

ANALIZA ZALEŻNOŚCI POMIĘDZY CENĄ A LOKALIZACJĄ NIERUCHOMOŚCI NA PRZYKŁADZIE KRAKOWA*

ANALYSIS OF THE RELATIONSHIP BETWEEN REAL ESTATE PRICE AND LOCATION: THE CASE OF KRAKÓW

**Joanna Branna¹, Katarzyna Madej¹, Michał Będkowski¹, Michał Serdeń¹,
Patryk Sosiński¹, Małgorzata Luc²**

¹UNIGIS, Uniwersytet Jagielloński

²Zakład SIGKiT, Instytut Geografii i Gospodarki Przestrzennej, Uniwersytet Jagielloński

**Słowa kluczowe: lokalizacja nieruchomości, GIS, regresja liniowa, regresja geograficznie
ważona**

Keywords: real estate location, GIS, geographically weighted regression, linear regression

Wprowadzenie

Rynek nieruchomości stanowi dziedzinę gospodarczą i naukową, która dotychczas rozwijała się całkowicie niezależnie od rozwoju technologii informatycznych. Okazuje się jednak, że ze względu na charakter przestrzenny i jednocześnie dużą aplikacyjność badania tego rynku znalazły one swoje miejsce w geomarketingu, a do jego analizy można z powodzeniem wykorzystywać metody z zakresu GIS&T. Świadczą o tym między innymi polskie publikacje Bieleckiej i współautorów (2011a, 2011b, 2011c) oraz Cichocińskiego (2006, 2007, 2009). Pokazują one różne aspekty zastosowania narzędzi geoinformatycznych w zarządzaniu nieruchomościami, od budowania reguł topologicznych czy dostępności komunikacyjnej, po wykorzystanie metod geostatystycznych oraz budowę schematu aplikacyjnego. Schemat w postaci serwisu internetowego również proponuje Fan Yao-Min i inni (2009), aby zwiększyć efektywność wyszukiwania informacji o nieruchomościach.

Pomimo tak szerokiego traktowania zagadnienia rynku nieruchomości, w literaturze polskiej nie można znaleźć przykładu podjęcia tematu zależności cen nieruchomości od ich lokalizacji, a na podstawie opublikowanych badań rynku (Raport, 2008) okazuje się, że spo-

* Artykuł powstał w ramach przedmiotu *Projekt grupowy* na studiach podyplomowych UNIGIS na Uniwersytecie Jagiellońskim.

śród czynników decydujących o wyborze mieszkania równoległe z ceną pojawia się właśnie jego lokalizacja. Lokalizacja rozumiana jest tutaj jako grupa czynników przestrzennych determinująca ceny nieruchomości, które stanowią przedmiot sprzedaży w ostatnim roku. Takie podejście można natomiast znaleźć w literaturze zagranicznej. Ciekawe zestawienie czynników prezentują Ismail i Buyong (1998). Okazuje się, że można wyznaczyć ich od kilku bardzo podstawowych do 21 szczegółowych i w efekcie utworzyć system służący przede wszystkim do aktualizacji danych, wyszukiwania i wizualizacji obiektów według atrybutów ich lokalizacji, do analizy podobieństw oraz wyliczania najwyższej możliwej wartości nieruchomości (ang. *open market value*). Przegląd możliwości wykorzystania GIS w analizach rynku nieruchomości został zamieszczony w pracy Zantuck (2004), w której autorka koncentruje się na aktualizacji bazy danych oraz wyborze obiektów według zadanych cech.

Przyjęte założenia

W przeprowadzonych badaniach skoncentrowano się na przeanalizowaniu zależności, jakie zachodzą pomiędzy cenami nieruchomości w gminie miejskiej Kraków, a ich przestrzennym rozmieszczeniem. Wybór obszaru zdeterminowała dostępność danych, ale również zaobserwowane przez autorów w ciągu ostatnich lat ciekawe procesy zachodzące na tym obszarze w analizowanym zakresie. Należy jednak zwrócić uwagę na fakt, że ponieważ założono użyteczność i uniwersalność metody badawczej, istnieje możliwość zastosowania jej na dowolnym obszarze, oczywiście pod warunkiem dostępności odpowiednich danych. Ze względów użytecznych uznano, że projekt musi zostać przygotowany w sposób stosunkowo zrozumiały dla przeciętnego, niewykwalifikowanego w zakresie GIS użytkownika oraz łatwy do wdrożenia bez wykorzystania specjalistycznych narzędzi sprzętowych i programowych. Ważny element stanowi także przejrzysta, interesująca wizualizacja wyników. Wydaje się, że w podobnych pracach należy dużą wagę nadać skalowalności rozwiązania, projekt bowiem musi się łatwo rozbudowywać przez dostosowanie analizy do nowych obszarów oraz zwiększenie liczby analizowanych czynników.

Projekt badawczy ma charakter o tyle nowatorski, że niewiele można znaleźć publikacji świadczących o podejmowaniu prób zmierzających do określenia najlepszej metody wyznaczenia czynników przestrzennych wpływających na ceny nieruchomości, z drugiej zaś strony ukazuje potencjał zastosowania podobnego projektu w tworzeniu nowatorskiego na polskim rynku portalu rynku nieruchomości. Obecnie bowiem nie jest możliwe wyszukanie w internecie ofert sprzedaży nieruchomości w odniesieniu do innych obiektów przestrzennych, które odgrywają znaczną rolę w handlu nieruchomościami, to jest obiektów handlowych i usługowych, placówek oświatowych czy służby zdrowia, jak również dogodnej komunikacji.

