

METODY DOBORU CECH DIAGNOSTYCZNYCH DO WYZNACZANIA PÓL CENNOŚCI I MODELOWANIA WARTOŚCI KATASTRALNEJ

Jan Kuryj, Oksana Kuryj-Wysocka

Uniwersytet Warmińsko-Mazurski

Streszczenie. Realizacja efektywnej gospodarki rynkowej wymaga znajomości wartości nieruchomości stanowiącej podstawę do podejmowania różnorodnych decyzji w zakresie gospodarowania i zarządzania nieruchomościami. Większość celów, dla których wymagana jest znajomość wartości nieruchomości, wynika z regulacji prawnych dotyczących gospodarowania nieruchomościami stanowiącymi własność publiczną. Jednym z rodzajów wartości jest wartość katastralna, ustalana w procesie masowej wyceny, wprowadzona prawnie w celu zreformowania systemu podatków od nieruchomości, określenia wartości nieruchomości publicznych, a także dokonywania czynności urzędowych. Sprawność przeprowadzenia masowej wyceny oraz dokładność otrzymanych wyników wskazują na potrzebę prowadzenia badań, których wynikiem byłoby określenie cech diagnostycznych umożliwiających wyznaczenie pól o jednorodnej cenności, modelowania masowej wyceny nieruchomości oraz cech rynkowych wpływających na wartość katastralną, metod ich doboru i selekcji. W opracowaniu zaproponowano metody statystyczne do wyboru cech diagnostycznych, które mogą być użyteczne w trakcie wyznaczenia pola cenności i wpływają na wartość katastralną nieruchomości. Przedstawiono także propozycję wyznaczenia ich procentowego udziału w kreowaniu tej wartości.

Słowa kluczowe: gospodarka nieruchomościami, wartość katastralna, wycena masowa, pola cenności

WSTĘP

Realizacja efektywnej gospodarki rynkowej wymaga znajomości wartości nieruchomości stanowiącej podstawę do podejmowania różnorodnych decyzji związanych z gospodarowaniem i zarządzaniem nieruchomościami.

Adres do korespondencji – Corresponding author: Jan Kuryj, Katedra Zasobów Nieruchomości, Uniwersytet Warmińsko-Mazurski w Olsztynie, ul. M. Oczapowskiego 2, 10-719 Olsztyn, e-mail: jkuryj@uwm.edu.pl

© Copyright by Wydawnictwo Uniwersytetu Warmińsko-Mazurskiego w Olsztynie, Olsztyn 2015

Wartość nieruchomości ma różne znaczenia. Raz jest to kwota, za którą nieruchomość można sprzedać czy kwota, za którą nieruchomość można wynająć lub wydzierżawić. Innym razem może być postrzegana jako koszt odtworzenia lub zastąpienia istniejącej zabudowy lub potencjalny dochód, który generuje lub może generować nieruchomość przy określonym sposobie użytkowania.

Większość celów, dla których wymagana jest znajomość wartości nieruchomości, wynika z obowiązujących w Polsce przepisów prawnych dotyczących głównie gospodarowania nieruchomościami stanowiącymi własność Skarbu Państwa czy jednostek samorządu terytorialnego. Coraz częściej również osoby prywatne i firmy deweloperskie podejmują swoje decyzje dotyczące inwestowania w oparciu o znajomość wartości nieruchomości. W procesach decyzyjnych najczęściej wykorzystywana jest wartość rynkowa.

Jednym z rodzajów wartości jest wartość katastralna. Wartość ta wprowadzona została art. 150, ust.1, pkt. 3 oraz 151, ust. 3 Ustawy z dnia 21 sierpnia 1997 r. o gospodarce nieruchomościami [Dz.U. z 2015 r. poz. 1774 i zm.]. Zgodnie z art.162, ust. 2 ww. ustawy wartość katastralna będzie wykorzystywana do następujących celów:

- ustalania podstawy opodatkowania podatkiem od nieruchomości;
- określania wartości nieruchomości stanowiących własność Skarbu Państwa lub jednostek samorządów terytorialnych w szczególnych przypadkach uregulowanych w ustawie o gospodarce nieruchomościami czy w odrębnych przepisach prawnych;
- wykonywania czynności urzędowych, do wykonywania których niezbędna jest znajomość wartości nieruchomości.

Ustalenie tej wartości odbywa się w procesie powszechnej taksacji nieruchomości, zwanej także wyceną masową. Wprowadzona ma być ona m.in. w celu reformy polskiego systemu podatkowego, dotyczącego sfery majątkowej, co ma ogromne znaczenie dla poprawy stanu finansów samorządów terytorialnych oraz doprowadzenia do zgodności z rozwiązaniami obowiązującymi w krajach Unii Europejskiej.

Jest to proces specyficzny, gdyż wymaga określania wartości dla dużej liczby nieruchomości w wydzielonych wcześniej obszarach o jednorodnych polach cenności (strefach taksacyjnych). Takie podejście wymaga użycia procedur możliwie szybkich i o sprawdzalnym stopniu wiarygodności oraz metod wykrywających zależności między atrybutami i pozwalających na szacowanie stabilnych w czasie, o ustalonej precyzji wartości [Kuryj 2001].

Wiarygodność przeprowadzenia masowej wyceny oraz dokładność otrzymanych wyników wymaga dostępu do kompletnych, odpowiedniej jakości, wiarygodnych i aktualnych danych o nieruchomościach – cech diagnostycznych. Posiadanie takich danych i określenie ich udziału w kreowaniu poziomu cen nieruchomości na danym rynku jest istotne z dwóch powodów. Po pierwsze, modelowanie masowej wyceny nieruchomości wymaga przeprowadzania analiz na dużych liczbowo zbiorach i po drugie, bez danych identyfikujących i charakteryzujących nieruchomości przeprowadzenie analiz przestrzennych za pomocą technologii informatycznej byłoby niemożliwe [Kuryj 2007].

