement spatial des biens a été considéré en utilisant les coordonnées mesurées à l'aide d'une grille cartésienne (X, Y). Les distances au centre de la ville et les plus petites distances aux principaux centres commerciaux et écoles ont été calculées, étant mesurés en km. Le prix de vente était à l'origine en monnaie brésilienne (Reais), et a été converti en euros sur la base du taux de change de décembre, 2017. Certaines mesures statistiques sont présentées dans le tableau 1.

**Tableau 1**  
*Mesures statistiques des données recueillies*

Variable	Unité	La base de données entière Moyenne	L'échantillon des 256 cas Moyenne	Propriété sélectionnée Quantités
Prix	Euros	64.706,12	182.783,42	156.250,00
Surface de construction	m <sup>2</sup>	94,70	121,49	159,00
Chambres	-	1,76	2,23	2
Place de parking	-	1,57	1,23	1
Duplex avec terrasse	-	0,06	0,43	1
Qualité de construction	-	2,34	2,46	2
Âge des bâtiment	An	20,90	37,80	39
Mois de la vente	Mois	153,68	172,25	233
Qualité du quartier	-	33,68	52,88	35
Taille du terrain	m <sup>2</sup>	4.846,60	1.864,31	725,00
Distance aux le centre de la ville	km	5,36	1,98	1,41
Distance aux centres commerciaux	km	5,16	1,95	2,42
Distance aux l'écoles	km	2,75	1,72	0,78
X	km	6,847	4,896	4,397
Y	km	13,225	14,041	13,789

Une propriété a été sélectionnée pour tester l'application du modèle. La propriété choisie est un appartement de 159 m<sup>2</sup> de surface, situé à Rue Santo Antônio, 618, dans le quartier Bom Fim (Qualité du quartier= 35). La propriété dispose de deux chambres, une place de parking couverte et est de type duplex avec terrasse. Il a 39 ans et est de moyenne norme (Qualité de construction= 2). Le taille du terrain est 725 m<sup>2</sup>. Ce bâtiment est situé aux coordonnées (4,397, 13,789). La distance au centre de la ville est de 1,41 km, et la distance au commerce est de 2,42 km. La distance aux l'écoles est 0,78 km. La valeur estimée devrait être ajustée pour le mois de décembre 2017 (Mois = 233). Cette bien a été évaluée par une banque à 156 250 Euros (Tableau 1).

### 5.2. Sélection de données

Le formulaire de sélection pour l'évaluation individuelle utilisait RàPC associant les caractéristiques pertinentes de la propriété à une mesure de la distance entre les propriétés. Les coefficients hédoniques ont été calculés en utilisant toute la base de données. En calibrant les coefficients  $k_0 = 0,025$  et  $k_1 = 0,000333$ . La mesure de la similarité des cas de la base à l'évaluation a ensuite été calculée par:

$$SIM_{a,b} = 1 / [(1 + 0,025 * ((X - 4,397)^2 + (Y + 13,789)^2)^{0,5}) * (1 + 0,000333 * (940,276 * |Surface - 159| + 2907,491 * |Chambres - 2| + 7881,243 * |Parking - 1| + 8422,495 * |Duplex - 1| - 56,447 * |Âge - 39| + 58487,842 * |Qualité de construction - 2| + 312,615 * |Qualité du quartier - 35| - 43,290 * |Distance au centre - 1,41| - 90,022 * |Distance au commerce - 2,42| - 16,184 * |Distance aux l'écoles - 0,78| - 19,864 * |Mois - 233|))] * 100\%. \quad (6)$$

La similitude a été calculée pour tous les biens immobiliers dans la base de données. Après avoir ordonné les cas de base pour leur similitude avec celui évalué, les cas avec similitude supérieure à 90% ont été sélectionnés, résultant en 256 cas, formant un échantillon de taille raisonnable pour une utilisation avec des techniques d'intelligence artificielle, d'une part, et un ensemble de grande homogénéité, de l'autre. La description de ces données c'est aussi dans le tableau 1. Les cas résultants ont été divisés en deux fichiers à des fins de formation et de test, contenant respectivement 80% et 20% de données. L'équation 6 a été testée, vérifiant l'exactitude d'échantillon récupéré.