Dane wykorzystane w analizie

Wykonalność analiz zależności między czynnikami lokalizacyjnymi a ceną nieruchomości zależy od dostępności danych wejściowych, na co zwracają uwagę też autorzy wspomnianych wyżej publikacji. Szczególnie istotnym wydaje się konieczność dysponowania możliwie jak najbardziej obszerną bazą danych nieruchomości. Dane brano pod uwagę w analizie dotyczyły charakterystyk nieruchomości oraz ich lokalizacji w stosunku do wybranych obiek-

tów użyteczności publicznej i odległości od centrum miasta. Do głównych źródeł danych należy zaliczyć:

- Bazę Danych Obiektów Topograficznych (BDOT) pozyskaną z Wojewódzkiego Ośrodka Dokumentacji Geodezyjnej i Kartograficznej w Krakowie, w której warstwy zawierają następujące informacje: punkty adresowe; ulice; tory kolejowe i tramwajowe; jezdnie; obszary wód stojących i płynących; budynki z informacjami o ich funkcji; granice administracyjne miasta; granicę centrum miasta;
- bazę ofert nieruchomości zebraną z serwisów internetowych zawierających ogłoszenia o nieruchomościach (z dokładnością przestrzenną ograniczoną do nazwy ulicy), która składała się z 743 obiektów; zostały one wyznaczone spośród 260 tys. mieszkań ze spisu z 2002 r., dodatkowo założono, że średnio w latach 2008-2011 oddawano do użytku 6700 mieszkań rocznie (BDL, GUS), co zwiększa liczebność do ok. 320 tys. (przy przyjętych założeniach: poziom ufności 0,95, błąd statystyczny 0,036, wielkość frakcji 0,5); oraz ich atrybutów: dzielnica, ulica, cena za metr kwadratowy i całkowita, stan mieszkania (deweloperski/wykończony), powierzchnia, liczba pokoi, kuchnia (aneks kuchenny/osobne pomieszczenie), piętro, rok budowy, liczba pięter w budynku, winda (występuje/nie występuje), preferencje dzielnicy (wg ankiety: Home Broker, 2010). Zostały podjęte starania o Bazę Rynku Nieruchomości, ale ze względu na poufność danych nie została ona nam udostępniona;
- OpenStreetMap (OSM) – cmentarze, parki i skwery (skorygowane na podstawie ortofotomapy i mapy topograficznej w skali 1:10 000), ciekі;
- bazę danych kodów pocztowych oraz miejscowości i ulic;
- mapę topograficzną w skali 1:25 000 – obszar Nowej Huty;
- Krajowy System Obszarów Chronionych – obszary chronione.

Nie udało się pozyskać pełnych i wiarygodnych danych dla przystanków komunikacji publicznej. Podjęto więc próbę wykorzystania OSM, jednak dane uzyskane w ten sposób były niepełne. Komunikację publiczną wyłączono ostatecznie z opracowania. Stwierdzono także brak rezerwatu przyrody „Skałki Przegorzalskie”, co zostało uzupełnione. Natomiast w przypadku budynku, który spełniał kilka funkcji, np. sklepy w bloku na parterze, wówczas jako funkcję szczegółową ustalono budynek wielorodzinny, a informacja o sklepie załączono w polu „informacja dodatkowa”. Wykonano także obliczenia odległości od każdej z ofert, które poddano procesowi geokodowania do najbliższego z obiektów przestrzennych zawartych w tych warstwach.

Zmienne brane pod uwagę w analizach dotyczyły charakterystyk nieruchomości, tj.: roku budowy, liczby kondygnacji, powierzchni, dostępności windy oraz lokalizacji konkretnych nieruchomości w stosunku do centrum miasta, placówek użyteczności publicznej, dostępu do komunikacji miejskiej i atrakcyjności otoczenia. Integracja danych obejmowała dwa etapy:

- 1) geokodowanie bazy ofert (poprzez odpowiednie przekształcenie warstwy odcinki jezdni) z tabelą ulice oraz bazą kodów pocztowych;
- 2) przygotowanie warstw oraz ich atrybutów niezbędne do prawidłowego obliczenia regresji (*Geographically Weighted Regression* – GWR), a następnie konwersję do odpowiedniego formatu i układu współrzędnych.

Wykaz nazw ulic z GUS przekodowano do standardu z bazy TERYT, a następnie zaznaczono na mapie ich lokalizację. Proces geokodowania podzielono na trzy etapy. Etap pierwszy to geokodowanie automatyczne czyli połączenie bazy ofert z warstwą ulic według ich nazwy, w wyniku czego otrzymano warstwę punktową ofert rozmieszczonych w środkach

odcinków ulic. W ten sposób wiele poddanych geokodowaniu ofert nakładało się na siebie, a ich lokalizacja nie nawiązywała do zabudowy wielorodzinnej w mieście. W drugim etapie przeprowadzono korektę ręczną. Wokół budynków wielorodzinnych utworzono bufory o wielkości 100 m w taki sposób, aby w obrębie strefy punktu znajdowało się możliwie najwięcej budynków wielorodzinnych położonych przy ulicy o tej samej nazwie, która występowała w ofercie. Korekta polegała na przesunięciu punktów w miejsca zabudowy wielorodzinnej. Trzeci etap wiązał się z weryfikacją i określeniem dokładności. Na podstawie kilkunastu ofert, dla których udało się ustalić adres nieruchomości (adres był podany, lub na podstawie zdjęć w ofertach) zweryfikowano proces geokodowania i ustalono ich dokładność do 400 m. Ostatecznie przygotowano mapy prezentujące lokalizacje wybranych obiektów infrastruktury miejskiej oraz dokonano wizualnej oceny wpływu lokalizacji na ceny (rys. 1).

