

**ZASTOSOWANIE REGRESJI WAŻONEJ GEOGRAFICZNIE
DO MODELOWANIA MIEJSKIEJ WYSPY CIEPŁA WE WROCŁAWIU**

**APPLICATION OF GEOGRAPHICALLY WEIGHTED REGRESSION
FOR URBAN HEAT ISLAND MODELLING IN WROCŁAW**

Mariusz Szymanowski¹, Maciej Kryza²

¹ Zakład Kartografii,
Instytut Geografii i Rozwoju Regionalnego Uniwersytetu Wrocławskiego

² Zakład Klimatologii i Ochrony Atmosfery,
Instytut Geografii i Rozwoju Regionalnego Uniwersytetu Wrocławskiego

SŁOWA KLUCZOWE: regresja ważona geograficznie, interpolacja przestrzenna, GIS, miejska wyspa ciepła, Wrocław

STRESZCZENIE: Artykuł prezentuje zastosowanie regresji ważonej geograficznie (GWR) do interpolacji przestrzennej temperatury powietrza w warunkach wystąpienia miejskiej wyspy ciepła (UHI) we Wrocławiu. Rezultaty interpolacji GWR porównano z uzyskanymi za pomocą innych metod: regresji wieloczynnikowej (MLR) – model globalny, i krigingu resztowego dla obydwu modeli regresyjnych (odpowiednio: RK i GWRK). Dane wejściowe stanowiły pomiary temperatury powietrza w 206 nieregularnie rozmieszczonych punktach oraz warstwy dodatkowych zmiennych objaśniających, utworzone głównie na podstawie mapy pokrycia terenu i zdjęć satelitarnych (Landsat TM). Walidację interpolacji przeprowadzono metodą oceny krzyżowej (CV), analizując miary diagnostyczne (BIAS, MAE, RMSE) i rozkład przestrzenny błędów CV. Porównanie rezultatów modeli regresyjnych wskazało zasadność zastosowania GWR w przypadku niestacjonarnego procesu przestrzennego (UHI). Kalibrację modelu lokalnego przeprowadzono w zmiennej macierzy sąsiedztwa (tzw. kernelu) z uwzględnieniem zachowania możliwości fizycznej interpretacji procesu. Kriging reszt, przeprowadzony dla lokalnego i globalnego modelu regresji poprawił jakość interpolacji. Za optymalną w analizowanych przypadkach UHI metodę interpolacji uznano kriging resztowy dla regresji ważonej geograficznie (GWRK).

1. WSTĘP

Miejska wyspa ciepła (ang. *Urban Heat Island* – UHI) jest jedną z najsilniej zaznaczonych lokalnych, antropogenicznych modyfikacji klimatu. UHI powstaje w wyniku

oddziaływania czynników fizycznych związanych z przekształceniem środowiska naturalnego w procesach urbanizacyjnych, głównie poprzez zmiany właściwości radiacyjnych, termicznych i aerodynamicznych przestrzeni miejskiej (Oke, 1987). UHI wyraża się wzrostem temperatury postępującym w kierunku centralnych partii miast. Pole temperatury w takich przypadkach przyjmuje na ogół kształt koncentryczny, zaburzony przez lokalne obniżenia i podwyższenia temperatury związane ze zmianami w pokryciu i użytkowaniu terenu (Oke, 1976). Występowanie UHI jest szczególnie czytelne podczas tzw. pogody radiacyjnej, a więc przy bezchmurnej i bezwietrznej atmosferze. Centralne partie miast mogą być wtedy nawet o kilkanaście stopni cieplejsze w porównaniu do obszarów pozamiejskich, co ma istotny wpływ na warunki bioklimatyczne, społeczno-gospodarcze i ogólny komfort życia mieszkańców miast. Dane o natężeniu oraz przestrzennej strukturze UHI są ważnym składnikiem procedur planistycznych, a także stanowią kluczową informację wejściową dla modelowania innych zjawisk, np. dyspersji zanieczyszczeń atmosferycznych.