### 5.3. Développement des modèles

Certains modèles de régression traditionnels ont été développés avec les données. En fin de compte, le meilleur modèle a utilisé le logarithme du prix (équation 7). Le modèle généré présentait un degré d'ajustement raisonnable, avec un coefficient de régression de  $R^2 = 0,8336$ , ce qui représente environ 83%

de l'explication des variations de prix, ainsi que les coefficients des modèles conventionnels et une significativité inférieure à 5%, selon les statistiques t calculé. La signification du modèle, utilisant la distribution F, était également inférieure à la limite de 5%. En appliquant les données de propriété à l'étude dans l'équation 7, le prix calculé est 155.993,39 Euros :

$$\begin{aligned} \text{Prix} = & \exp(10,9285 + 0,002765 * \text{Surface} + 0,1687 * \text{Chambres} + 0,1486 * \text{Parking} + \\ & + 0,1512 * \text{Duplex} - 0,007532 * \text{Âge} + 0,1055 * \text{Qualité de construction} + \\ & + 0,0001487 * \text{Qualité du quartier} + 0,0001278 * \text{Mois}). \end{aligned} \quad (7)$$

Le système basé sur des règles diffuses a été développé. La surface du bâtiment a été utilisée comme élément de fuzzification car c'est la variable qui a montré la plus grande importance dans l'analyse statistique (plus grand t). Pour déterminer les règles, les données ont été classées par surface et divisées en ensembles de la même taille. Systèmes avec deux, trois et quatre règles ont été testés, obtenant de meilleures performances avec le système des trois règles.

Ainsi, le système de règles était basé sur trois ensembles flous avec deux extrémités trapézoïdales et un triangle central. Les valeurs de référence étaient de 81,29 m<sup>2</sup> (« petit appartement »), 140,61m<sup>2</sup> (« milieu ») et 278,64 m<sup>2</sup> (« grand ») comme le montre la Fig. 1. La ligne pointillée indique la propriété sélectionnée, qui se situe approximativement au milieu de l'intervalle considéré.

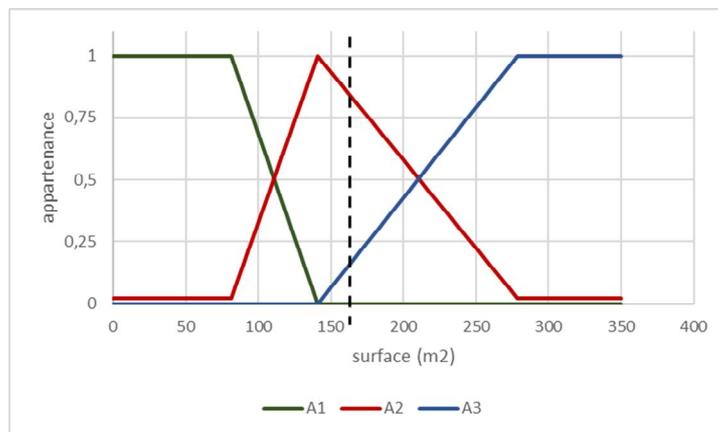


Fig. 1 – Ensemble flous.