Analiza przestrzenna i statystyczna

Jak już wspomniano, lokalizację uważa się za jeden z najważniejszych czynników wpływających na cenę za metr kwadratowy nieruchomości. Należy jednak zwrócić uwagę na wysoki współczynnik subiektywizmu, jakim jest obarczone takie podejście. Dobra lokalizacja, bowiem, to pojęcie względne, którego ocena zależy od położenia w stosunku do obiektów przestrzennych uznanych za „dobre” czy też korzystne. Biorąc pod uwagę powyższe oraz treść powszechnie znanego pierwszego prawa Toblera: *wszystko [w przestrzeni] jest związane ze wszystkim innym, przy czym bliższe rzeczy są bardziej związane niż rzeczy odległe* (Miler, 2004), przyjęto założenie, że istnieją czynniki przestrzenne, które wpływają na cenę nieruchomości mocniej lub słabiej. Wpływ na kształtowanie się końcowej ceny nieruchomości przeanalizowano w dwóch grupach elementów:

- 1) charakterystyka mieszkania
 - liczba pokoi w mieszkaniu
 - powierzchnia mieszkania
 - piętro, na którym mieszkanie jest położone
 - liczba kondygnacji budynku, w którym mieszkanie jest zlokalizowane
 - rok budowy nieruchomości
 - występowanie windy (tak/nie)
 - rodzaj kuchni (aneks kuchenny, osobna kuchnia)
 - stan danej nieruchomości (wykończony, developerski)
 - preferencje danej dzielnicy wg ankiety
- 2) położenie danej nieruchomości w stosunku do (dane odległościowe)
 - centrum
 - obiektów handlowych (centra handlowe, sklepy spożywcze, supermarkety-dyskonty, targowiska)
 - placówek oświatowych (żłobki, przedszkola, szkoły podstawowe)
 - placówek służby zdrowia (przychodnie zdrowia, szpitale)
 - obszarów przemysłowych (w tym huty ArcelorMittal)
 - istotnych linii komunikacyjnych (ruchliwe ulice, czynne linie kolejowe)
 - terenów zielonych (parki-skwerki, obszary chronione, wody powierzchniowe)
 - cmentarzy.

Wpływ czynnika przestrzennego na cenę nieruchomości zbadano posługując się analizą wieloczynnikową przy użyciu regresji liniowej metodą najmniejszych kwadratów (OLS) oraz regresją ważoną geograficznie (GWR) (Kluczycki, Ligas, 2007; Charlton, Fotheringham, 2009). Założono, że lepsze dopasowanie modelu dla tego typu danych uzyska się przy wykorzystaniu modelu regresji GWR. Ponadto, z uwagi na to, że dysponowano dwoma rodzajami danych, odległością i opisem nieruchomości, zdecydowano się na porównanie regresji z wyłączonymi/włączonymi zmiennymi odległościowymi. Stąd drugie założenie badań mówiące, że po porównaniu parametrów obu regresji dla modelu z danymi odległościowymi i bez nich, lepsze parametry uzyskuje regresja zbudowana na modelu wraz z tego typu danymi. Model OLS stosuje się do oszacowania najlepszego liniowego związku pomiędzy zmiennymi. Natomiast model GWR zakłada, że parametry można oszacować w dowolnym miejscu obszaru badań, a lokalizacja zmiennych zależnych i niezależnych jest znana. Założono też, że obiekty, które znajdują się bliżej tej lokalizacji przyjmują większą wagę w obliczeniach niż te, które znajdują się w dalszej odległości. Wyniki analizy zestawiono z regresją liniową, która pomijała zależności regresji ważonej geograficznie. Mają one potwierdzić założenie o występowaniu wpływu czynników przestrzennych na kształtowanie się cen nieruchomości oraz określić, które z nich odgrywają najistotniejszą rolę.

Analizy wykonano w czterech podstawowych etapach:

- 1) korelacja 28 zmiennych zależnych (9 opisujących nieruchomości wystawione do sprzedaży, 19 zmiennych przestrzennych tj. odległości obliczonych przy użyciu oprogramowania GIS) z ceną na podstawie współczynnika korelacji Pearsona oraz korelacja zmiennych niezależnych w celu zidentyfikowania ewentualnych współliniowości; dla każdej ze zmiennych wykonano autokorelację przestrzenną Moran I i zidentyfikowano wskaźnik Morana;
- 2) testowanie i budowanie modeli dla regresji OLS oraz GWR w celu utworzenia par regresji z tym samym zestawem zmiennych, w wyniku czego powstały następujące trzy pary regresji OLS i GWR oraz jedna regresja OLS:
 - I. regresje „charakterystyka mieszkań” – z wykorzystaniem zmiennych opisujących nieruchomości bez zmiennych odległościowych,
 - II. regresje „z Centrum” – z wykorzystaniem zmiennych opisujących nieruchomości z odległością do centrum jako kluczową zmienną przestrzenną,
 - III. regresje „przestrzenna” – z wykorzystaniem zmiennych opisujących nieruchomości z wyselekcjonowanymi zmiennymi przestrzennymi,
 - IV. regresja OLS „max” – z wykorzystaniem zmiennych opisujących nieruchomości z wyselekcjonowanymi zmiennymi przestrzennymi oraz zmiennymi odrzuconymi przez wcześniejsze GWR (przedstawia wykorzystanie 14 czynników istotnych statystycznie);
- 3) zestawienie parametrów opisujących powstałe modele, do porównania których wykorzystano parametry opisujące zmienne w modelu oraz opisujące sam model, przy czym w pierwszym przypadku mamy do czynienia ze współczynnikiem B oraz standaryzowanym Beta (dla OLS), następnie wśród wskaźników opisujących model wybrano indeksy: współczynniki determinacji R² i skorygowane R², istotność statystyczną testu F, Centrum Informacyjne Akaike’go AIC, wskaźnik przestrzennej autokorelacji Moran I (dla reszt regresji OLS);
- 4) wykonanie 7 rastrowych map resztowych (2 przykładowe zostały zaprezentowane na rysunkach 2 i 3) na podstawie wytworzonych reszt standaryzowanych GWR i OLS przy zastosowaniu interpolacji IDW, Power 3.

