ISSN 1644-0757
www.acta.media.pl

ACTA SCIENTIARUM POLONORUM *Oeconomia 14 (2) 2015, 27–38*

MULTIFACETED ANALYSIS OF THE PROCESS OF PRICE DEVELOPMENTS ON THE LOCAL REAL ESTATE MARKET BY MEANS OF THE REGRESSION TREES METHOD (C&RT)

Mariusz Dacko, Katarzyna Szajdecka
University of Agriculture in Krakow

Abstract. Attention of rational purchaser or seller of real estate is focused on the prices and properties of similar real estates. In the light of the theory of valuation, differentiation in prices on the local market is explained by variation in a few most important properties of marketed real estate. Thus, in the practice of valuation, a important issue is to decide which of many possible attributes are significant and what is their individual contribution to the market price that can be achieved. In this paper, this problem has been solved using the model of regression trees (C&RT). The study covered undeveloped land designed in local plans for single-family housing development, situated on the northern outskirts of Krakow. A regression tree permitted a simple assessment of the impact of the different attributes on real estate prices. This method enables prediction of the market value, but it requires knowledge of at least several tens, or preferably many hundreds, of transactions relating to similar real estates. For that reason, it is predestined for mass valuations.

Key words: real estate market, regression trees C&RT

INTRODUCTION

Initiated in 1989, radical changes in the functioning of the Polish economy connected with the shift from the centrally planned economy to the market system, gave rise to a dynamic development of the Polish real estate market. Already in 1990, over 220 thousand market transactions were entered into in the Polish real estate sector. Compared to 1980, their number grew by 37%, which confirmed that in the new economic reality the Polish real estate market entered a phase of a spectacular boom. Over the following five

Corresponding author: Mariusz Dacko, University of Agriculture in Krakow, Economic and Social Institute, Adama Mickiewicza 21, 30-120 Kraków, Poland, e-mail: m.dacko@ur.krakow.pl
© Copyright by Warsaw University of Life Sciences Press, Warsaw 2015

years, turnover increased twice: in 1995 nearly 446 thousand of real estate purchase/sales transactions were entered into in Poland.

In 1999, the reform changing Poland's administrative division introduced a three--level structure of territorial division. With that, new actors appeared on the local real estate markets. These were local governments, which had legal personality and own real estate resources. These actors (in particular nearly 2.5 thousand gminas) became important participants of the real estate market, not only in the sphere of building land supply. In the institutional sphere, they were also entrusted with a range of important functions in the area of real estate management. For local governments, the question about the price and market value became one of the most important questions connected with real estate – especially that it appeared not only in the situation of purchase/sale, but also during calculating betterment taxes and planning fees, granting the right of perpetual *usufruct* and establishing permanent management. Valuation was also required in new expropriation procedures, as the property right could be deprived only against just compensation (monetary or in the form of replacement real estate). According to the provisions of the real estate management act, the amount of such compensation (equal to the market value of the deprived property right) should be clearly established at the stage of an administrative trial and paid to the owner within 14 days from the validation of the decision about expropriation[1].

Summing up, it should be highlighted that every time rights to fixed property are transferred or granted, knowledge of their value is a fundamental issue. It is also needed when immovable property is divided, contributed to enterprises being created or used to secure a bank's claim.

The question about real estate value may take on an unprecedented scale in our country with the tax system reform. The possibility of replacing the existing real property tax with an *ad valorem* ("according to value") tax has been speculated about for over 20 years. The methodical, technical and organisational aspects of its implementation were tested in pilot projects conducted, among other things, in Kraków, Łódź, Pabianice, Szczecin, Olsztyn and Ostróda [Kuryj 2001]. However, this huge project raises serious concerns and leads to social resistance. Making the amount of this tax dependent on the value of real estate would certainly significantly influence prices and sales on the local markets until new levels of market prices were established (and the first transactions taking them into account). Many owners of real estate situated in centres and other attractive spheres of large cities (with high *ad valorem* value) would be subject to many times higher tax assessment than under the existing system that took into account square footage. We can thus conclude that the challenge faced by public administration would be to ensure in the first years after introducing the *ad valorem* tax that its rates are socially acceptable. These rates would be based not on the physical size of the object of taxation but on its value, established on a mass scale by means of mathematical models comprehensible to the taxpayer.