Les modèles de régression ont été utilisés pour calculer les fonctions pour la partie conséquente des règles, en utilisant chacun des trois ensembles. Chacun a été examiné selon les critères traditionnels, obtenant  $R_2$  entre 0,84 et 0,86, ainsi que les tests t et F et l'analyse des erreurs a été réalisée

individuellement, avec des résultats satisfaisants. Le système de règles généré était le suivant:

$$R_1: \text{SI } Surface \text{ est } A_1 \text{ ALORS } Prix_1 = \exp(10,7271 + 0,002342 * Surface + 0,1949 * Chambres + 0,1799 * Parking + 0,1821 * Duplex - 0,006923 * \text{Âge} + 0,1121 * \text{Qualité de construction} + 0,0001253 * \text{Qualité du quartier} + 0,0003147 * Mois), \quad (8)$$

$$R_2: \text{SI } Surface \text{ est } A_2 \text{ ALORS } Prix_2 = \exp(10,4256 + 0,003246 * Surface + 0,2182 * Chambres + 0,1676 * Parking + 0,1847 * Duplex - 0,003172 * \text{Âge} + 0,1085 * \text{Qualité de construction} + 0,0001029 * \text{Qualité du quartier} + 0,0005418 * Mois), \quad (9)$$

$$R_3: \text{SI } Surface \text{ est } A_3 \text{ ALORS } Prix_3 = \exp(10,6372 + 0,002613 * Surface + 0,1988 * Chambres + 0,1746 * Parking + 0,1926 * Duplex - 0,004193 * \text{Âge} + 0,1348 * \text{Qualité de construction} + 0,0001132 * \text{Qualité du quartier} + 0,0002798 * Mois). \quad (10)$$

L'application du système se fait par la moyenne des prix calculées par les règles  $R_i$  pondérées par l'appartenance du cas à chaque ensemble flou  $A_i$  (équation 11). Considérant les ensembles flous  $A_i$  (Fig. 1) et la surface de  $159 \text{ m}^2$ , l'appartenances de cette propriété sont :

- $\mu_{A_1} = 0$
- $\mu_{A_2} = 0,868$
- $\mu_{A_3} = 0,132$
- avec  $\mu_A = \mu_{A_1} + \mu_{A_2} + \mu_{A_3} = 1$

En utilisant les informations de propriété, les prix calculés par les règles

$R_i$  sont:

- $R_1 = 144.812,00$
- $R_2 = 155.367,24$
- $R_3 = 161.570,26$

Ainsi, le prix estimé pour la propriété sélectionnée est:

$$\text{Prix} = 0 * 144.812,00 + 0,868 * 155.367,24 + 0,132 * 161.570,26 = 156.185,49. \quad (11)$$

## 6. Analyse des modèles d'évaluation

Le modèle de régression présente les variables conventionnelles, y compris la superficie de construction, le nombre de place de parking, la norme de construction (qualité du bâtiment), l'âge et le mois de vente de la propriété. La distance au centre, au l'écoles et au commerce n'a pas participé au modèle, probablement parce qu'il y a une homogénéité de localisation, c'est-à-dire que la distance est indifférente car elle est similaire pour toutes les données. Le surface du terrain n'a pas été sélectionné non plus.

Les règles qui composent le système des règles floues sont semblables les unes aux autres, et elles sont similaires au modèle hédonique. Après le

développement des modèles plusieurs tests ont été effectués, particulièrement en ce qui concerne les erreurs commises par les modèles. Le tableau 2 présente les résultats pour les modèles.

**Tableau 2**  
*Résultats pour les modèles*

Modèle	Entraînement - 205 cas		Test – 51 cas		Prix estimés pour la propriété sélectionnée (€)
	MAE	MSE	MAE	MSE	
Régression	52,18	11.034	53,72	11.385	155.993,39
SBRF	48,97	10.333	49,01	10.956	156.185,49

Il est vérifié qu'il existe un équilibre entre les techniques explorées, dans tous les paramètres étudiés. Le modèle avec des règles floues fonctionne légèrement mieux. Les Prix estimés sont également très similaires, proches de 156 000 Euros.