Na potrzeby utworzenia obu modeli regresji, dane pozbawiono wartości ekstremalnych oraz zbadano istotność, wpływ i dopasowanie poszczególnych zmiennych. Współczynniki opisujące oba modele porównano ze sobą w celu ustanowienia wpływu czynnika przestrzennego na cenę mieszkań. Przeprowadzono analizę przestrzenną i prezentację reszt dla obu modeli regresji. Badanie to miało potwierdzić założenie o występowaniu wpływu czynnika przestrzennego na kształtowanie się ceny nieruchomości dla rynku wtórnego (związanego ze wspomnianym prawem Toblera).

Wyniki

Analizę związku zmiennych niezależnych z ceną przeprowadzono przy pomocy wskaźnika korelacji Pearsona (tab. 1).

Wśród zmiennych związanych z charakterystyką mieszkania największą ujemną korelację ze zmienną zależną tj. ceną za m² wykazują zmienne "liczba pokoi" i „powierzchnia”. Wartości współczynnika korelacji (r) tych zmiennych są zbliżone do siebie, co oznacza, że ich wpływ na cenę mieszkania jest podobny. Znacznie mniejszy wpływ na cenę (określony ujemną wartością r) ma liczba pięter znajdujących się w budynku i piętro, na którym znajduje się mieszkanie, przy czym ten pierwszy charakteryzuje nieco większy wpływ. Bardzo niewielkie znaczenie (określone niską ujemną wartością r) dla ceny mieszkania ma również rok

Tabela 1. Współczynnik korelacji Pearsona

Wskaźniki	Współ. Pearsona do zmiennej "cena"	Wskaźniki	Współ. Pearsona do zmiennej "cena"
Powierzchnia	-0,267**	Szkoły podstawowe	-0,012
Liczba pokoi	-0,281**	Szpital	-0,157**
Piętro	-0,056	Targowisko	-0,026
Liczba pięter	-0,077*	Żłobki	0,002
Rok budowy	-0,035	Cmentarz	-0,233**
Winda	-0,002	Centrum	-0,473**
Stan	0,096*	Przemysł	-0,014
Preferecje	0,183**	Nowa Huta	0,272**
Centra handlowe	0,005	Ulice ruchliwe	0,021
Gimnazja	-0,059	Kolej czynna	0,001***
Przedszkola	0,019	Obszary chronione	-0,334**
Przychodnie zdrowia	-0,103**	Wody powierzchniowe	0,135**
Sklepy spożywcze	-0,082*	Parki skwerki	-0,103**
Supermarkety-dyskonty	0,071		

*korelacja jest istotna na poziomie istotności 0,05; **korelacja jest istotna na poziomie istotności 0,01; ***korelacja nieistotna statystycznie

budowy. Z analizy wartości wskaźnika korelacji wynika także, że w zakresie charakterystyki mieszkania większy wpływ na jego cenę ma jego „stan” niż wiek.

Analizując znaczenie zmiennych odległościowych w świetle wskaźników korelacji, należy zaznaczyć, że cena mieszkania (a konkretnie jej spadek) zależy w znacznym stopniu od odległości od centrum, którego wartość (r) jest największa ze wszystkich badanych zmiennych. Duży wpływ na spadek ceny ma także odległość od obszarów chronionych, szpitali, przychodni zdrowia i, co ciekawe, od cmentarzy. Cena mieszkania rośnie natomiast najbardziej wraz ze wzrostem odległości od Nowej Huty, wód powierzchniowych i supermarketów.

Do porównania siły oddziaływania poszczególnych zmiennych niezależnych na zmienną zależną wykorzystano standaryzowany współczynnik Beta (generowany tylko dla regresji OLS). Najbardziej stabilną zmienną okazuje się „powierzchnia”. Występuje ona w obu regresjach i oddziałuje na model z podobną siłą. Ponadto współczynnik Beta standaryzowany jest w sposób zbliżony do współczynnika korelacji, co może świadczyć o dość niewielkiej szansie na skorelowanie tego współczynnika z innymi zmiennymi. Podobną sytuację wykazuje także odległość od centrum, choć tutaj różnice wskaźnika Beta we wszystkich regresjach są nieco bardziej zróżnicowane. Pozostałe zmienne wykazują mniejsze lub większe odchylenia, co może świadczyć o współliniowości, jednak ze względu na istotny wpływ na wyniki regresji pozostają one w modelu. Współczynnik B stanowi parametr równania regresji dla poszczególnych zmiennych, a jego interpretację zaprezentowano w tabeli 2.