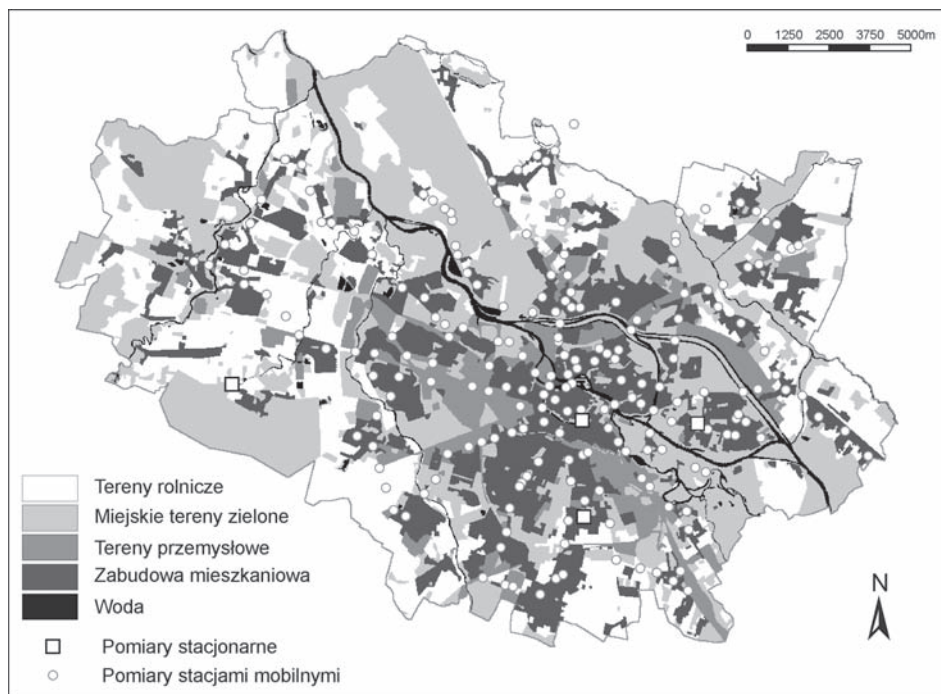
Przestrzenna interpolacja ma szczególne znaczenie w meteorologii i klimatologii, gdzie pole elementów jest podstawową informacją analityczną (Dobesch *et al.*, 2007; Tveito *et al.*, 2008). Algorytmy interpolacyjne zaimplementowane w środowisku GIS zapewniają powtarzalność procedur dla zmieniających się parametrów, a uniwersalne metody oceny jakości wyników pozwalają na wybór metody optymalnej dla każdego analizowanego przypadku. W rezultacie znacznej poprawie, w porównaniu do tradycyjnych, manualnych metod, ulega jakość końcowego produktu – mapy.

Dotychczasowe prace nad przestrzenną interpolacją UHI dla Wrocławia wskazywały na metody liniowej regresji wieloczynnikowej (ang. *multiple linear regression* – MLR) i jej rozszerzenia: krigingu resztowego (ang. *residual kriging* – RK) jako optymalnych dla interpolacji temperatury powietrza w mieście przy nieregularnym rozmieszczeniu danych wejściowych (Szymanowski, Kryza, 2008, 2009). Zastosowanie globalnego modelu regresji (MLR) w sytuacji niestacjonarności procesu przestrzennego, z jakim mamy do czynienia w przypadku wystąpienia UHI, budzi jednak pewne zastrzeżenia metodyczne. Potencjalnym sposobem rozwiązania tego problemu jest zastosowanie lokalnego modelu regresji, znanego jako regresja ważona geograficznie (ang. *geographically weighted regression* – GWR), ewentualnie wraz z jej rozszerzeniem o kriging reszt (GWRK).

Celem artykułu jest wybór optymalnej metody interpolacji przestrzennej oraz jej zastosowanie do uzyskania ciągłej przestrzennie informacji o polu temperatury powietrza w mieście dla wybranych przypadków UHI we Wrocławiu.