## 7. Conclusions

Plusieurs techniques dans le domaine de l'intelligence artificielle ont un potentiel d'utilisation dans les évaluations immobilières. Nous avons présenté une application hybride. La sélection des cas pour les modèles d'évaluation a été effectuée grâce à un mécanisme de RàPC basé sur des mesures de positionnement spatial et de différenciation entre les propriétés de base et le cas analysé (une propriété sélectionnée pour évaluation) en utilisant une formulation spécialement développée.

Le mécanisme de sélection des cas a bien fonctionné, indiquant un groupe de cas bien ajusté. Cette tâche est généralement effectuée par l'évaluateur, mais au prix d'un certain temps. À partir de ces données, certains modèles ont été générés, avec le soin de réserver une partie des données pour tester le modèle.

L'équation de régression présentait de bonnes performances statistiques. Les modèles de régression ont servi aussi de base pour déterminer la fonction de similarité, ainsi que les parties conséquentes des règles flous.

Le modèle des règles floues avait des performances statistiques satisfaisantes. Nous n'avons pas détecté complexité significative dans ce modèle, qui peut alors être considéré comme une alternative dans l'évaluation immobilière.

Enfin, les exigences de base requises pour les modèles d'évaluation, tels que la précision et la facilité de compréhension, ont été suivies de la même

manière par les techniques utilisées. On voit qu'il y a un équilibre dans les niveaux d'erreur.

**Remerciement.** L'auteur souhaite remercier les agences de recherche brésiliennes CAPES, CNPq et FAPERGS et à la mairie de Porto Alegre.

## REFERENCES

- Aamodt A., Plaza E., *Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches*, *AI Communications*, **7**, 1, 39-59 (1994).
- Abidoye R. B., Chan A.P.C., *Artificial Neural Network in Property Valuation: Application Framework and Research Trend*, *Property Management*, **35**, 5, 554-571 (2017).
- Abidoye R.B, Chan A.P.C., *Research Trend of the Application of Artificial Neural Network in Property Valuation*, Proc. of the 33rd CIB W78 Conf. 2016. Brisbane, Australia, 1-9.
- Aguiar M.M., Simões R., Golgher A.B., *Housing Market Analysis Using a Hierarchical–Spatial Approach: the Case of Belo Horizonte, Minas Gerais, Brazil*, *Regional Studies, Regional Science*, **1**, 1, 116-137 (2014).
- Alcantud J., Rambaud S., Torrecillas M.J.M., *Valuation Fuzzy Soft Sets: a Flexible Fuzzy Soft Set Based Decision Making Procedure for the Valuation of Assets*, *Symmetry*, **9**, 11 (2017).
- Alkhatib H., Weitkamp A., Zaddach S., Neumann I., *Realistische Unsicherheits-schätzung des Verkehrs wertes durch ein Fuzzy-Bayes-Vergleichswertverfahren*, *ZFV - Zeitschrift für Geodasie, Geoinformation und Landmanagement*, **141**, 3, 169-178 (2016).
- Amri S., Tularam G.A., *Performance of Multiple Linear Regression and Nonlinear Neural Networks and Fuzzy Logic Techniques in Modelling House Prices*, *Journal of Mathematics and Statistics*, **8**, 4, 419-434 (2012).
- \* \* \* *The Appraisal of Real Estate*, Appraisal Institute, 14<sup>th</sup> ed, Appraisal Institute, Chicago, 2013.
- Bagnoli C., Smith H., *The Theory of Fuzz Logic and its Application to Real Estate Valuation*, *Journal of Real Estate Research*, **16**, 2, 169-200 (1998)
- Braga L.F., Alves T.W., *Valuation of Properties in Rio Grande do Sul: an Analysis from Spatial Regression*, *A Economia em Revista-AERE*, **22** ,1, 85-111 (2015).
- Chen J.-H., Ong C.F., Zhang L., Hsu S.-C., *Forecasting Spatial Dynamics of the Housing Market Using Support Vector Machine*, *Internat. J. of Strategic Property Manag.*, **9179**, 3, 273-283 (2017).
- Cruz B.O., Morais M.P., *Demand for Housing and Urban Services*, Proc. of ENHR 2000 Conf. 26-30 June 2000. Gavle, Sweden, 1-29.
- D'Amato M., *Comparing Rough Set Theory with Multiple Regression Analysis as Automated Valuation Methodologies*, *Internat. Real Estate Rev.*, **10** ,2, 42-65 (2007).