Określona przepisami ustawy o gospodarce nieruchomościami i Rozporządzenia Rady Ministrów z dnia 29 czerwca 2005 r. w sprawie powszechnej taksacji nieruchomości [Dz.U. z 2005 r. nr 131, poz. 1092] procedura administracyjno-prawna przeprowadzania

powszechnej wyceny nieruchomości wskazuje na potrzebę prowadzenia badań. Efektem tych badań byłoby określenie standardów danych nie tylko parametrycznych, ale także rynkowych. Dzięki badaniom zaproponowanoby także metody doboru i selekcji cech diagnostycznych wpływających na wartość katastralną, niezbędne do prowadzenia analiz na etapie monitoringu rynku i ustalaniu zasięgu pól cenności (stref taksacyjnych) z wykorzystaniem narzędzi GIS [Cellmer 2014]. Ma to ogromne znaczenie w tworzeniu nowoczesnego zintegrowanego systemu informacji o nieruchomościach opartego na katastrze nieruchomości, katastrze fiskalnym i systemie elektronicznych ksiąg wieczystych.

TEORETYCZNE PODSTAWY WYZNACZANIA POLA CENNOŚCI I MODELOWANIA WARTOŚCI

Opierając się na teorii taksacji [Przewłocki 1994] wartość rynkową można opisać następującą zależnością:

$$W_r = W_o \cdot c \quad (1)$$

gdzie: W_r – wartość rynkowa,
 W_o – wartość początkowa,
 c – natężenie pola cenności.

W tym przypadku pole cenności zdefiniowano jako przyporządkowanie punktowi w określonej przestrzeni cenności określonej klasy obiektu. Natężenie tego pola jest wielkością niemianowaną i określającą stosunek danego człowieka do określonego obiektu pozwalający na zaspokojenie ludzkich potrzeb.

W procesie powszechnej taksacji nieruchomości pojawiła się wartość katastralna, która jest ustalana w procedurze masowej wyceny i stanowi między innymi podstawę do naliczenia podatku *ad valorem*. Wartość ta nie pojawia się na rynku, a tym samym nie jest przez ten rynek weryfikowana, pozostaje jedynie wartością ustaloną.

Po takim założeniu formuła teorii taksacji w zastosowaniu do masowej wyceny nieruchomości powinna przyjąć następującą postać:

$$W_k = W_o \cdot c_{st} \quad (2)$$

gdzie: W_k – wartość katastralna części składowej nieruchomości,
 W_o – wartość początkowa części składowej nieruchomości,
 c_{st} – natężenie pola cenności strefy taksacyjnej.

Wartość początkową poszczególnych składników nieruchomości w procesie masowej wyceny można określić na wiele sposobów, np.: poprzez wybór nieruchomości modelowej różnymi sposobami i jej wycenę [Kuryj 2007], poprzez uśrednienie skorygowanych ze względu na upływ czasu cen sprzedaży nieruchomości na obszarze danej strefy czy też urzędowo na podstawie map średnich cen transakcyjnych gruntów, opracowywanych przez Głównego Geodetę Kraju zgodnie z Rozporządzeniem RM

z dnia 3 października 2011 r. w sprawie rodzajów kartograficznych opracowań tematycznych i specjalnych [Dz.U. nr 222, poz. 1328].

Pole cenności (c_{st}) jest to pole subtelne, którego wielkością skalarną jest cenność określonej klasy obiektów w odniesieniu do nieruchomości. Może być przedstawione w postaci rozkładu natężenia zjawisk przestrzennych w czasie, jako suma natężenia pól cenności wyselekcjonowanych obiektywnych cech (diagnostycznych) przedmiotowych nieruchomości (c_i) po zweryfikowaniu ich przez rynek w momencie transakcji:

$$c_{st} = \sum_{i=1}^n c_i \quad (3)$$

W oparciu o informacje z monitoringu rynku można oszacować natężenia pól cenności rynkowej wyselekcjonowanych cech rynkowych nieruchomości poprzez badanie ich wpływu na ulepszenie wartości początkowej nieruchomości. Liczba cech zakwalifikowana do badania cenności powinna być możliwie największa, a każda z nich jednoznacznie zdefiniowana. Jak wskazuje w swojej rozprawie M. Krajewska, jedną z cech wpływających na zmienność pola cenności dla ściśle określonej lokalizacji jest jego przeznaczenie [Krajewska 2008].

Selekcja cech diagnostycznych do wyznaczenia natężenia pola cenności powinna być oparta na modelach regresji krokowej postępującej i wstecznej oraz na analizie korelacji z uwzględnieniem ich zmienności w czasie spowodowanych lokalnymi uwarunkowaniami rynku nieruchomości.

PRZYJĘTE METODY BADAWCZE

Objaśnienie reguł złożoności badanego zjawiska uzależnione jest od wielu czynników objaśniających to zjawisko. Czynniki te mogą być powiązane ze sobą w sposób bezpośredni – dają się zaobserwować albo pomierzyć. Powiązane mogą być także w sposób pośredni – nie są możliwe do zaobserwowania, a ich wpływu jedynie można się domyślać ([Kuryj 2007]. Czynniki pośrednie często zakłócają dane zjawisko i w konsekwencji mogą przyczynić się do błędnego jego opisu.

W naukach przyrodniczych, ekonomiczno-społecznych i technicznych [Kowal 1998, Walesiak 1996], a także w badaniach związanych z przeprowadzeniem masowej wyceny nieruchomości [Adamczewski 2006, Bitner 2010, Cellmer 2014, Czaja i Parzych 2015, Kuryj 2007], gdy badane zjawisko opisywane jest za pomocą większej liczby zmiennych (atrybutów), poczesne miejsce, w objaśnianiu jego złożoności, zajmują metody statystyczne (tj.: analizy wielowymiarowe i analiza regresji wielokrotnej).

Kierując się tymi przesłankami w prezentowanym opracowaniu zaadoptowano dwie metody statystyczne, tj.: analizę korelacji oraz analizę regresji krokowej (wstecznej i postępującej) do selekcji atrybutów mniej istotnych w objaśnianiu poziomu kreowania wartości i do wyboru cech diagnostycznych istotnych w wyznaczaniu zasięgu pól cenności i modelowaniu wartości nieruchomości.

Stosowanie analizy regresji wielokrotnej wymaga zgromadzenia próby badawczej – dostatecznej liczby danych o transakcjach rynkowych (o cenach transakcyjnych i atrybutach nieruchomości) oraz wyselekcjonowania najbardziej statystycznie dopasowanej grupy zmiennych niezależnych – cech diagnostycznych nieruchomości. Szczególnie przydatna dla tego celu może okazać się analiza regresji krokowej, zarówno wstecznej, jak i postępującej ([Kuryj 2001]. Regresja krokowa wsteczna polega na eliminacji z modelu w kolejnych iteracjach atrybutów, które w najmniejszym stopniu wyjaśniają wariancję zmiennej zależnej. Wyznaczone tą metodą współczynniki regresji mogą posłużyć do ustalenia wpływu poszczególnych cech na cenę transakcyjną. Regresja krokowa postępująca pozwala na utworzenie optymalnego równania regresji, poprzez dodanie do niego, w kolejnych iteracjach, tylko tych cech, które w największym stopniu wyjaśniają wariancję zmiennej zależnej.