2. DANE

Wrocław (pow. 293 km²; ludn. ~640 tys.) jest położony w SW Polsce na wysokości około 120 m n.p.m. Zabudowa miasta jest rozmieszczona wzdłuż głównej rzeki – Odry. 31.4% powierzchni miasta zajmują tereny zabudowy mieszkaniowej i przemysłowej (Rys. 1). Na pozostałą część składają się miejskie tereny zielone (36.6%), rolnicze z nieużytkami (28.9%) oraz wody powierzchniowe (3.1%). Brak różnicowania rzeźby



Rys. 1. Mapa pokrycia terenu oraz rozmieszczenie punktów pomiaru temperatury powietrza we Wrocławiu.

terenu oraz stosunkowo czytelna struktura urbanistyczna sprawiają, iż Wrocław jest cennym obszarem badań i weryfikacji modeli relatywnie niezaburzonego klimatu miasta. Zjawisko UHI ma tu charakter typowy, co do intensywności i struktury przestrzennej, dla kilkuset tysięcy miast umiarkowanych szerokości geograficznych (Szymanowski, 2004, 2005).

2.1. Dane meteorologiczne

Dane meteorologiczne zostały zebrane podczas sesji pomiarowych przeprowadzonych w latach 2001-2002. Do pomiarów wykorzystano 2 mobilne stacje meteorologiczne, umożliwiające pomiar temperatury powietrza na standardowej wysokości 2 m n.p.g. Każda stacja miała w trakcie sesji do pokonania trasę o długości około 90 km, co przy sugerowanej prędkości podczas pomiaru 30 km h⁻¹ wymagało 3÷4 godzin na zgromadzenie kompletu materiału z wyznaczonych 206 punktów referencyjnych (Rys. 1). Sesje pomiarowe prowadzono w godzinach nocnych podczas fazy największej intensywności UHI. Zmiana temperatury w trakcie trwania sesji została skorygowana na podstawie rejestracji z 4 automatycznych stacji meteorologicznych, których rozmiesz-

czenie uwzględniało wpływ zróżnicowanego pokrycia terenu na przebieg temperatury. Przeprowadzono 7 sesji pomiarowych podczas sprzyjających rozwojowi UHI warunków pogodowych (bezchmurne lub słabo zachmurzone niebo, prędkość wiatru poza miastem nieprzekraczająca 4 m s^{-1}). Na potrzeby tego opracowania analizie poddano 2 przypadki wyspy ciepła obserwowanej w różnych porach roku (22.05.2001, 3.01.2002; Szymanowski, Kryza, 2009). Wybrane do analizy przypadki są reprezentatywne dla silnie zaznaczonej UHI o genezie radiacyjnej. W takich sytuacjach można obserwować zgodność położenia strefy maksymalnej intensywności UHI z centrum miasta (sytuacja bezwietrzna lub występowanie bryzy miejskiej ze zbieżnymi w kierunku centrum liniami prądów) lub też przesunięcie struktury UHI *en bloc* w kierunku zawietrznym (przy słabym, ale stabilnym co do kierunku wietrze regionalnym). Wybrane przypadki reprezentują te dwie sytuacje: pierwszą – 22.05.2001, drugą – 3.01.2002. Jednocześnie są to przypadki najintensywniej zaznaczonej UHI spośród zmierzonych podczas wszystkich sesji pomiarowych.

2.2. Dodatkowe zmienne objaśniające

Miejska wyspa ciepła powstaje w rezultacie kompleksowego oddziaływania szeregu czynników. Informacje przestrzenne o części parametrów środowiska miejskiego determinujących natężenie i pole UHI mogą zostać wykorzystane jako dodatkowe zmienne objaśniające w wielowymiarowych algorytmach interpolacyjnych, do których zaliczamy modele regresyjne. Warstwy środowiskowe zastosowane w opracowaniu zostały sporządzone na podstawie różnorodnego materiału wejściowego: cyfrowej mapy pokrycia terenu (16 klas), map topograficznych w skali 1:10 000, cyfrowej bazy budynków (dostępnej dla fragmentów miasta), ortofotomapy oraz zobrazowań satelitarnych (Landsat TM). W oparciu o te materiały sporządzono mapy rastrowe sześciu parametrów silnie skorelowanych z polem temperatury:

- współczynnika szorstkości – z_0 [m],
- udziału czynnych powierzchni sztucznych w danej klasie pokrycia terenu (z uwzględnieniem powierzchni ścian budynków) – AS [%],
- udziału powierzchni naturalnych w danej klasie pokrycia terenu – NS [%],
- indeksu wegetacji NDVI,
- admitancji termicznej – μ [$\text{J m}^{-2} \text{s}^{-1/2} \text{K}^{-1}$],
- emisji ciepła sztucznego – Q_A [W m^{-2}].