Natomiast współczynnik determinacji R^2 odzwierciedla dopasowanie funkcji liniowej do danych rzeczywistych, występuje na poziomie od 25% i osiąga maksymalną wartość w ostatniej regresji tj. 48% danych. Różnica pomiędzy współczynnikiem R^2 a skorygowanym R^2 świadczy o reprezentatywności modelu. Regresję OLS charakteryzuje więc większa reprezentatywność aniżeli GWR. AIC bierze pod uwagę złożoność modelu, przy określonej liczbie zastosowanych zmiennych, czyli im niższa wartość AIC tym lepszy model. AIC w parach regresyjnych uzyskało lepsze rezultaty dla regresji GWR, jednakże to regresja OLS Max (nieposiadająca odpowiednika GWR) prezentuje najniższe wartości tego współczynnika, co tłumaczy największą liczbę zmiennych zastosowanych w tym modelu. Wskaźnik przestrzennej autokorelacji Morana wykonano tylko dla reszt regresji OLS, a jego wysokie wartości (bliskie 1) mogą świadczyć o pominięciu znaczących zmiennych kluczowych (ESRI forum). Potwierdzają to regresje tutaj prezentowane, dla których wskaźnik Morana maleje wraz ze zwiększeniem się liczby zmiennych wyjaśniających. Interesujące zależności można także zauważyć porównując różnice współczynników R^2 , AIC w poszczególnych parach regresji OLS i GWR. W pierwszej regresji oprócz gorszych rezultatów wartości samych współczynników widoczne są wyraźnie różnice między tymi współczynnikami. Wraz z dodaniem do kolejnych modeli zmiennych przestrzennych różnice maleją. Można więc wyciągnąć z tego wniosek, że wraz z włączeniem do modelu większej liczby zmiennych, różnice pomiędzy dopasowaniem modeli GWR i OLS maleją. Zależności te prezentuje tabela 3.

Tabela 2. Interpretacja współczynnika B dla różnych modeli regresji

Regresja	Czynniki decydujące o spadku/wzroście ceny	Spadek/wzrost ceny za m ² [PLN]	
		spadek	wzrost
Regresja OLS "Charakterystyka mieszkań"	wzrost powierzchni nieruchomości o 1 m ²	15	
	podwyższenie budynku o jedno piętro	132	
	występowanie w budynku windy		640
	występowanie widnej kuchni	922	
	fakt wykończonego mieszkania		776
	zwiększenie preferencji wyboru dzielnicy o 1%		24
Regresja OLS "z centrum"	wzrost powierzchni nieruchomości o 1 m ²	15	
	podwyższenie budynku o jedno piętro	140	
	występowanie w budynku windy		579
	występowanie widnej kuchni	589	
	fakt wykończonego mieszkania		548
	zwiększenie odległości od centrum o 100 m	25	
Regresja OLS "przestrzenna"	wzrost powierzchni nieruchomości o 1 m ²	15,5	
	podwyższenie budynku o jedno piętro	110	
	występowanie w budynku windy		540
	występowanie widnej kuchni	555	
	fakt wykończonego mieszkania		603
	zwiększenie odległości od centrum handlowego o 100 m	33	
	zwiększenie odległości od cmentarza o 100 m		7
	zwiększenie odległości od centrum o 100 m	27,5	
Regresja OLS "MAX"	wzrost powierzchni nieruchomości o 1 m ²	16	
	podwyższenie budynku o jedno piętro	111	
	zwiększenie roku budowy budynku o 1 rok		15
	występowanie w budynku windy		425
	występowanie widnej kuchni	570	
	fakt wykończonego mieszkania		774
	zwiększenie odległości od centrum handlowego o 100 m		47
	zwiększenie odległości od gimnazjum o 100 m	15	
	zwiększenie odległości od cmentarza o 100 m		7
	zwiększenie odległości od centrum o 100 m	36,5	
	zwiększenie odległości od czynnej kolei o 100 m	9	
	zwiększenie odległości od obszarów chronionych o 100 m	9	
zwiększenie odległości od wód powierzchniowych o 100 m		23	
	zwiększenie odległości od parków i skwerów o 100 m	20	

Tabela 3. Porównanie współczynników: R2 (dopasowanie funkcji liniowej do danych rzeczywistych), Skorygowane R2 (dopasowanie do innej próby z tej samej populacji), AIC (kryterium wyboru pomiędzy modelami statystycznymi o różnej liczbie predyktorów) i Moran I (określający skoncentrowanie przestrzennie obiektów) dla poszczególnych modeli regresji

Współczynnik	"Charakterystyka mieszkań"		"Z centrum"		"Przestrzenna"		OLSmax
	OLS	GWR	OLS	GWR	OLS	GWR	OLS
R2	0,258	0,413	0,369	0,452	0,429	0,453	0,490
Skorygowane R2	0,251	0,384	0,363	0,422	0,422	0,431	0,480
AIC	12074,19	11947,95	11960,66	11905,83	11894,07	11890,39	11823,01
Moran I	0,154	0	0,082	0	0,039	0	0,021

Analiza map reszt

Mapy reszt przedstawiają rozmieszczenie różnic pomiędzy obserwowaną wartością zmiennej zależnej i wartością przewidywaną przez model regresji w wersji standaryzowanej. Zawarte są w nich wszystkie te informacje, które nie zostały ujęte w regresjach. Przeglądając powstałe w ten sposób mapy należy zwrócić uwagę na następujące zależności:

- zaznacza się lokalność regresji GWR – mapy reszt GWR dla tych samych zmiennych w mniejszym stopniu przejawiają radykalne wartości i bardziej dążą w kierunku wartości zerowej;
- widoczny jest wpływ punktów, na bazie których wykonano interpolację – większa liczba punktów umożliwiłaby bardziej dogłębną analizę, mniejsza liczba obserwacji na obrzeżach w porównaniu z obszarami centralnymi miasta także zniekształca obraz;
- duża różnica w wartościach reszt pojawia się po włączeniu do równania regresji danych o odległości z centrum, szczególnie w rejonie Nowej Huty;
- dodanie zmiennej „obszary chronione” do regresji przestrzennych mogło wpłynąć na wyłączenie wartości ekstremalnych reszt z obrzeży miasta.