Nowadays, from 208 thousand to 324 thousand sales contracts are concluded annually on the Polish real estate market. In 2012, their total value was PLN 710 billion, and in land property transactions alone – PLN 184 billion. So, there is allocation of huge

[1] Real estate management act of 21 August 1997 (Journal of Laws 1997 issue no 115 item 741).

amount of capital on the real estate market. And although the annual real estate transactions on a national scale are not at such high level as in mid-1990s, in 2012 their value was 2.5 higher than state revenue [Obrót nieruchomościami w 2012 r., 2013].

The development of the real estate market is always accompanied by growing demand for reliable knowledge of value. Reliable information about the value of a house, flat or plot even became a value in itself. It makes it easier to take rational decisions not only in situations of selling real estate, but also when taking investment decisions. Nowadays, value management, or active shaping or influencing value, refer not only to commercial facilities. Compared to other goods, real estate is very costly, so it's worth making sure that capital is put in appropriate real estate and is not subject to depreciation. It should be stressed that economic consequences of the lack of comprehensive knowledge about value and its determinants can be very severe for market participants, as they are adequate to the money paid for real estate. So, what do real estate prices on the local market depend on? What factors and attributes establish these prices at a given level? These questions are asked every day by millions of potential purchasers and sellers of houses, flats and plots. These questions bother local governments, developers and banks.

According to the economic theory, price is the result of the market play between supply and demand forces. They, in turn, are determined by factors of the market environment [Wrzosek 1994]. This environment includes legal, economic, social and spatial factors. We can say that these factors shape a certain level of price on the local market. But it is individual properties of real estate, called attributes, that are responsible for its differentiation [Cellmer 1999, Dacko 2002]. Economic assessment of the impact of the different attributes on market price developments is not a simple task. Every real estate has a range of attributes, but usually only some of them significantly impact the price [Mączyńska et al. 2004]. Some of them may also be strongly interrelated. For instance, the neighbourhood of a plot may be strongly related with the existence of a network of utility infrastructure and a paved road. Building plots situated far from larger concentration of such plots are less often equipped with such costly facilities. Greenery and quiet location are valuable attributes of every real estate, but statistically they less often co-occur with the location in the strict centre, public transport accessibility and close location of social infrastructure facilities. It should also be stressed that the contribution of the different attributes to price differentiation varies across local markets. Thus, we cannot automatically use the results of an analysis of one local market to draw conclusion about prices on another local market. The traditional way to determine the impact of a particular attribute on the market value of real estate has been to compare the prices of real estates that extremely differ in an attribute, *ceteris paribus* [Mączyńska et al. 2004]. Other popular tools for such evaluation include simple and multiple regression models [Cellmer 1999, Dacko 2002]. However, with the development of calculation methods and techniques in the field of real estate market analyses, we can see an increasingly stronger position of methods classified as data mining (machine learning, neural networks, decision trees). Indeed, these methods start to be used by banks, statistical services and local governments. In situations when the question about value is on a mass scale, these methods allows for minimization of the unit workload in appraisal of real estate. This paper examines the process of price developments on the local real

estate market in three places in the Gmina Zielonki using the model of a C&RT binary regression trees. The study covered the market of land property intended in local plans for single-family housing development. The study was exploratory in nature, and its aim was to evaluate results offered in mass appraisal by the regression trees C&RT method.

MATERIAL AND METHOD

The appraisal process is in fact a process of predicting the price that can be obtained on the local market depending on the attributes of a specific real estate. In this process we refer to historical events – past transactions relating to similar real estates. Thus, in modelling, apart from attributes, also the dates when the transactions were concluded are taken into account, as prices may change over time. For the purpose of this paper, market purchase/sale transactions between January 2008 and July 2013 in three places of the Gmina Zielonki (Zielonki, Bibice and Węgrzce), near Kraków, were examined. These places are considered to be dormitory towns of Kraków – many residents of this city look for plots to build houses. All transactions referred to undeveloped real estate intended in local plans for single-family housing development. When creating a database, the following transactions were excluded: transactions conducted by a tendering procedure and those concerning sale of plots intended for improvement of the conditions of the neighbouring real estate development. Plots which, due to their large size, were bought with the intention of divisions, building small housing estates and reselling them, were also not taken into account. The final database contained 109 transactions. After analysing notarial acts, maps and satellite imagery, the collection of prices was supplemented with data about square footage, utility infrastructure, shape and localisation of plots, date of entering into transaction, access and neighbourhood, and about potential encumbrances (mortgage, easement of passage or laying utilities) and status of the parties to transactions (Table 1).