Metoda analizy korelacji, jak wskazują Czaja i Parzych [2015], może okazać się przydatna do określenia zależności statystycznej między zmiennymi objaśniającymi (atrybutami) przyjętymi do analizy a zmienną objaśnianą (ceną transakcyjną). Miernikiem tej zależności jest współczynnik korelacji liniowej Pearsona (r).

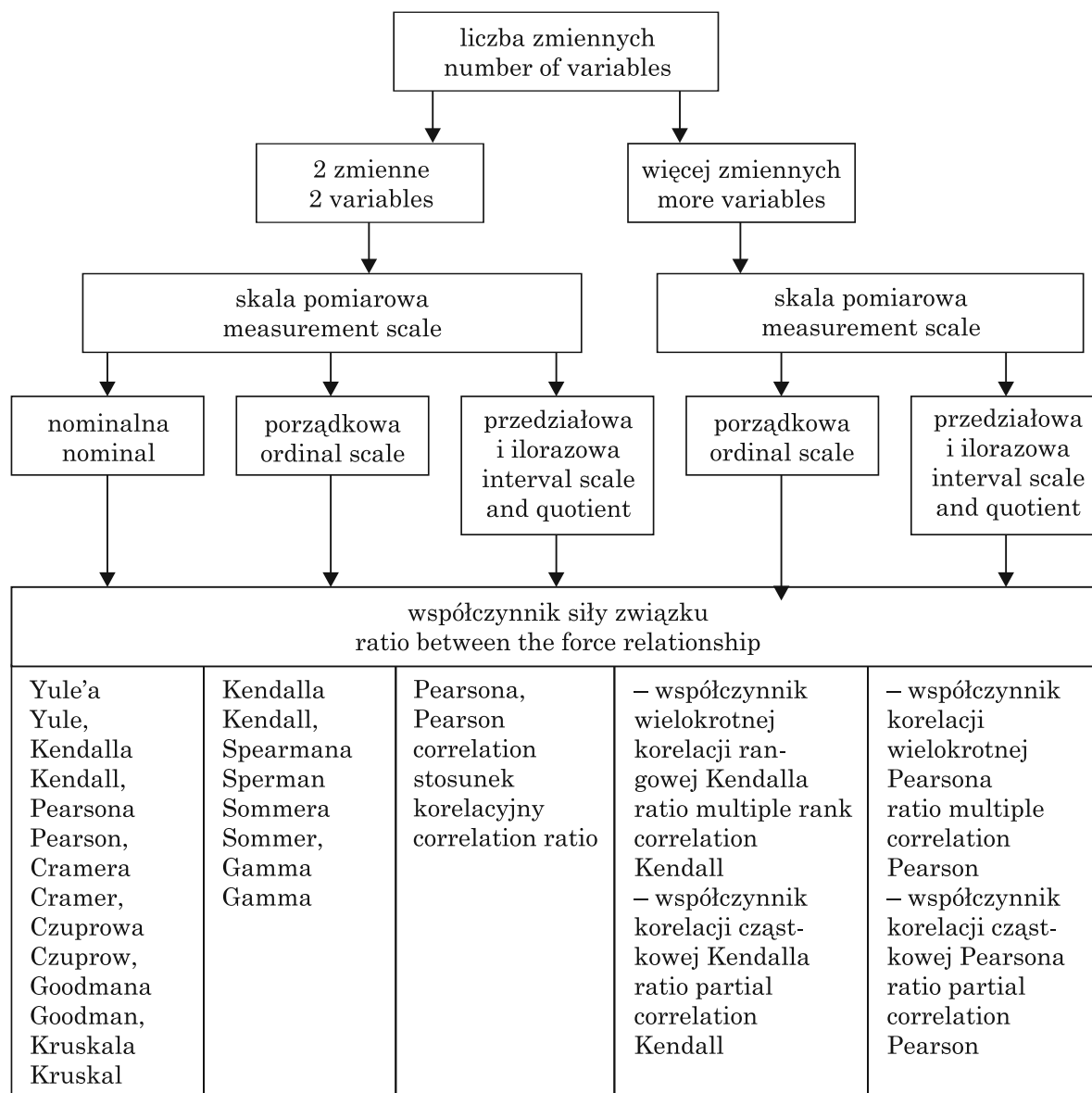
Współczynnik ten przyjmuje wartości z przedziału od -1 do 1. Znak współczynnika korelacji wskazuje na kierunek zależności, a jego wartość bezwzględna na siłę zależności między dwiema zmiennymi. Siłę tę, jak sugerują Czaja i Parzych [2015], na potrzeby wyznaczania pól cenności, jak i modelowania masowej wyceny nieruchomości, można określić na czterech poziomach, tj.:

- dla: $|r| \leq 0,3$ – korelacja słaba;
- dla: $0,3 \leq |r| \leq 0,6$ – korelacja przeciętna;
- dla: $0,6 \leq |r| \leq 0,9$ – korelacja silna;
- dla: $> 0,9 \leq |r|$ – korelacja bardzo silna.

PROPOZYCJA METODYKI WYBORU CECH DIAGNOSTYCZNYCH

W celu zaproponowania metody doboru cech diagnostycznych do przeprowadzenia wyceny masowej i określenia pól cenności wpływających na wartość nieruchomości postawiono tezę: występuje istotny związek liniowy między cechami opisującymi nieruchomości a ceną sprzedaży nieruchomości. W tym celu dokonano analizy metod statystycznych pomocnych w rozwiązywaniu klasycznego problemu, który stanowi ocena siły związku między zmiennymi. Istotą metod badania zależności między zmiennymi jest wybór i zastosowanie optymalnego miernika – współczynnika współzależności lub korelacji.

Udowadniając sformułowaną hipotezę, zaproponowano procedurę doboru możliwych do zastosowania współczynników, uzależniając ich dobór od liczby zmiennych i ich skal pomiarowych (rys. 1).