Na rysunkach 2 i 3 zaprezentowano przykładowe wyniki poszczególnych regresji w postaci mapy reszt.

Potencjał wykorzystania zaproponowanego rozwiązania

Wydaje się, że wyniki zaproponowanej metody, ze względu na swoją specyfikę, mogą znaleźć wielu odbiorców końcowych w obszarach związanych z rynkiem nieruchomości. Do tego grona zaliczyć można następujące grupy:

- nabywcy lub sprzedający nieruchomości – metoda umożliwia zdobycie informacji o kształtowaniu się cen nieruchomości w zależności od czynników przestrzennych;
- deweloperzy i inwestorzy – metoda może wspomagać analizę kosztów oraz korzyści dla procesu wyszukiwania nowych miejsc pod inwestycje mieszkaniowe, jak również, gdy jest przeprowadzana regularnie, wskazać trendy zmian;
- rzeczoznawcy majątkowi – obecnie lista czynników uwzględnianych przez rzeczoznawców jest ograniczona;

- banki – wspomaganie w procesie analizy ryzyka przyznawanego kredytu hipotecznego za pomocą weryfikacji wartości obszaru kredytowanej inwestycji;
- urzędy statystyczne – analiza wykonywana w sposób regularny umożliwia określenie trendów zmian cen nieruchomości w zależności od czynników lokalizacyjnych;
- biura nieruchomości – metoda umożliwia wyszukanie lokalizacji odpowiadającej specyficznym wymaganiom klienta.

Podsumowanie

Na podstawie przeprowadzonej analizy poszczególnych współczynników modelu zostały wyciągnięte wnioski zaprezentowane graficznie w tabeli 4.

Okazuje się, że regresja GWR poprawia dopasowanie modelu do wartości obserwowanych określane przez parametr R kwadrat oraz wiąże się ze złożonością modelu charakteryzowaną za pomocą parametru AIC. Z kolei regresja OLS wykazuje większą reprezentatywność wśród hipotetycznych modeli tworzonych na tej samej populacji, co wyraża parametr skorygowany R². Za większość różnic pomiędzy parametrami modeli GWR i OLS odpowiada lokalność modelu GWR. Szacując parametry modelu dla danej lokalizacji zakłada się, że wartości obiektów, które znajdują się bliżej tej lokalizacji powinny przyjąć większą wagę w obliczeniach niż te, które znajdują się w dalszej odległości (Charlton, Fotheringham, 2009). Ponieważ współczynniki R² oraz AIC (wyższe dla modeli GWR) identyfikują lepsze dopasowanie modelu, to można potwierdzić przyjęte na początku założenie, że lepsze dopasowanie modelu dla tego typu danych uzyska się przy wykorzystaniu modelu regresji GWR. Natomiast drugą hipotezę – mówiącą że po porównaniu parametrów regresji OLS i GWR dla modelu z danymi odległościowymi i bez nich, lepsze parametry uzyskuje regresja zbudowana na modelu wraz z tego typu danymi – potwierdzają zarówno wysokie współczynniki korelacji niektórych zmiennych przestrzennych (odległość od centrum), współczynniki standaryzowane BETA w regresjach OLS, jak również lepsze wyniki regresji ze zmiennymi odległościowymi.

Tabela 4. Cechy modeli w poszczególnych typach regresji

Cecha	Regresja OLS	Regresja GWR
Dopasowanie modelu		
Reprezentatywność		
Złożoność modelu		
Zastosowany model	OGÓLNY	LOKALNY

Literatura

- Bielecka E., Latos D., 2011a: Koncepcja systemu zarządzania nieruchomościami policji za pomocą GIS. *Studia i materiały Towarzystwa Naukowego Nieruchomości*, vol. 19, nr 2: 107-116.
- Bielecka E., Latos D., Latos M., Kamiński P., Szura P., 2011b: Wybrane elementy o charakterze przestrzennym schematu aplikacyjnego do zarządzania nieruchomościami. *Roczniki Geomatyki* t. 9, z. 2(46): 7-16, PTIP, Warszawa.
- Całka B., Bielecka E., 2011c: Analiza funkcjonalności wybranych programów wspierających proces zarządzania nieruchomościami i ich możliwość zastosowania w systemie policji. *Studia i materiały Towarzystwa Naukowego Nieruchomości*, vol. 19, nr 2: 71-82.
- Charlton M., Fotheringham S., 2009: Geographically Weighted Regression. White Paper, Maynooth, Ireland.
- Cichociński P., 2006: Modelowanie dostępności komunikacyjnej nieruchomości jako atrybutu niezbędnego w procesie wyceny. *Roczniki Geomatyki* t. 4, z. 3: 71-80, PTIP, Warszawa.
- Cichociński P., 2007: Zastosowanie zaawansowanych reguł topologicznych w procesie budowania baz danych przestrzennych wspomagających wyceną nieruchomości. *Roczniki Geomatyki* t. 5, z. 3: 29-37, PTIP, Warszawa.
- Cichociński P., 2009: Próba zastosowania metod geostatystycznych do taksacji nieruchomości. *Roczniki Geomatyki* t. 7, z. 4(34): 17-30 PTIP, Warszawa.
- ESRI Forum, Spatial Statistic, Geographically Weighed Regression GWR and Ordinary Least Squares OLS woes.
- Fang Yao-Min, Lin Li-Yu, Huang Chua-Huang, Chou Tien-Yin, 2009: An integrated information system for real estate agency-based on service-oriented architecture. *Expert Systems with Applications* 36, 11039-11044. HELP ArcGIS 10, Interpreting GWR Results.
- Home Broker, 2010: http://www.inwestycje.pl/resources/Attachment/2009/09_02/file4749.doc
- Ismail S., Buyong T., 1998: Residential property valuation using geographic information system. *Buletin Geoinformasi*, Jld. 2, No.2: 249-266.
- Kluczycki M., Ligas M., 2007: Regresja ważona geograficznie jako narzędzie analizy rynku nieruchomości. *Geomatics and Environmental Engineering*, vol. 1, No 2.
- Miler H.J, 2004: Tobler's First Law and Spatial Analysis. *Annals of the Association of American Geographers* vol. 94, no. 2: 284-289.
- Raport, 2008: Rynek mieszkaniowy w Warszawie, Analiza czynników wpływających na zakup mieszkania, CB Richard Ellis i Murator EXPO.
http://me-mieszkania.pl/pl/analizy_i_raporty/badania_ryнку/art7.html
- Zantuck J., 2004: Developing a Geodata base and GIS Software for a Real Estate Industry Application.
<http://ecite.utas.edu.au/32502> i <http://www.docstoc.com/docs/24403530/Developing-a-Geodatabase-and-GIS-Software-for-a-Real-Estate>