Literature [Bitner 2007] recommends reduction of the number of variables in market analysis by grouping attributes. According to this principle, utility infrastructure can be treated as one attribute that has a few statuses (e.g. full, partial, lack). However, in this case, utility infrastructure was the most important price influencing factor. Already the preliminary analyses showed that treating utility infrastructure as one attribute with a few statuses resulted in the loss of information, which had a significant impact on reducing predictive capabilities of tree models. Therefore, each utility is analysed separately. Better results were obtained when the plot square footage was expressed in absolute numbers than when it was divided into various classes (e.g. small, medium-sized, large). Therefore this variable was expressed in absolute numbers. The other attributes were coded in accordance with their language description. Contrary to multiple regression models, regression trees are not subject to restrictions in the form of the requirement to code data by means of numbers. This is undoubtedly a significant convenience when analysing the real estate market. After all most attributes have qualitative character here (e.g. good location, convenient access, non-burdensome neighbourhood or regular shape of the plot).

Table 1. Variables selected for analysis

Variable	Variable name	Variable type	Variable variant
X_1	Plot square footage	Quantitative	number of square metres
X_2	Transaction date	Quantitative	yyyy-mm-dd
X_3	Seller	Quantitative	natural person, legal person
X_4	Buyer	Quantitative	natural person, legal person
X_5	Plot shape	Quantitative	favourable, unfavourable
X_6	Neighbourhood	Quantitative	favourable, unfavourable
X_7	Electrical connection (E)	Quantitative	exists, lack
X_8	Water supply connection (W)	Quantitative	exists, lack
X_9	Sewage connection (S)	Quantitative	exists, lack
X_{10}	Gas supply connection (G)	Quantitative	exists, lack
X_{11}	Encumbrances	Quantitative	exist, lack
X_{12}	Access	Quantitative	favourable, unfavourable
X_{13}	Location	Quantitative	Zielonki, Bibice, Węgrzce

Source: Own study.

Regression trees are widely used to solve data mining problems. They help to identify patterns and relationships hidden in data and thus extract knowledge from it. They work well in the case of data sets that even aren't complete. A C&RT algorithm-based tree searches for a set of logical "if-then" split conditions [Łapczyński 2010].

C&RT trees have numerous advantages compared to other methods. It is mainly simplicity of results: a tree model permits a quick classification of new cases and explanation of rules for this classification [Dacko 2007]. Moreover, C&RT trees represent a non-parametric and non-linear method, therefore it is not necessary to make preliminary assumptions regarding relationships between predictors and a dependent variable. This method works well with data mining problems, where the *a priori* knowledge about the existence of relationships between variables is often only intuitive [http://www.statsoft.pl/ textbook/stathome_stat.html]. Jasińska [2012] also appreciates no necessity to standardise explanatory variables and resistance of C&RT algorithm to predictors that do not affect the dependent variable. The authors of C&RT algorithm – Breiman et al. [1998] – stress that there is no need to convert quantitative predictors into dummy variables. Such conversion is recommended only when a qualitative variable has more than 15 variants.

Numerous advantages confirm attractiveness of regression trees as a tool used in market analysis with reference to neural networks or regression models. But trees, like other models, apart from expert knowledge of the market, require that the researchers understand the nature of modelling phenomena in order to properly select and prepare data for analysis [Dacko 2007].

The process of constructing a regression tree is about examining possibilities of identifying dichotomic sub-sets (nodes) which differ in the average value of the dependent variable. Splitting of each node consists in distinguishing two sub-sets: positive cases, which pass the test ascribed to the node, and negative cases, which do not

pass this test. For each node, variance is estimated, which is its impurity measure. The aim is for average values in nodes to have the smallest possible variance

$$s^2(P) = \frac{1}{P}\sum_{k \in P}\left(y_k - \overline{y}_p\right)^2$$

where: P – a set of cases referred to a daughter node;
$\overline{y}_p$ – average value of the target function of the cases in set P.