Rys. 1. Procedura wyboru współczynnika badania siły związku między zmiennymi

Fig. 1. Procedure for selecting the study ratio of strength relationship between the variables

Źródło: opracowanie na podstawie Kowala [1998]

Source: study based on literature: Kowal [1998]

W zaproponowanej procedurze wyróżniono następujące etapy:

- określenie liczby zmiennych (cech) opisujących badane obiekty, zdefiniowanie skal pomiarowych, dokonanie transformacji z różnych skal pomiarowych w celu ujednolicenia opisu cech;
- określenie liczebności zbioru obserwacji i wylosowanie próby badawczej;
- wyszczególnienie możliwych do zastosowania współczynników z uwzględnieniem przyjętych założeń, zgodnie z przedstawionym schematem (rys. 1);
- wybór optymalnego współczynnika, który przy występujących ograniczeniach jest obciążony najmniejszym błędem i obliczenie jego wartości;

- e) przetestowanie hipotezy zerowej o braku związku liniowego między zmiennymi za pomocą testu istotności;
- f) wybór zmiennych (cech diagnostycznych) o najwyższej wartości współczynników korelacji po odrzuceniu hipotezy zerowej, przyjmując zasadę minimalizacji liczby zmiennych.

Uwzględniając charakterystyki atrybutów opisujących nieruchomości, rynek nieruchomości oraz założenia wynikające z przedstawionej na rysunku 1 procedury do badania wpływu atrybutów na kreowanie cen transakcyjnych, a tym samym na ustalenie wartości katastralnej, zastosowano dwie metody statystyczne:

- analizę korelacji z oceną wartości współczynnika korelacji liniowej Pearsona;
- analizę regresji krokowej postępującej i wstecznej z wykorzystaniem zestandaryzowanych współczynników regresji BETA.

Za wyborem tych metod przemawia fakt, że zarówno analiza regresji, jak i korelacji dotyczą wielowymiarowej zmiennej losowej, w której jedna zmienna stanowi zmienną zależną, zaś pozostałe zmienne mają charakter zmiennych objaśniających (niezależnych). Analogicznie w analizie statystycznej rynku, przeprowadzanej na potrzeby masowej wyceny, jednocześnie rozpatruje się cenę i kilka (kilkanaście) atrybutów opisujących nieruchomości. Związki między nimi można opisywać za pomocą jednego modelu wielowymiarowej regresji wielokrotnej lub za pomocą kilku niezależnych modeli dwuwymiarowej regresji prostej.

W celu przetestowania zaproponowanej procedury przeprowadzono monitoring rynku nieruchomości w południowo-wschodniej części Olsztyna. Monitoringiem objęto grunty niezabudowane i zabudowane przeznaczone pod budownictwo mieszkaniowe jedno- i wielorodzinne na osiedlach Jaroty i Brzeziny, które tworzą obszar jednorodny pod względem użytkowania.

Do opisu działek gruntu przyjęto zestaw cech parametrycznych, zgodnie z §6 ust. 1. rozporządzenia Rady Ministrów z 2005 r. w sprawie powszechnej taksacji nieruchomości oraz z §19 ust. 2. Wytycznych przeprowadzenia powszechnej taksacji nieruchomości [Adamczewski i in. 2002], które nadal mają status projektu. Ten zestaw cech powiększono o cechy dodatkowe wysondowane w badaniach ankietowych jako ważne dla respondentów na badanym obszarze (jako potencjalne atrybuty wpływające na cenę).

Łącznie zgromadzono 132 transakcje nieruchomościami, które opisano dziesięcioma atrybutami, tj.: położenie (X_1), przeznaczenie w planie miejscowym (X_2), stan wyposażenia w urządzenia infrastruktury technicznej (X_3), stan zagospodarowania (X_4), intensywność zabudowy (X_5), klasa gleboznawcza gruntu (X_6) oraz dostępność do obiektów przyrodniczych i rekreacyjnych (X_7), powierzchnia działki gruntowej (X_8), dojazd i dostęp do działki (X_9) i uciążliwość sąsiedztwa (X_{10}).

Cechy te pomierzono zgodnie z wytycznymi przeprowadzania PTN na skali porządkowej pięciostopniowej. Pomiar przeprowadzono dla wszystkich cech wg przyjętego założenia, że 1 jest najgorszą rangą dla mierzonej cechy, a 5 – najlepszą. Cechy mierzalne, jak: X_1 , X_5 , X_8 pomierzono na skali ilorazowej, a następnie przeprowadzono transformację na skalę porządkową rangową. W tym celu zastosowano metodę najmniejszych kwadratów wygładzenia linii symulującej rozkład cen względem wymienionych cech ważoną odległością.

Do selekcji i wyboru cech diagnostycznych przyjęto nieruchomości będące przedmiotem transakcji na badanym rynku po skorygowaniu ich cen transakcyjnych trendem zmiany cen, wyznaczonym metodą regresji prostej, na koniec roku kalendarzowego.

PROPOZYCJA METODY WYBORU CECH DIAGNOSTYCZNYCH

Metoda analizy korelacji

W celu ustalenia siły związku między przyjętymi cechami nieruchomości oraz cechami a ceną transakcyjną zaadaptowano metodę analizy korelacji, w której jako miernik siły związku i istotności cech zastosowano współczynnik korelacji zupełnej Pearsona „ r ”.

Własności współczynnika korelacji Pearsona „ r ” wywodzą się wprost z własności kowariancji. Znak współczynnika korelacji wskazuje na kierunek zależności, a wartość bezwzględna mówi o sile zależności między dwiema zmiennymi. Im wartość bezwzględna współczynnika jest bliższa jedności, tym silniejsza jest zależność liniowa między badanymi zmiennymi.

Do wyboru cech diagnostycznych przyjęto skalę wskazaną przez Czaję i Parzycha [2015], modyfikując ją i dostosowując do warunków badanego rynku nieruchomości. Po modyfikacji przedziały oceny istotności cech przyjęły następujące wartości współczynnika Pearsona „ r ”:

- dla: $|r| \leq 0,2$ – korelacja słaba, cechy nie istotne;
- dla: $0,2 \leq |r| \leq 0,5$ – korelacja przeciętna, cechy mogą być; uznane za istotne;
- dla: $0,5 \leq |r| \leq 0,8$ – korelacja silna, cechy istotne i wpływają; na poziom cen;
- dla: $|r| > 0,8$ – korelacja bardzo silna, cechy bardzo istotne.