Ponieważ temperatura powietrza nie zmienia się skokowo na granicy klasy pokrycia terenu, ale strefa przejściowa sięga, w zależności od warunków meteorologicznych, od kilkudziesięciu do kilkuset metrów, wygenerowano warstwy pochodne z zastosowaniem funkcji sąsiedztwa (ang. *focalmean*), uśredniających informację w kołowym otoczeniu o zadanym promieniu ($50 \div 1000 \text{ m}$) od danego rastra. Szczegółowy opis metodyki pomiarowej oraz metodyki przygotowania warstw dodatkowych zmiennych objaśniających został zawarty we wcześniejszych publikacjach (Szymanowski, Kryza, 2008, 2009).

3. METODY

Poszukując optymalnej metody przestrzennej interpolacji dla analizowanych przypadków należy zwrócić uwagę na dwie charakterystyczne cechy rozmieszczenia meteorologicznych danych wejściowych. Po pierwsze, rozkład punktów pomiarowych jest nieregularny, o zmiennej przestrzennie gęstości. Po drugie, znaczne obszary na obrzeżach miasta są pozbawione pomiarów, co oznacza, iż poszukiwana metoda powinna dawać wiarygodne wyniki nie tylko jako interpolator (wewnątrz strefy pokrytej pomiarami), ale także jako ekstrapolator przestrzenny. We wcześniejszych pracach poddano ocenie szereg stosowanych, także w klimatologii, metod, jak: metoda ważonych odwrotnych odległości, funkcje sklepane, kriging zwyczajny oraz regresję wieloczynnikową i kriging resztowy (Szymanowski, Kryza, 2008, 2009). Najlepsze rezultaty otrzymano dla dwóch ostatnich algorytmów, w związku z czym będą one stanowić tło dla rozważań nad wynikami GWR i jej rozszerzenia przez kriging reszt. Dla odróżnienia od RK stosowanego dla globalnego modelu regresji (MLR), kriging resztowy dla modelu lokalnego zostanie oznaczony jako GWRK.

Pełna procedura poszukiwania optymalnego algorytmu interpolacyjnego dla analizowanych przypadków składa się z:

- 1) opracowania warstw potencjalnych predyktorów (zmiennych objaśniających) pola temperatury (rozdział 2.2),
- 2) doboru optymalnych zmiennych objaśniających metodą regresji krokowej (rozdział 3.2),
- 3) doboru optymalnych parametrów metod interpolacji (modelowanie wariogramu dla RK i GWRK - rozdział 3.3; dobór wielkości macierzy sąsiedztwa dla GWR – rozdział 4) oraz wykonania interpolacji,
- 4) walidacji rezultatów uniwersalną metodą oceny krzyżowej (rozdział 3.4).

W opracowaniu wykorzystano system GIS GRASS oraz pakiet statystyczny R z bibliotekami gstat i spgwr.

3.1. Regresja wieloczynnikowa – model globalny

Metoda regresji liniowej w przestrzennej interpolacji danych może być zaliczana zarówno do grupy metod probabilistycznych, jak i do deterministycznych. Jeśli regresję traktujemy jako metodę deterministyczną głównym wyznacznikiem specyfikacji funkcji jest możliwość wyjaśnienia procesu przestrzennego przyczynami fizycznymi. Przy podejściu stochastycznym podstawowym założeniem jest brak przestrzennej autokorelacji i normalność rozkładu reszt regresji (Tveit *et al.*, 2008). Do metod regresyjnych, oprócz regresji prostej, wykorzystującej jedną zmienną objaśniającą, zaliczamy MLR, która pozwala rozszerzyć zbiór predyktorów. Matematycznie model globalny może być wyrażony jako:

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \varepsilon_i \quad (1)$$

gdzie y oznacza zmienną interpolowaną, zmienne niezależne (predyktory) opisane są jako x , współczynniki regresji liniowej jako β , a ε oznacza reszty regresji.