Abstract

The paper presents only one aspect of a broad issue of the real estate market analysis, that is the relationship between the price of the property and its location. The study used a geographically weighted regression. Parameters in this method can be estimated anywhere in the area of the research, and location of independent and dependent variables is known. The data taken into account in the analyses were related to the characteristics of the properties and their location in relation to certain public facilities as well as to the distance from the city center. The results have been weighted. When estimating the parameters established for a given location, it was assumed that the objects that are closer to this location should adopt a greater weight in the calculations than those that are further away. The analysis results were compared with the linear regression, which omitted those relationships of geographically weighted regression. The obtained results confirm the assumption of the existence of spatial factors that influence the property prices and determine which ones are the most important.

mgr Joanna Branna
joannagluza@gmail.com

mgr inż. Katarzyna Madej
katarzyna.madej@gmail.com

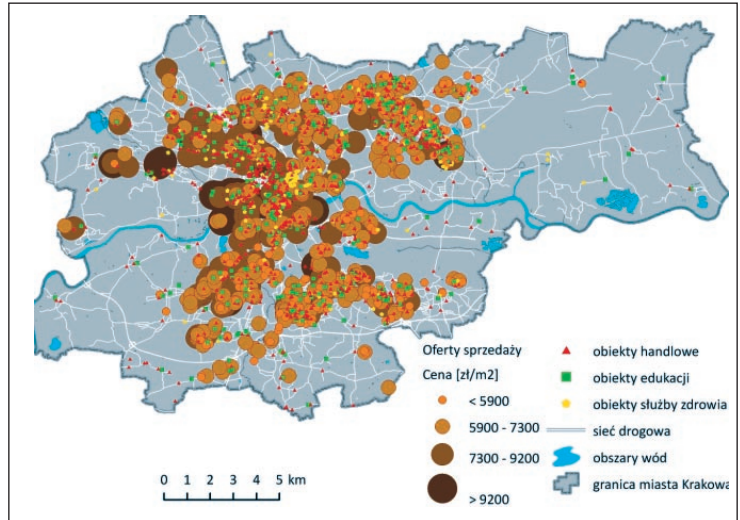
mgr Michał Będkowski
bedkowskim@gmail.com

mgr Michał Serdeń
m.serden@analizygis.pl

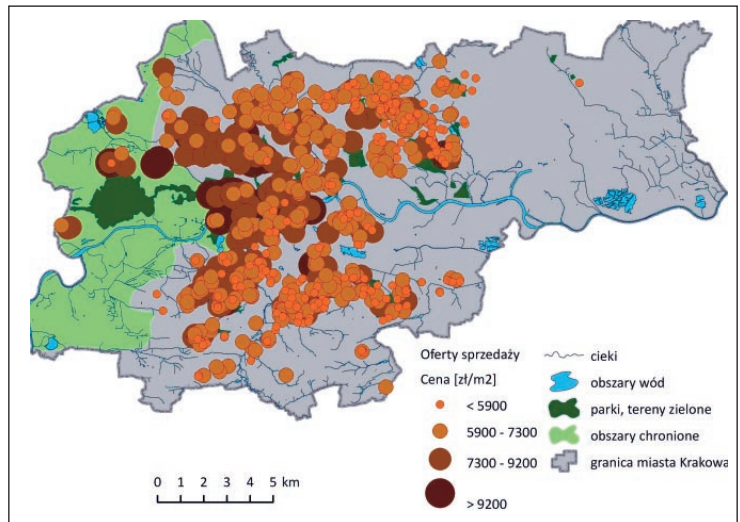
mgr Patryk Sosiński
patryk.sosinski@gmail.com

dr Małgorzata Luc
mluc@gis.geo.uj.edu.pl
tel. 12 664 53 22
<http://www.gis.geo.uj.edu.pl/ZGIS/onas/ml.html>