Selekcję zmiennych niezależnych, na potrzeby wyznaczenia cenności pól i modelowania masowej wyceny, przeprowadzono w macierzy korelacji (tab. 1). Oceniono wartości współczynnika Pearsona, badając korelację przyjętych cech opisujących nieruchomości z ceną transakcyjną, z zachowaniem warunku braku korelacji lub słabego skorelowania zmiennych objaśniających (niezależnych).

Do obliczenia współczynników korelacji liniowej Pearsona wykorzystano oprogramowanie firmy StatSoft Statistica. Wyniki badania siły związków między cechami oraz cechami i ceną transakcyjną zaprezentowano w macierzy korelacji (tab. 1). Oznaczone współczynniki korelacji są istotne z $p < 0,05$.

Wybór cech rynkowych wpływających na ceny transakcyjne poprzedzono badaniem istotności obliczonych współczynników poprzez testowanie hipotezy zerowej o braku związku liniowego ($H_0: r = 0$). Do tego celu wykorzystano statystykę T, która ma rozkład t-Studenta i obliczana jest według wzoru:

$$t = r \sqrt{\frac{N-2}{1-r^2}} \quad (5)$$

gdzie:

N – liczebność próby,

r – wartość współczynnika korelacji.

Tabela 1. Obliczenie wartości współczynnika korelacji Pearsona dla gruntów
 Table 1. The calculation of the Pearson correlation factor for land

Zmienne Variables	Oznaczone wsp. korelacji są istotne z $p < 0,05000$, $N = 132$ (braki danych usuwano przypadkami) Labeled et all correlation is significant with $p < 0.05000$, $N = 132$ (lack of data removed cases)										
	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	C_{lm}^2
X_1	1,000	-0,013	0,028	-0,234	0,033	0,226	0,595	0,155	-0,561	0,612	0,641
X_2	-0,013	1,000	-0,097	0,009	-0,127	-0,184	0,258	0,476	-0,461	0,261	0,111
X_3	0,028	-0,097	1,000	0,479	-0,267	0,047	0,009	0,303	0,087	0,025	0,041
X_4	-0,234	0,009	0,479	1,000	-0,293	-0,126	-0,011	0,221	0,106	-0,220	-0,091
X_5	0,033	-0,127	-0,267	-0,293	1,000	0,031	-0,163	0,110	0,238	0,317	0,250
X_6	0,226	-0,184	0,047	-0,126	0,031	1,000	0,205	-0,312	-0,174	-0,126	0,084
X_7	0,595	0,258	0,009	-0,011	-0,163	0,205	1,000	0,137	-0,626	0,042	0,392
X_8	0,155	0,476	0,303	0,221	0,110	-0,312	0,137	1,000	-0,025	0,635	0,449
X_9	-0,561	-0,461	0,087	0,106	0,238	-0,174	-0,626	-0,025	1,000	-0,185	-0,353
X_{10}	0,612	0,261	0,025	-0,220	0,317	-0,126	0,042	0,635	-0,185	1,000	0,595
C_{lm}^2	0,641	0,111	0,041	-0,091	0,250	0,084	0,392	0,449	-0,353	0,595	1,000

Źródło: obliczenia własne z wykorzystaniem oprogramowania StatSoft Statistica
 Source: own study from using software StatSoft Statistica

Hipotezę H_0 odrzucano na korzyść hipotezy alternatywnej o istnieniu związku liniowego między analizowanymi zmiennymi, gdy wartość t obliczona ze wzoru (5) była większa od wartości krytycznej odczytanej z tablic rozkładu t -Studenta.

Na podstawie macierzy korelacji wybrano następujące cechy diagnostyczne opisujące nieruchomości gruntowe i skorelowane z ceną transakcyjną, tj.: położenie (X_1), intensywność zabudowy (X_5), powierzchnia działki gruntowej (X_8), dojazd i dostęp do działki (X_9). Cechy diagnostyczne pogrubiono w tabeli 1 i zdefiniowano jako cechy wpływające na poziom cenności obszaru i ustalenie wartości katastralnej.

Wybierając te cechy, uwzględniono wartość obliczonego wskaźnika istotności t dla hipotezy alternatywnej, w badaniu siły związku między cechami a ceną transakcyjną, oraz występowanie hipotezy zerowej w przypadku badania zależności między poszczególnymi cechami.

Z macierzy korelacji wynika, że silna i przeciętna korelacja z ceną transakcyjną występuje także przy cechach: dostępność do obiektów przyrodniczych i rekreacyjnych (X_7) oraz uciążliwość sąsiedztwa (X_{10}), ale cechy te są silnie skorelowane z innymi cechami opisującymi nieruchomości, dlatego nie zostały wytypowane jako cechy diagnostyczne.

Stosując analizę wariancji, można wskazać, że współczynnik Pearsona r ma ścisły związek z wiarygodnością prognozy zmiennej zależnej (ceny transakcyjnej). Zatem bezwzględna jego wartość może stanowić wagę dokładności prognozy, a tym samym procentowy wskaźnik wpływu cechy na poziom kształtowania cen i natężenia pola cenności

tej cechy. Przyjmując tę zależność, zaproponowano algorytm (6) obliczenia natężenia pola cenności cechy i jej wpływu na wartość początkową. Wielkość natężenia pola cechy można przedstawić jako wielkość niemianowaną lub wielkość procentową poprzez transformację bezwzględnych wartości współczynników korelacji według następującej formuły matematycznej:

$$c_i = \frac{|r_i|}{\sum_{j=1}^n |r_j|} \cdot 100\% \quad (6)$$

gdzie:

- c_i – wielkość natężenia pola dla i -tej cechy i jej wpływu na wartość w %;
- $|r_i|$ – bezwzględna wartość współczynnika korelacji i -tej cechy z ceną transakcyjną;
- n – liczba wyselekcjonowanych cech diagnostycznych.

Obliczenie procentowych wielkości natężenia pola cenności dla wyselekcjonowanych cech według wzoru (6) przedstawiono w tabeli 2.