Dla każdego interpolowanego przypadku UHI dodatkowe zmienne objaśniające były dobierane z całego zbioru potencjalnych predyktorów za pomocą regresji krokowej. Analizowano istotność statystyczną korelacji zmiennej niezależnej z temperaturą powietrza oraz usuwano zmienne nadmiarowe w wypadku wzajemnej korelacji predyktorów. Bardzo ważnym kryterium był kierunek zależności (dodatni lub ujemny współczynnik β), który przy założeniu fizycznej interpretowalności modelu powinien być zgodny z naturą oddziaływania czynnika środowiskowego.

3.2. Regresja ważona geograficznie – model lokalny

Jednym z podstawowych problemów metodycznych przy stosowaniu globalnego modelu regresji jest założenie stacjonarności procesu przestrzennego. Oznacza to, iż w każdej części analizowanego obszaru proces podlega takim samym uwarunkowaniom, a określone oddziaływanie czynników sprawczych przynosi takie same rezultaty. Bardzo trudno o spełnienie takiego założenia w klimatologii, gdzie z reguły mamy do czynienia z procesami niestacjonarnymi. Rozwiązaniem może być zastosowanie lokalnego modelu regresji, w którym przyjmuje się założenie niestacjonarności procesu przestrzennego. Taką metodą jest regresja ważona geograficznie (Fotheringham *et al.*, 2002; Lloyd, 2007).

GWR może być wyrażona jako:

$$(2) \quad y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i)x_{ik} + \varepsilon_i$$

gdzie (u_i, v_i) to współrzędne i -tego punktu, a $\beta_k(u_i, v_i)$ jest realizacją funkcji $\beta_k(u, v)$ w punkcie i . Podstawową cechą odróżniającą GWR od MLR jest uzależnienie procesu od położenia w przestrzeni (2), w związku z czym współczynniki regresji (β) nie są stałe, jak w MLR, ale zmieniają się wraz z lokalizacją. Istotą GWR jest konstrukcja lokalnych modeli regresji opartych na macierzy sąsiedztwa (ang. *spatial kernel*) z zastosowaniem funkcji wagowej (ang. *weighting function, kernel function*) (Kulczycki, Ligas, 2007). Wagi zmniejszają się wraz z odległością punktu danych od punktu regresji. W GWR możliwe jest zastosowanie stałej lub zmiennej macierzy sąsiedztwa. Stała stosowana jest z reguły przy równomiernym rozmieszczeniu punktów w przestrzeni i definiowana maksymalnym promieniem poszukiwań punktów pomiarowych, na podstawie których określany jest model lokalny. W przypadku nieregularnej sieci punktów jej użycie będzie skutkowało dużą wariancją lokalnych estymatorów (przy małej gęstości) lub zamaskowana zostanie lokalna zmienność interpolowanego elementu (przy dużej gęstości). Macierz zmienna dopasowuje się do przestrzennego rozmieszczenia obserwacji. Definiuje się ją stałą liczbą punktów służących do kalibracji modelu (Fotheringham *et al.*, 2002).

Nierównomierne rozmieszczenie punktów pomiarowych w analizowanych przypadkach zdecydowało o wyborze zmiennej macierzy sąsiedztwa. Jej wielkość określono porównując miary modeli opartych na tych samych zmiennych objaśniających: odchylenie

nia standardowego reszt regresji (σ) i skorygowanego kryterium informacyjnego Akaike (AICc) (Hurvich *et al.*, 1998) oraz analizując estymowane parametry regresji (β) zgodnie z przyjętym założeniem fizycznej interpretowalności. Optymalne parametry GWR ustalone zostały poprzez porównanie rezultatów dla wszystkich możliwych wielkości zmiennej macierzy sąsiedztwa.

3.3. Kriging reszt

Zarówno globalny, jak i lokalny model regresji wyjaśniają jedynie część zmienności przestrzennej analizowanego elementu, pozostawiając składnik losowy – resztę (ε) w każdym punkcie regresji. Uwzględnienie czynnika losowego w procesie interpolacji przestrzennej możliwe jest, zarówno w przypadku MLR jak i GWR, poprzez zastosowanie krigingu reszt regresji. Jest to procedura dwuetapowa. W pierwszym kroku następuje wyjaśnienie części deterministycznej za pomocą dodatkowych zmiennych objaśniających (MLR lub GWR), a następnie dokonuje się interpolacji reszt regresji metodą krigingu zwyczajnego (uprzestrzennienie części stochastycznej). Wynik jest sumą rozkładów uzyskanych w obydwu etapach.