As indicated by Dudek [2014], the division of a set of observations at a parent node into two sub-sets is justified only when it leads to decrease in variance in accordance with the following formula

$$p_1\, s^2(P_1) + p_0\, s^2(P_0) < s^2(P_n)$$

where: P_n, P_1, P_0 – respectively sets of cases at a parent node, at a child node which is reached by positive cases, and at a child node which is reached by negative cases;
p_1, p_0 – fractions of positive and negative cases in set P_n.

The splitting criterion (i.e. variable and its value) is selected using the so-called greedy algorithm, which in each node analyses all predictors and all possibilities of using them to divide a set of observations into two subdets [Dudek 2014]. This algorithm is resistant to irrelevant predictors – loosely related or unrelated to the response variable. Such predictors do not participate in constructing a tree. Each split is justified only when it leads to decrease in variance at child nodes in accordance with the above presented formula.

Thus, in the process of learning, a number of different predictors can be presented to trees to obtain valuable information regarding their importance. However, real estates are specific goods, characterised by great variety. One cannot find two identical real estates. Therefore, prices of the rights to real estate are not easy to model [Dacko 2002]. So, a few questions arose in the context of the research subject and method: will C&RT algorithm be able to classify real estate with similar unit prices to specific tree nodes? Will the selected predictors have a sufficiently strong relationship with market prices for tree classifications to have a practical value? Can the tree method be recommended as useful for mass appraisals? The answer to the last question seems to be positive because of simplicity and transparency of results alone. If predictors are related with the dependent variable strongly enough (this relationship may even be non-monotonic), a tree diagram can explain the nature of the differentiation of the variable being analysed and quickly classify its new cases.

The tree-building process requires defining parameters that will allow for evaluation of the model quality and prevention of its excessive growth. In this study, the procedure for creating new nodes was to prune on variance and criterion of minimum number of observations in a node ($n \geq 10$). Quality control of obtained results was conducted by means of V-fold cross-validation ($V = 10$). This test consisted in a random selection of V sub-samples from the set of possessed observations. This method generates a whole sequence of trees from which the best tree is selected. A tree of the specified size is computed V

times each time leaving out one of the sub-samples from the computations and using that sub-sample as a test sample for cross-validation. Each sub-sample is used *V*-1 times in the learning sample and once as the test sample (http://www.statsoft.pl/ textbook/stathome_stat.html). The cross-validation costs (CV) are calculated as an averaged cost from V test samples, and this average is estimate of the CV cost. If the growth of a tree is not accompanied by a satisfactory growth of its predictive capabilities, the splitting process is stopped. In a special case, if none of the predictors contributed to explaining the dependent variable, no splitting would be justified. Then, the resubstitution and cross-validation costs would assume maximum and similar values matching the variance characterising the whole set of observations. If, however, predictors are well selected, the CV and re-substitution costs show desirable fall – which is similar at the beginning. As the trees grow, the difference between the costs becomes sharper, but in the final choice of the right tree the most important role is played by the CV costs. It is their decrease that shows growing capability of recognising regularities hidden in data, and their proper generalisation, which leads to increased predictive accuracy in the test sample. The compromise between complexity and accuracy is ensured by applying the principle of one standard error, suggested by the authors of C&RT algorithm: the "right-sized" tree is the smallest-sized tree whose cross-validation costs are not higher than the lowest (in the whole sequence of trees) costs of cross-validation increased by the value of one standard error of such costs.

RESULTS

The calculation approach used in the regression trees method (C&RT) makes it unaffected by the problem of excessive correlation of explanatory variables or the problem of loose relationship between some of them and the response variable. Moreover, as was already mentioned, trees cope with non-monotonic relationships, and even tolerate missing data. The latter does not have to be specially coded. Possibility of introducing language descriptions makes it much easier to analyse real estate. Such a description was used in this study in the case of as many as 11 predictors. The model was built using the data mining module from Statistica 10 suite. The software generated a sequence of 20 trees (Fig. 1) characterised by different degrees of detail and shapes whose measures were variances estimated for the learning sample (cost of re-substitution) and test sample (cost of cross-validation). The division into the learning sample and test sample was performed by means of V-fold cross-validation, which has been explained in the section describing the research method.