Tabela 2. Obliczenia natężenia pola cenności cechy i jej wpływu na wartość początkową
Table 2. Calculation of field strength characteristics of the preciousness and its impact on the initial value

Cechy Attribute (n)	$ r $	Waga atrybutu The weight attribute [%]	Natężenie pola cenności Field strengththe precious c_i [%]
X_1	0,6413	37,86	38
X_5	0,2498	14,75	15
X_8	0,4494	26,54	26
X_9	0,3532	20,85	21
Σ	1,6938	100,00	100

Źródło: obliczenia własne
Source: own study

Metoda regresji krokowej postępującej i wstecznej

W celu porównania wyników otrzymanych z analizy korelacji zaadoptowano do wyznaczenia istotności cech diagnostycznych metodę regresji krokowej postępującej i wstecznej jako kolejne metody. Według tych metod analizowana jest wariancja zmiennej zależnej. W poszczególnych krokach w regresji krokowej postępującej wprowadzane są do równania regresji predyktory – zmienne niezależne (cechy), które w największym stopniu wyjaśniają wariancję zmiennej zależnej. Następnie odejmowana jest wariancja wyjaśniona przez tę zmienną, obliczane są nowe korelacje dla każdej z potencjalnych

zmiennych niezależnych i dla poprawionej zmiennej zależnej. W efekcie wybierana jest zmienna niezależna z najwyższą korelacją w danym kroku. W ten sposób do równania predyktywnego są dodawane zmienne do momentu, gdy wszystkie predyktory istotne zostaną wytypowane, lub gdy procedura zostanie zastopowana w oparciu o wcześniejsze założenia.

Na potrzeby prowadzonej analizy przyjęto założenie, że równanie zostanie rozwiązane wówczas, gdy w następnym kroku wariancja będzie wyjaśniona mniej niż 1%. Poprawność otrzymanego modelu regresji zbadano za pomocą trzech statystyk:

- współczynnika korelacji wielokrotnej – r ;
- współczynnika determinacji – r^2 ;
- zestandaryzowanych współczynników regresji dla zmiennych – BETA.

Do wyboru reprezentacji z próby generalnej zgromadzonych transakcji na potrzeby prowadzonej analizy zaadaptowano miernik – odległość Cooka, która jest miarą oddziaływania danego przypadku (transakcji) na rozwiązanie równania regresji. Odległość ta wskazuje różnicę między obliczoną wartością współczynników regresji B a taką samą wartością obliczoną po założeniu, że dana obserwacja (transakcja) została wyłączona z równania regresji. Jak wynika z przeglądu literatury przedmiotu, wszystkie odległości powinny być tego samego rzędu i zbliżone do zera. Jeśli nie są, to można przypuszczać, że przypadki, dla których obliczona odległość Cooka jest największa, mają istotny wpływ na obciążenie równania regresji. W związku z tym przyjęto założenie, że do rozwiązania równania regresji wykorzystane zostaną tylko te transakcje, dla których obliczona odległość Cooka jest mniejsza od 0,01. Wyniki zamieszczono w tabeli 3.

Tabela 3. Podsumowanie regresji krokowej postępującej dla zmiennej zależnej: C_{1m}^2
Table 3. Summary of stepwise regression progressive for the dependent variable: C_{1m}^2

Cechy diagnostyczne Diagnostic attributes $N = 72$	$R = 0,75747440$ $F(3,28)=12,564$		$R^2 = 0,57376746$; $p < ,00002$;		Poprawione – Corrected: $R^2 = 0,52809969$ Bł. std. estymacji – Error estimation: 22,895	
	BETA	bł. std.	b	bł. std.	$t(67)$	p
W. wolny			132,49	24,771	5,3484	0,0000
X_1	0,5825	0,1249	17,37	3,725	4,6637	0,0000
X_8	0,3376	0,1256	9,50	3,533	2,6883	0,0119
X_5	0,1937	0,1241	6,31	4,042	1,5607	0,1298

Źródło: obliczenia własne z wykorzystaniem oprogramowania StatSoft Statistica

Source: own study from using software StatSoft Statistica

Stosując metodę regresji krokowej postępującej, wybrano następujące cechy diagnostyczne, które w największym stopniu wyjaśniają wariancję ceny transakcyjnej, tj.: położenie (X_1), intensywność zabudowy (X_5), powierzchnia działki gruntowej (X_8). Cechy te wraz z wyznaczonymi charakterystykami dla modelu regresji krokowej zamieszczono w tabeli 3.

Współczynnik determinacji R^2 wskazuje, że 57% zmienności zmiennej zależnej (ceny transakcyjnej) jest wyjaśniane za pomocą wytypowanych zmiennych niezależnych.

Regresja krokowa wsteczna polega na działaniu odwrotnym niż ma to miejsce w regresji postępującej. W kolejnych iteracjach (krokach) odrzucane są z równania regresji predyktory – zmienne niezależne (cechy), które w najmniejszym stopniu wyjaśniają wariancję zmiennej zależnej. W efekcie pozostają w równaniu tylko zmienne niezależne z najwyższą korelacją w danym kroku. W ten sposób z równania predyktywnego są odrzucane zmienne do momentu, kiedy wszystkie predyktory istotne zostaną wytypowane, lub gdy procedura zostanie zastopowana w oparciu o wcześniejsze założenia. Wyniki podsumowania regresji krokowej wstecznej zamieszczono w tabeli 4.

Tabela 4. Podsumowanie regresji krokowej wstecznej dla zmiennej zależnej: C_{1m}^2
Table 4. Summary of stepwise regression backward for the dependent variable: C_{1m}^2

Cechy diagnostyczne Diagnostic attributes $N = 72$	$R = 0,76187126$; $F(4,27) = 9,3386$;		$R^2 = 0,58044782$; $p < ,00007$;		Poprawiona – Corrected: $R^2 = 0,51829194$ bł. std. estymacji – error estimation: 23,132	
	BETA	bł. std.	b	bł. std.	$t(67)$	p
W. wolny			149,68	36,250	4,12921	0,00031
X_1	0,52286	0,15561	15,59	4,640	3,36005	0,00234
X_5	0,21998	0,13165	7,16	4,287	1,67095	0,10628
X_8	0,34147	0,12704	9,60	3,574	2,68805	0,01216
X_9	-0,10389	0,15845	-4,22	6,435	-0,65568	0,51758

Źródło: obliczenia własne z wykorzystaniem oprogramowania StatSoft Statistica
Source: own study from using software StatSoft Statistica

Stosując metodę regresji krokowej wstecznej, wybrano następujące cechy diagnostyczne, opisujące nieruchomości gruntowe, które w największym stopniu wyjaśniają wariancję ceny transakcyjnej, tj.: położenie (X_1), intensywność zabudowy (X_5), powierzchnia działki gruntowej (X_8), dojazd i dostęp do działki (X_9). Cechy te wskazano w tabeli 4 i zdefiniowano jako wpływające na poziom cenności obszaru i ustalenie wartości katastralnej.