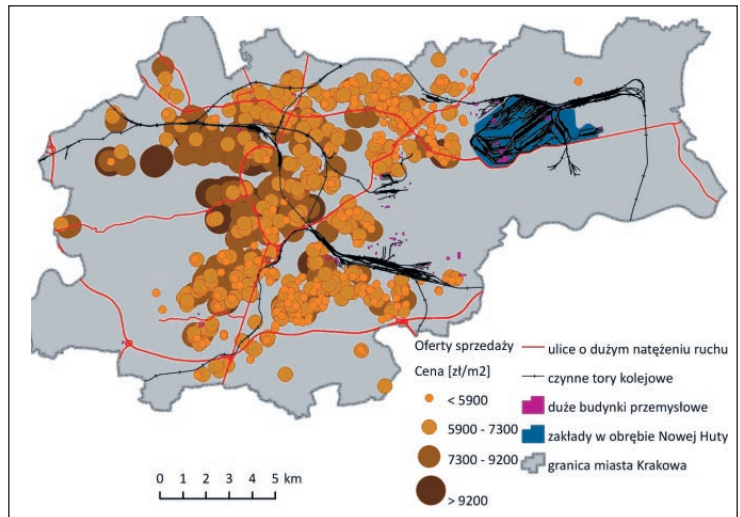
a



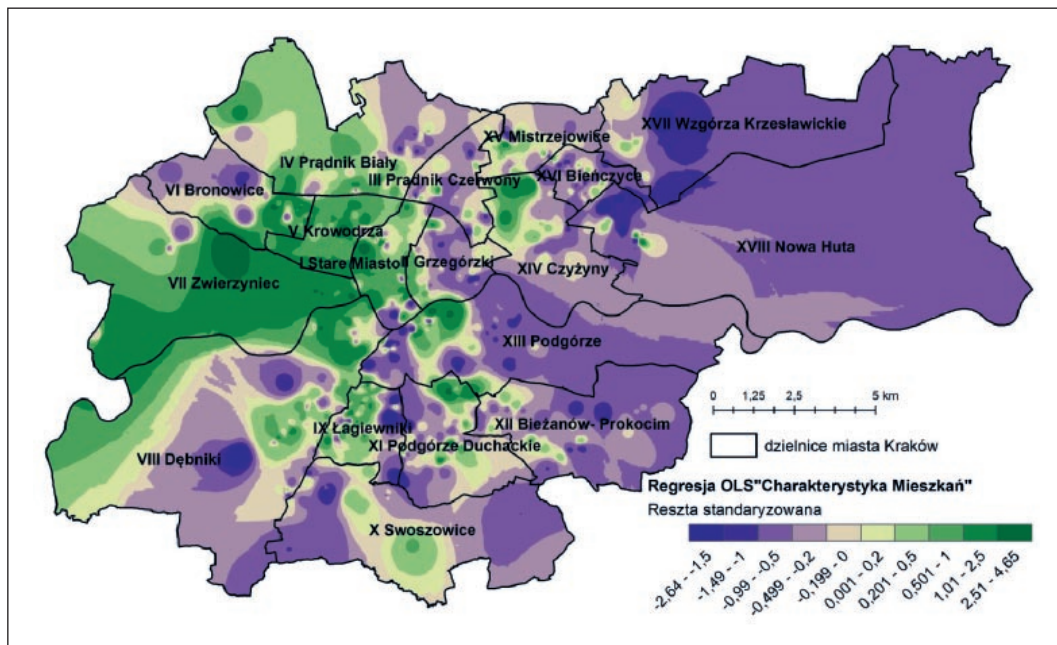
b



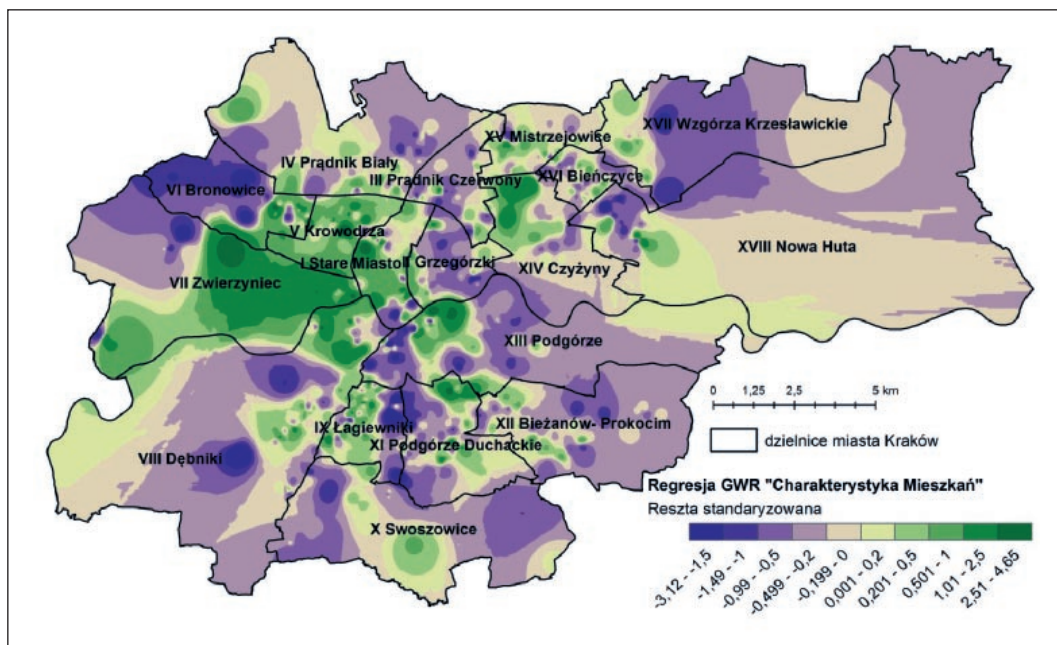
c



Rys. 1. Mapa rozmieszczenia ofert nieruchomości względem:
a – punktów usługowych i użyteczności publicznej,
b – obiektów przyrodniczych,
c – uciążliwego sąsiedztwa



Rys. 2. Wynik regresji OLS „Charakterystyka mieszkań”



Rys. 3. Wynik regresji GWR „Charakterystyka mieszkań”