- D'Amato M., Kauko T. (org), *Advances in Automated Valuation Modeling AVM - After the Non-Agency Mortgage Crisis*, Cham, Switzerland: Springer Internat. Publishing AG, 2017.
- Fourati F., *Modélisation et commande intelligentes de systèmes dynamiques*, Saint-Denis, Editions Publibook, 2017.
- Furtado B A, *Neighbourhoods in Urban Economics: Incorporating Cognitively Perceived Urban Space in Economic Models*, *Urban Studies*, **48**, 13, 2827-2847 (2011).
- Gasparini L., Escudero W.S., *Implicit Rents from Own-Housing And Income Distribution: Econometric Estimates for Greater Buenos Aires*, *J. of Income Distribution*, **12**, 1-2, 32-55 (2003).
- Gerek I.H., *House Selling Price Assessment Using Two Different Adaptive Neuro-Fuzzy Techniques*, *Automation in Construction*, **41**, 33-39 (2014).
- González M.A.S., Formoso C.T., *Mass Appraisal with Genetic Fuzzy Rule-Based Systems*, *Property Management*, **24**, 1, 20-30 (2006).
- González M.A.S, *Aplicação de técnicas de descobrimento de conhecimento em bases de dados e de inteligência artificial em avaliação de imóveis*, Thèse de doctorat, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brésil, 2002.
- Graczyk M., Lasota T., Trawiński B., *Comparative Analysis of Premises Valuation Models Using KEEL, RapidMiner, and WEKA*, in Nguyen N.T., Kowalczyk R., Chen S.M. (eds), *Computational Collective Intelligence. Semantic Web, Social Networks and Multiagent Systems*, ICCCI 2009, Lecture Notes in Computer Science, **5796**, Berlin, Heidelberg: Springer, 800-812 (2009).
- Guan J., Zurada J., Levitan A.S.A., *Comparison of Regression and Artificial Intelligence Methods in a Mass Appraisal Context*, *Journal of Real Estate Research*, **3**, 350-387 (2011).
- Hefferan M J., Boyd T., *Property Taxation and Mass Appraisal Valuations in Australia – Adapting to a New Environment*, *Property Management*, **28**, 3, 149-162 (2010).
- Horváth Á., Imre B., Sápi Z., *The International Practice of Statistical Property Valuation Methods and the Possibilities of introducing Automated Valuation Models in Hungary*, *Financial and Economic Review*, **15**, 4, 45-64 (2016).
- Kauko T., D'Amato M. (Org.), *Mass Appraisal Methods. An International Perspective for Property Valuers*, Oxford, Blackwell, 2008.
- Kauko T., *From Modelling Tools Towards the Market itself an Opportunity for Sustainability Assessment?*, *Internat. J. of Strategic Property Manag.*, **12**, 2, 95-107 (2008).
- Kauko T., *Towards the 4th Generation - an Essay on Innovations in Residential Property Value Modelling Expertise*, *J. of Property Research*, **21**, 1, 75-97 (2004).
- Kolodner J., *Case-Based Reasoning*, San Francisco, Morgan Kaufmann, 2014.
- Król D., Lasota T., Nalep W., Trawiński B., *Fuzzy System Model to Assist with Real Estate Appraisals*, in Okuno H.G, Ali M. (eds), *New Trends in Applied Artificial Intelligence*, IEA/AIE 2007. Lecture Notes in Computer Science, **4570**, Berlin, Heidelberg: Springer (2007a).