Tree 16, composed of five terminal nodes and four deep splits, was selected for further examination. Cross-validation costs estimated for this tree complied with the rule of one standard deviation:

$$RE^{CV} \leq \min\left(RE^{CV}\right) + \sigma_{\min\left(RE^{CV}\right)}$$

where: RE^{CV} – cross-validation costs for a given tree;
$\min(RE^{CV})$ – minimum costs of cross-validation;
$\sigma_{\min\left(RE^{CV}\right)}$ – standard deviation of minimum costs of cross-validation.

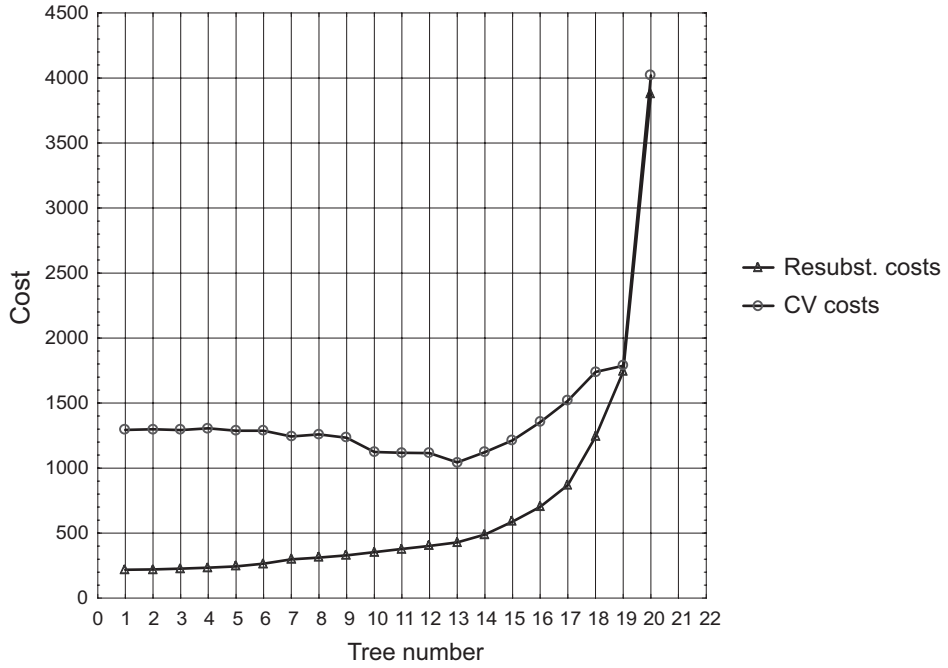

Fig. 1. Chart of costs for a sequence of C&RT regression trees
Source: Own study.

The tree suggested by Statistica was characterised by CV costs equal to 1353. 15 other trees also fulfilled the rule of one standard deviation, but according to the main principle of modelling simplicity implemented in the tree method, the model with the smallest number of terminal nodes should be chosen. A moderately complicated tree should ensure desired interpretation and predictive capabilities. The latter were positively evaluated based on the correlation coefficient between observed and predicted values of the model (w_k = 0.94). Diagram of the tree suggested by Statistica software has been presented in Figure 2.

The tree diagram showed that the process of price developments on the market analysed had a multifaceted character and was determined by availability of technical infrastructure elements. The first split (into branches 2 and 3) took place based on the variable "water supply connection". The plots which due to their location near water pipe network could be equipped with water supply connection achieved an average unit price of PLN 247 per m^2 (87 observations). The average price of plots without access to this connection (22 observations) was PLN 131 per m^2. The criterion for splitting node 2 (into branches 4 and 5) was gas supply connection. Its availability significantly differentiated average unit prices of plots with access to water pipe (207 *vs* PLN 262 per m^2). Another split, into branches 6 and 7, took into account the fact that in transactions relating to plots partially equipped with utilities where the selling party was a natural person average unit prices reached a lower level (PLN 258 per m^2). The last split (into branches 8 and 9) was based on the criterion of the presence of a sewage system. Real estate with access to a sewage