Zestandaryzowane współczynniki regresji BETA posłużyć mogą do określenia wpływu dobranych do równania regresji cech diagnostycznych na kreowanie cenności pola cechy i wpływu jej na wartość.

Stosowanie tych współczynników jest szczególnie przydatne w przypadku zmiennych wyrażonych na różnych skalach pomiarowych, gdyż trudno je wówczas między sobą porównywać.

Standaryzacji dokonuje się poprzez przekształcenie zmiennych i wyrażane ich w jednostkach odchylenia standardowego. W wyniku takiego przekształcenia wyznaczone współczynniki BETA są pierwotnymi współczynnikami regresji B , podzielonymi przez odchylenie standardowe.

Zmienne z wyższymi wartościami współczynników BETA mają większe znaczenie w predykcji zmiennej zależnej. Wielkość wpływu danej cechy diagnostycznej na jej pole cenności i wartość katastralną można ustalić, stosując algorytm:

$$c_i = \frac{|BETA_i|}{\sum_{j=1}^n |BETA_j|} \cdot 100\% \quad (7)$$

gdzie: c_i – wielkość natężenia pola i -tej cechy i wpływu jej a wartość w %;
 $|BETA_i|$ – bezwzględna wartość zestandaryzowanego współczynnika BETA dla i -tej cechy diagnostycznej;
 n – liczba wprowadzonych do równania cech diagnostycznych.

Wyniki obliczenia wpływu wytypowanych w metodzie regresji krokowej postępującej i wstecznej cech diagnostycznych zaprezentowano w tabeli 5 i tabeli 6.

Tabela 5. Wpływ cech diagnostycznych wyznaczonych regresją krokową postępującą na natężenie pola cenności

Table 5. The effect of diagnostic characteristics designated stepwise regression progressive on the field the precious

Cechy Attribute (n)		Waga cechy The weight attribute [%]	Natężenie pola cenności Field strength the precious c_i [%]
X_1	0,58255	52,29	52
X_8	0,33767	30,32	30
X_5	0,19376	17,69	18
Σ	1,11398	100,00	100

Źródło: obliczenia własne

Source: own study

Tabela 6. Wpływ cech diagnostycznych wyznaczonych regresją krokową wsteczną na natężenie pola cenności

Table 6. The effect of diagnostic characteristics designated stepwise regression backward on the field the precious

Cechy Attribute (n)	BETA	The weight attribute [%]	Natężenie pola cenności Field strength the precious c_i [%]
X_1	0,52286	44,00	44
X_5	0,21998	18,51	18
X_8	0,34148	28,74	29
X_9	0,10389	8,74	9
Σ	1,18821	100,00	100

Źródło: obliczenia własne

Source: own study

W każdej z zaproponowanych metod doboru cech diagnostycznych wybierane zmienne niezależne w największym stopniu objaśniają zmienną zależną. Te same zmienne typowane były w każdej z zastosowanych metod i przyjmują zbliżoną wartość procentowego wpływu na wyznaczenie pola cenności cechy i wartości katastralnej, co ilustruje zestawienie w tabeli 7.

Tabela 7. Zestawienie wpływu cech na wyznaczenie pola cenności cech i wpływu na ustalenie wartości katastralnej

Table 7. Summary of the impact on the determination of the characteristics of the field valuables characteristics and impact on the determination of the cadastral value

Cechy Attributes	Analiza korelacji Analysis of the correlation [%]	Regresja postępująca The stepwise regression progressive [%]	Regresja wsteczna The stepwise regression backward [%]
X_1	38	52	44
X_5	15	18	18
X_8	26	30	29
X_9	21	–	9

Źródło: obliczenia własne

Source: own study

Liczba zmiennych niezależnych (cech diagnostycznych) zastosowanych do opisu nieruchomości, które mogą być istotne na badanym rynku lokalnym, w wyniku zastosowania zaproponowanych metod zmniejszyła się z dziesięciu (przyjętych na wstępie) do czterech, a w przypadku regresji krokowej postępującej do trzech.

Jak wynika z oceny współczynnika determinacji R^2 , zmniejszenie liczebności cech nie wpływa w istotnym stopniu na pogorszenie wyników określenia wartości katastralnej modelem ekonometrycznym, gdyż R^2 po wykorzystaniu wszystkich cech przyjętych do opisu nieruchomości kształtował się na poziomie 59,90% wyjaśnienia ceny transakcyjnej, a po wyselekcjonowaniu tylko czterech cech diagnostycznych (istotnych) przyjmuje wartość 58,04%. Z kolei R^2 dla trzech cech wybranych metodą regresji krokowej postępującej wynosi 57,38%.

Na rynku nieruchomości ceny transakcyjne (zmienne zależne) cechują się dużą zmiennością, podyktowaną zaangażowaniem emocjonalnym zbywców i nabywców, zależą od przebiegu negocjacji oraz innych czynników zewnętrznych wcześniej wskazanych w prezentowanym opracowaniu. Zdarza się, że podobne nieruchomości będące przedmiotem zbycia mają różne ceny sprzedaży. Dlatego też wartość współczynnika determinacji na poziomie ok. 60% jest wielkością zadowalającą do wyjaśnienia zjawisk zachodzących na tym rynku. Zawężenie liczby cech diagnostycznych z dziesięciu do czterech czy nawet do trzech, jak w przypadku cech wyselekcjonowanych metodą regresji krokowej postępującej, spowodowało zmniejszenie adekwatności modelu regresji o ok. 2%. Na tej podstawie można wnioskować, że nie wpłynie to znacząco na jakość modelowania wartości nieruchomości i ustalenia pola cenności.

PODSUMOWANIE

Z przeprowadzonych badań wynika, iż zaproponowana procedura i przetestowane metody statystyczne można zastosować do wyboru: cech diagnostycznych użytecznych do wyznaczania pola cenności, istotnych zmiennych niezależnych do budowy modeli ekonometrycznych oraz do określania na ich podstawie wartości (w tym wartości katastralnej). Zmienne niezależne definiowane są jako istotne cechy diagnostyczne wpływające na wartość i poziom cen transakcyjnych na danym rynku lokalnym.