- Król D., Lasota T., Trawiński B., Trawiński K., *Comparison of Mamdani and TSK Fuzzy Models for Real Estate Appraisal*, in Apolloni B, Howlett R J (eds), *Knowledge-Based Intelligent Information and Engineering Systems. KES 2007. Lecture Notes in Computer Science*, **4694**, Berlin, Heidelberg: Springer, 1008–1015 (2007b).
- Król D., Lasota T., Trawiński B., Trawiński K., *Investigation of Evolutionary Optimization Methods of TSK Fuzzy Model for Real Estate Appraisal*, *Internat. J. of Hybrid Intelligent Syst.*, **5**, 3, 111-128 (2008).
- Lasota T., Mazurkiewicz J., *Comparison of Data Driven Models for the Valuation of Residential Premises Using KEEL*, *Internat. Journal of Hybrid Intelligent Systems*, **7**, 1, 3-16 (2010).
- Lasota T., Pronobis E., Trawiński B., Trawiński K., *Exploration of Soft Computing Models for the Valuation of Residential Premises Using the KEEL Tool*, *Proc. of the First Asian Conf. on Intelligent Information and Database Systems*, 2009. ACIIDS 2009. Dong Hoi, Vietnam: IEEE, 253-258.
- Lasota T., Telec Z., Trawiński G., Trawiński B., *Empirical Comparison of Resampling Methods Using Genetic Fuzzy Systems for a Regression Problem*, in Yin H., Wang W., Rayward-Smith V. (eds), *Intelligent Data Engineering and Automated Learning*, IDEAL 2011. Lecture Notes in Computer Science, **6936**, Berlin, Heidelberg: Springer, 17-24 (2011).
- Lasota T., Trawinski B., Trawinski K., *An Attempt to Use the KEEL Tool to Evaluate Fuzzy Models for Real Estate Appraisal*, in Zgrzywa A., Choroś K., Siemiński A. (eds), *New Trends in Multimedia and Network Information Systems*. IOS Press, 125-139 (2008).
- Lasota T., Trawiński B., Trawiński K., *Evolutionary Optimization of TSK Fuzzy Model to Assist with Real Estate Appraisals*, in Rutkowski L., Tadeusiewicz R., Zadeh L.A., Zurada J. (eds), *Computational Intelligence: Methods and Applications*, Academic Publishing Housing EXIT, Warszawa, 232-243 (2008).
- Leake D.B., *Case-Based Reasoning*, New York, John Wiley and Sons, 2003.
- Lughofer E., Trawinski B., Trawinski K., Kempa O., Lasota T., *On Employing Fuzzy Modeling Algorithms for the Valuation of Residential Premises*, *Information Sciences*, **181**, 23, 5123–5142 (2011).
- Melanda E., Hunter A., Barry M., *Identification of Locational Influence on Real Property Values Using Data Mining Methods*, *European J. of Geography, Cybergeo* (2016).
- Muggenhuber G., Mansberger R., Navratil G., Twaroch C., Wessely R., *Kataster als Ausgangspunkt einer flächendeckenden Liegenschaftsbewertung*, *Wirtschaft und Gesellschaft*, **39** (2013).
- Musa A.G., Daramola O., Owoloko E.A., Olugbara O.O., *A Neural-CBR System for Real Property Valuation*, *J. of Emerging Trends in Computing and Information Sci.*, **4**, 8, 611-622 (2013).
- O’Roarty B., Patterson D., McGreal S., Adair A., *A Case-Based Reasoning Approach to the Selection of Comparable Evidence for Retail Rent Determination*, *Expert Syst. with Appl.*, **12**, 4, 417-428 (1997).