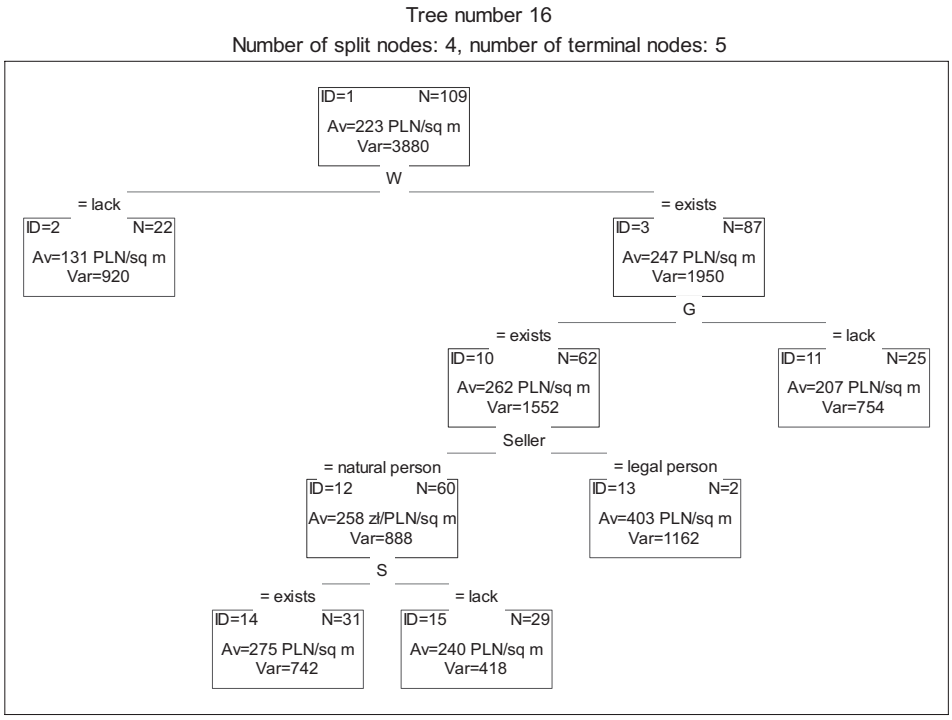

Fig. 2. Model of the final regression tree

Source: Own study.

system achieved an average price of PLN 275 per m^2. The average price of real estate without access to this utility was lower by PLN 35 per m^2.

The model discussed herein was characterised by clear asymmetry: the left branch of the tree was subject to further more detailed splits, whereas the right branch remained the terminal node. This was caused by the fact that a water supply system was the key utility that differentiated prices of building plots. If it was not available, other factors were not important – such plots were sold at a nearly two-fold lower price. Access to a water supply system determined further splits of the left branch of the tree. Then other connections became important. It was even important who was selling the plot.

The tree model is in fact a set of conditional statements the fulfilment of which led to one of five terminal nodes. The average characterising the different nodes was the basis for predicting prices. For instance, node number 9 was: if plots had water supply and gas supply connections and the selling party was a natural person, then if sewage connection was not available, their average unit price was PLN 240 per m^2.

Another advantage of regression trees is possibility of evaluating the importance of predictors. One of Statistica reports may be an ordered ranking of predictors (Fig. 3).

This ranking arranges predictors in terms of their importance, but it doesn't show whether the impact of a specific attribute on the price of real estate was directly proportional or inversely proportional. It should be reminded that in the trees method this impact