Metoda korelacji jest metodą pracochłonną i po części subiektywną, ponieważ o odrzuceniu zmiennych uważanych za nieistotne decyduje badacz. Wymaga to dużego doświadczenia i znajomości lokalnego rynku.

Metoda regresji krokowej jest metodą zautomatyzowaną. W pakietach statystycznych programów komputerowych, jak np.: STATISTICA, firmy StatSoft, jest opracowany moduł umożliwiający przeprowadzanie analiz tą metodą. Dobieranie zmiennych niezależnych do równania regresji z najwyższą korelacją w każdym kroku przebiega automatycznie, po uwzględnieniu przyjętego przez badacza założenia lub zdefiniowaniu liczby cech istotnych. W każdym kroku tej analizy dla wprowadzonych zmiennych istnieje możliwość oceny poprawności sformułowanego modelu po objaśnieniu zmiennej zależnej za pomocą trzech statystyk: współczynnika korelacji wielokrotnej, współczynnika determinacji oraz statystyki (testu) *F*-Snedecora.

Analiza regresji krokowej umożliwia wprowadzenie do modelu ekonometrycznego tylko tych cech nieruchomości (predyktorów), które w największym stopniu są skorygowane ze zmienną zależną i umożliwiają jej prognozę, tj.: prognozowanie wartości nieruchomości. Stosując tę metodę, z dużej liczby czasem zbędnych atrybutów (cech nieruchomości), można wyłonić te, które rzeczywiście mają znaczący wpływ na predykcję wartości. Co więcej, metoda regresji krokowej pozwala na eliminację problemu współliniowości, czyli silnie skorelowanych ze sobą predyktorów (cech).

Jak pokazują zaprezentowane wyniki, metoda regresji krokowej, zarówno postępującej, jak i wstecznej, może być z powodzeniem wykorzystywana jako optymalna metoda do wyboru cech diagnostycznych umożliwiających: wyznaczanie pól o jednorodnej cenności, modelowanie masowej wyceny nieruchomości oraz wybór cech rynkowych wpływających na kształtowanie poziomu cen na danych rynkach nieruchomości.

PIŚMIENNICTWO

Adamczewski, Z. (2006). Elementy modelowania matematycznego w wycenie nieruchomości. Podejście porównawcze. Oficyna Wydawnicza PW, Warszawa.

Adamczewski, Z., Bojar, Z., Telegra, T. (2002). Wytyczne przeprowadzenia powszechnej taksacji nieruchomości (projekt). Przegląd Geodezyjny 6, 6–11.

Bitner, A. (2010). O użyteczności metod statystycznych w wycenie nieruchomości. Infrastruktura i Ekologia Terenów Wiejskich 12, 145–158.

Czaja, J., Parzych, P. (2015). Szacowanie rynkowej wartości nieruchomości. AGH, Kraków.

Cellmer, R. (2014). Modelowanie przestrzenne w procesie opracowania map wartości gruntów. UWM, Olsztyn.

- Cellmer, R., Kuryj, J. (2003). Interpretacja wyników analizy statystycznej cen transakcyjnych w procesie sporządzania map wartości gruntów. Bogucki Wydawnictwo Naukowe UAM, Poznań, 257–266.
- Głuszak, M., Moranka, B. (2015). Podatek katastralny. Ekonomiczne uwarunkowania reform opodatkowania nieruchomości. Poltext, Warszawa.
- Kowal, J., (1998). Metody statystyczne w badaniach sondażowych rynku. PWN, Warszawa–Wrocław.
- Krajewska, M. (2008). Wartość nieruchomości zabudowanej w różnych stanach planistycznych. *Studia i Materiały Towarzystwa Naukowego Nieruchomości* 16(3), 47–54.
- Kuryj, J. (2001). Metody doboru i oceny danych w taksacji powszechnej nieruchomości. Rozprawa doktorska. UWM, Olsztyn (maszynopis).
- Kuryj, J. (2007). Metodyka wyceny masowej nieruchomości na bazie aktualnych przepisów prawnych. *Wycena* 4(81), 50–58.
- Model ustalania i weryfikacji stref taksacyjnych dla potrzeb powszechnej taksacji nieruchomości. I. (1994). Przewłocki, S. (red.). PŁ, Łódź
- Rozporządzenie Rady Ministrów z dnia 29 czerwca 2005 r. w sprawie powszechnej taksacji nieruchomości. *Dz.U. z 2005 r. nr 131, poz. 1092.*
- Rozporządzenie Rady Ministrów z 3 października 2011 r. w sprawie rodzajów kartograficznych opracowań tematycznych i specjalnych. *Dz.U. z 2011 r., nr 222, poz. 1328.*
- Walesiak, M. (1996). Metody analizy danych marketingowych. PWN, Warszawa.
- Ustawa z dnia 21 sierpnia 1997 r. o gospodarce nieruchomościami. *Dz.U. z 2015 r. poz. 1774 i zm.*

THE SELECTIVE METHODS OF DIAGNOSTIC FEATURES IN DETERMINING THE REAL ESTATE FIELD OF VALUE AND MODELING VALUE OF PROPERTY

Summary. The implementation of an efficient market economy requires knowledge of the value of the property, giving rise to make a variety of decisions regarding the management and property management. Knowledge of the value of property required many of purposes, which follow the laws relating to the management of real estate owned by the public. One of the types is the value calculated in the mass valuation process, cadastral valuation, entered legally in order to reform the system of property taxes, determine the value of the public property, as well as office duties. The efficiency of the conduct of mass valuation and the accuracy of the results indicate the need of research. That would be the determination of diagnostic characteristics for homogeneous fields of value, modeling of mass property valuation and market characteristics influencing on cadastral value, methods of their selection and assortment. The study suggests the possibility of adaptation the statistical methods to the selection of diagnostic characteristics influencing on the range of real estates' valuable fields, as well as ways of determining their percentage in creating value of properties.

Key words: land management, cadastral value, mass valuation, real estate field of value

Zaakceptowano do druku – Accepted for print: 15.12.2015

Do cytowania - For citation:

Kuryj, J., Kuryj-Wysocka, O. (2015). Metody doboru cech diagnostycznych do wyznaczenia pól cenności i modelowania wartości. *Acta Sci. Pol. Administratio Locorum* 14(4), 49–64.