3.4. Ocena jakości interpolacji

Zbiór wejściowy (liczba punktów pomiarowych) w obydwu analizowanych przypadkach był stosunkowo niewielki, a utrata części danych na rzecz podzbioru kontrolnego mogłaby istotnie pogorszyć jakość interpolacji. Dodatkowo, nieliczny zbiór kontrolny, wydzielony z całego zbioru obserwacji, nie pozwala na wiarygodną ocenę statystyczną błędów. Z tych powodów zdecydowano się na przeprowadzenie walidacji interpolacji przestrzennej metodą oceny krzyżowej (ang. *cross-validation* – CV) typu *leave-one-out*, polegającą na sekwencyjnej eliminacji danej z każdego punktu i dokonaniu interpolacji w tym punkcie w oparciu o pozostałe dane. Jest to metoda uniwersalna, szczególnie polecana przy niewielkiej liczebności zbioru i zalecana do stosowania w klimatologii i meteorologii (Daly, 2006). Zaletą CV jest możliwość analizy rozkładu błędów estymacji w przestrzeni. Metoda umożliwia także obliczenie charakterystyk rozkładu i szeregu miar diagnostycznych błędów oszacowania, jak: błąd średni (BIAS), średni błąd absolutny (MAE), pierwiastek średniego błędu kwadratowego (RMSE) i in., potrzebnych do weryfikacji wyników estymacji.

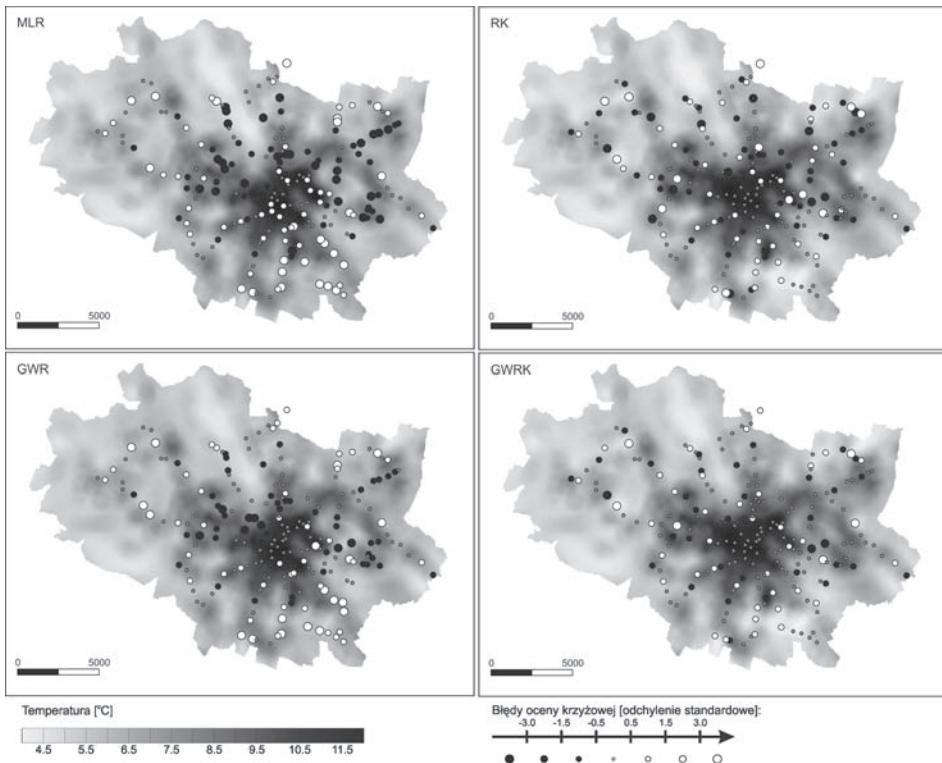
4. WYNIKI I DYSKUSJA

Początkową fazą procedury interpolacyjnej był dobór zmiennych objaśniających za pomocą regresji krokowej i kalibracja modelu MLR. W obydwu analizowanych przypadkach stwierdzono, iż znaczna część wariacji temperatury pozostaje niewyjaśniona: 23% dla 22.05.2001 i 34% dla 3.01.2002 (Tab. 1).