- Pacharavanich P., Wongpinunwatana N., Rossini P., *The Development of a Case-Based Reasoning System as a Tool for Residential Valuation in Bangkok*, Proc. of the Pacific-Rim Real Estate Society (PRRES) Conf., 2000.
- Rezayan H., Karimipour F., Amini J., Toose A.M., *Neighborhoods Development in Real Estate Appraisal Process*, Proc. of the FIG Working Week 2005 and GSDI-8 Cairo - From Pharaohs to Geoinformatics, 1-11.
- Riesbeck C.K., Schank R.C., *Inside Case-Based Reasoning*, Psychology Press, 2013.
- Rosen S., *Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition*, J. of Political Economy, **82**, 34-55 (1974).
- Sheppard S., *Hedonic Analysis of Housing Markets*, in Cheshire P.C., Mills E.S. (eds), *Handbook of Applied Urban Economics*, **3**, Chap.8, Elsevier, Amsterdam, 1999.
- Taffese W.Z., *Case-Based Reasoning and Neural Networks for Real Estate Valuation*, Artificial Intelligence and Applications, 98-104 (2007).
- Telec Z., Trawiński B., Lasota T., Trawiński K., *Comparison of Evolving Fuzzy Systems with an Ensemble Approach to Predict from a Data Stream*, in Bădică C., Nguyen N.T., Brezovan M. (eds), *Computational Collective Intelligence. Technologies and Applications*, ICCCI 2013. Lecture Notes in Computer Science, **8083**, Springer, Berlin, Heidelberg, 377-387 (2013).
- Tomić H., *Analiza geoprostornih podataka za potrebe vrednovanja nekretnina u urbanim područjima (Geospatial Data Analysis in Purpose of Real Estate Valuation in Urban Areas)*, Geodetski fakultet. University of Zagreb. Faculty of Geodesy (2010).
- Vada V.M., *Mass appraisal. Esperienze di valutazione immobiliare di massa*, Politecnico de Mila Scuola di Ingegneria Edile-Architettura, 2014.
- Watson I., *Applying Case-Based Reasoning: Techniques for Enterprise Systems*, San Francisco, Morgan Kaufmann, 1998.
- Xie J.C., Wu B., Sriboonchita S., *Fuzzy Estimation Methods and their Application in Real Estate Evaluation*, Internat. J. of Intelligent Technol. a. Appl. Statistics, **3**, 2, 187-202 (2010).
- Yakubovskiy V., Dimitrov G.P., Bychkov O., Shevchenko T., Panayotova G., *Combined Neural Network Model for Real Estate Market Range Value Estimation*, Proc. of The Fourth Internat. Conf. on Artificial Intell. A. Pattern Recognition (AIPR2017), Lodz, Poland: SDIWC, 1-16.
- Zurada J., Levitan A., Guan J., *A Comparison of Regression and Artificial Intelligence Methods in a Mass Appraisal Context*, J. of Real Estate Research, **33**, 3, 349-387 (2011).

#### STUDIU PRIVIND UTILIZAREA ANALIZEI DE CAZ ȘI A SISTEMELOR DE TIP FUZZY-LOGIC ÎN EVALUAREA DE PREȚ A BUNURILOR IMOBILIARE

(Rezumat)

Bunurile imobiliare sunt bunuri cu valori diferite și trebuie să se țină cont de caracteristicile principale ale acestora în evaluarea imobiliară. Efectele simultane a mai multor factori / caracteristici pot provoca dificultăți atunci când se dorește o comparație

directă a două sau mai multe bunuri imobiliare. În acest caz este necesară dezvoltarea unui model numeric pentru evaluarea corectă a acestor diferențe. Evaluarea imobiliară poate fi îmbunătățită prin utilizarea unui eșantion omogen selecționat dintr-o bază de date care conține cazuri cu un grad foarte mare de asemănare cu imobilul supus evaluării. Studiul de față propune un model hybrid ce combină analiza de caz cu sistemele de reglaj de tip fuzzy-logic și regresie, cu aplicare directă pe piața bunurilor imobiliare din Porto Alegre, un oraș situat în Brazilia. Rezultatele obținute demonstrează fezabilitatea modelului propus.