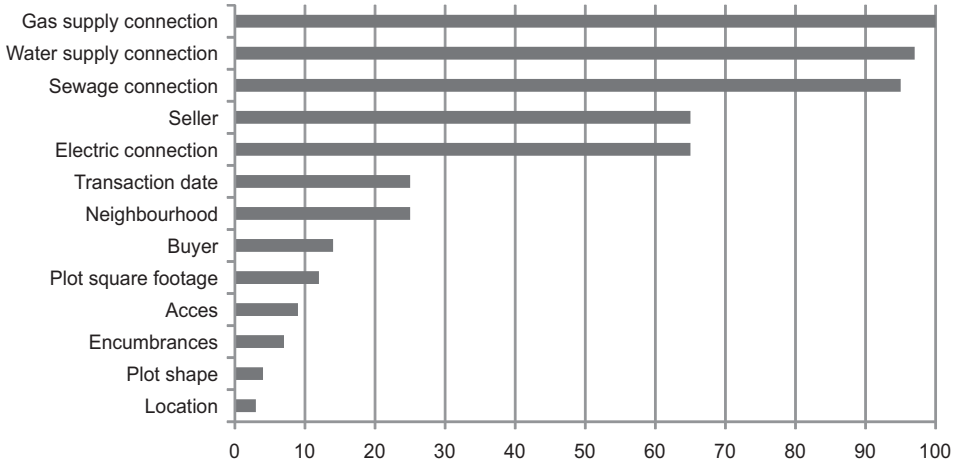

Fig. 3. Ranking of the importance of predictors

Source: Own study.

is often non-monotonic, which is reflected in a series of splits based on different values or classes of the same predictor. The character of relationships can be inferred only based on examination of the diagram of the final tree and only for those predictors that participated in its construction. The importance of predictors may not completely correspond with the final tree diagram. It is determined with reference to the whole sequence of trees of various complexity and all possible variants of their splits. A given predictor may achieve high place in the ranking of importance despite belonging to the main splitting criteria for the final tree. This explains why gas supply connection was the most important predictor in the ranking presented, although the first split of the final tree was performed based on the water supply connection variable. This difference results from the method for calculating the importance of predictors used in the data mining module of Statistica suite [Paluch and Sroka 2012].

The ranking presented confirms that utilities were of key importance in the process of average unit price developments in undeveloped plots intended for single-family house development (Fig. 3). Its elements, i.e. water supply, gas supply and sewage connections, were both used in the construction of the final tree and were at the top of the ranking of predictors. The party selling real estate and electrical connection were also important. These factors to the largest extent determined the process of the developments of unit prices of real estate in the areas analysed. The other attributes contributed much less, in particular: encumbrances on a plot, plot location and shape, and access to a plot.

The trees have not shown importance of the transaction date in the process of price developments, although the temporal scope of the market research was wide and covered the period of almost six years. In the ranking of the importance of market price predictors, the factor of lapse of time (here transaction date) came six and was overtaken by utilities and the selling party. Moreover, this factor has been left out in the structure of the final tree.

The theory of real estate valuation stresses the key importance of location. This fact is usually confirmed in practice. However, location has not been taken into account in the

final tree model. In the ranking of predictors, this important attribute came last, which could suggest glaring contradiction between the results obtained and the theory of valuation. According to the authors, regression trees method (C&RT) would certainly assign high position to this attribute, if only it resulted from differentiation of average unit prices in the areas analysed. As this was not the case, we should suppose that for a potential purchaser location in each of the areas under analysis was equally attractive. Being situated at similar distance from Kraków, they indeed represented one local market.

CONCLUSIONS

The regression tree presented in this paper explained the process of real estate price developments within three areas of gmina Zielonki near Kraków. The model approach made it possible to examine the impact of multiple attributes at the same time on price developments in building plots. The trees method indicated utilities on the land as the main price influencing attribute. This very crucial condition of each building process turned out to be much more important than such factors as neighbourhood, access to a plot, its square footage, shape, or encumbrances.

The basic principle of modelling is simplification. That's why no prediction of any economic phenomena or processes will ever be ideal. Models only allow us to come closer to identifying key factors in the development of a given phenomenon. Their accuracy depends not only on using appropriate methods but also on the appropriate selection of predictors. Regression trees seem to make this task much more easier. They tolerate language descriptions, and even missing data. The results are resistant to the presence of predictors that do not have impact on a response variable. Despite their simplicity C&RT regression trees can reflect non-monotonic relationships. One of the advantages of their use is easy interpretation of the result. It is usually sufficient to analyse a few decision rules presented in a clear graph to identify the general relationship hidden in even huge amount of data.

The model presented permitted achievement of two research goals: predictive one (i.e. predictions of average unit prices of real estate by assigning them to terminal nodes) and descriptive one (i.e. relationships described by split nodes identified in the collection of market data).

Models of regression trees are advanced tools for learning about the complex social and economic reality, including the identification of regularities and tendencies on local real estate markets. Knowledge acquired by means of such models may be valuable to investors, banks and local authorities responsible for managing the property of gmina. It is worth considering their application in the area of mass appraisal.

REFERENCES

Bitner, A. (2007). Konstrukcja modelu regresji wielorakiej przy wycenie nieruchomości. Acta Sci. Pol., Administratio Locorum, 6(4), 59–66.

Breiman, L. et al. (1998). Classification and Regression Trees. Chapman and Hall/CRC, New York.

Cellmer, R. (1999). Metodyka analizy wybranych elementów rynku nieruchomości na potrzeby gospodarki przestrzennej. Rozprawa doktorska, UWM, Olsztyn.
Dacko, M. 2007. Badanie rynku mieszkań w Olsztynie z zastosowaniem metod statystycznych. Studia i Materiały Towarzystwa Naukowego Nieruchomości, 15, 3–4.
Dacko, M. (2002). Analiza czynników kształtujących wybrane segmenty olsztyńskiego rynku nieruchomości. Rozprawa doktorska, UWM, Olsztyn.
Dudek, G. (2014). Prognozowanie krótkoterminowe obciążeń systemów elektroenergetycznych z wykorzystaniem rozmytych drzew regresyjnych. Przegląd Elektrotechniczny, 90, 4, 108–111.
Jasińska, E. (2012). Wybrane metody statystyczne w analizie rynku nieruchomości. Wydawnictwo AGH, Kraków.
Kuryj, J. (2001). Metody doboru i oceny danych w taksacji powszechnej nieruchomości. Rozprawa doktorska, UWM, Olsztyn.
Łapczyński, M. (2010). Drzewa klasyfikacyjne i regresyjne w badaniach marketingowych. Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie, Kraków.
Mączyńska, E. et al. (2004). Ile jest warta nieruchomość. Wydawnictwo Poltext, Warszawa.
Obrót nieruchomościami w 2012 r. 2013. GUS, Warszawa.
Paluch, Ł., Sroka, W. (2012). Identyfikacja i ocena czynników zrównoważonego rozwoju gmin wiejskich województwa małopolskiego. Roczniki Naukowe SERiA, XIV, 8, 81–87.
Wrzosek, W. (1994). Funkcjonowanie rynku. PWE, Warszawa.

WIELOWYMIAROWA ANALIZA PROCESU KSZTAŁTOWANIA SIĘ CEN NA LOKALNYM RYNKU NIERUCHOMOŚCI PRZY UŻYCIU METODY DRZEW REGRESYJNYCH C&RT

Streszczenie. Uwaga racjonalnych nabywców i sprzedawców nieruchomości koncentruje się na cenach i cechach obiektów podobnych. W świetle teorii wyceny zróżnicowanie cen tłumaczy się zróżnicowaniem kilku najistotniejszych cech, jakimi charakteryzują się nieruchomości stanowiące przedmiot obrotu. Stąd w praktyce szacowania wartości ważną kwestią jest rozstrzygnięcie, które z atrybutów mają istotne znaczenie i jaki jest ich indywidualny wkład w możliwą do uzyskania cenę rynkową. W artykule problem ten rozstrzygnięto przy wykorzystaniu modelu drzew regresyjnych C&RT. Badaniem objęto nieruchomości gruntowe niezabudowane, przeznaczone w planach miejscowych na cele budownictwa jednorodzinnego, położone na północnych peryferiach Krakowa. Drzewo regresyjne umożliwiło prostą ocenę wpływu poszczególnych atrybutów na ceny nieruchomości. Metoda drzew pozwala przewidywać wartość rynkową, ale wymaga wiedzy o co najmniej kilkudziesięciu, a najlepiej kilkuset transakcjach obiektami podobnymi. Z tego względu jest więc predestynowana do wycen o charakterze masowym.

Słowa kluczowe: rynek nieruchomości, drzewa regresyjne C&RT

Accepted for print: April 2015

For citation: Dacko M., Szajdecka K. (2015). Multifaceted analysis of the process of price developments on the local real estate market by means of the regression trees method (C&RT). Acta Sci. Pol., Oeconomia, 14 (2), 27–38.