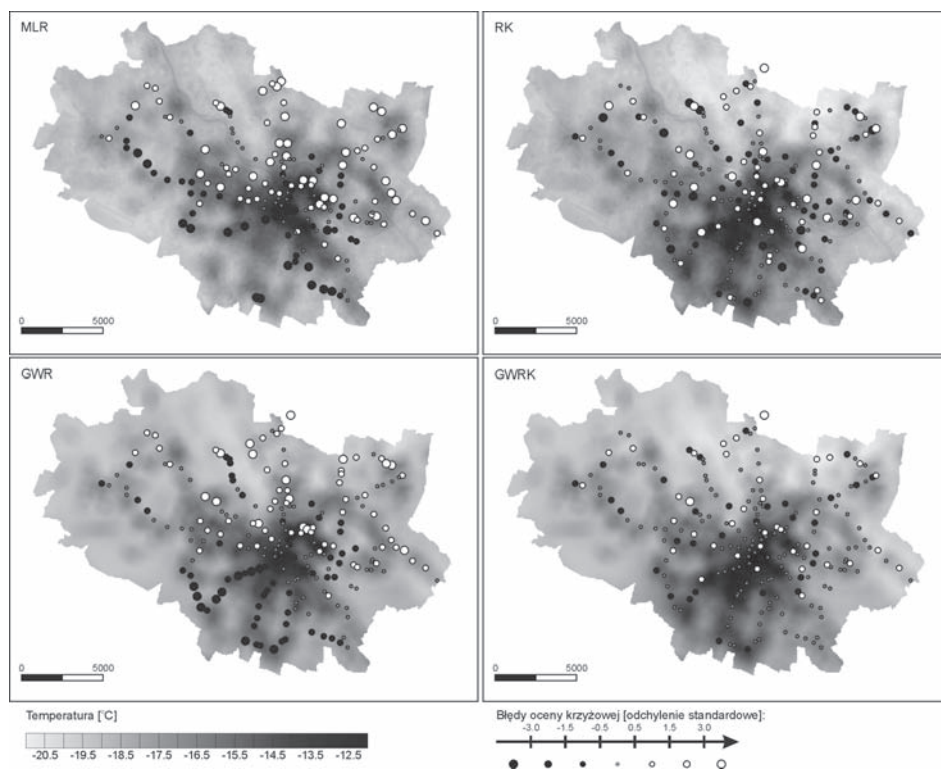
Tab. 1. Analiza regresji dla modelu globalnego (MLR).

Data	Predyktory (wg istotności)	Błąd standardowy [°C]	R ²	F (n=206)
22.05.2001	NDVI, z ₀ , Q _A	0.98	0.77	224.8
3.01.2002	Q _A , AS, NDVI	1.14	0.66	134.8

Po specyfikacji modelu globalnego MLR, dokonano przestrzennej interpolacji (Rys. 2, Rys. 3) oraz przeprowadzono ocenę krzyżową. Analiza rozmieszczenia przestrzennego błędów CV wskazała na występowanie czynników zewnętrznych prowadzących do systematycznych błędów w interpolacji. Szczególnie czytelne jest to w przypadku z 3.01.2002, kiedy stwierdzono wyraźną tendencję do przeszacowania wyników w części NE, a niedoszacowania w części SW miasta (Rys. 3). Analiza warunków meteorologicznych w tej sytuacji wskazała na wystąpienie słabego, ale stałego wiatru regionalnego z NE, powodującego przesunięcie struktury UHI *en bloc* na SW, a więc czynnika nie opisanego



Rys. 2. Temperatura powietrza i błędy oceny krzyżowej wybranych metod interpolacji dla przypadku miejskiej wyspy ciepła z 22.05.2001 r.



Rys. 3. Temperatura powietrza i błędy oceny krzyżowej wybranych metod interpolacji dla przypadku miejskiej wyspy ciepła z 3.01.2002 r.

przez zastosowane predyktory. Rozwiązaniem tego problemu było zastosowanie kriginu resztowego, które spowodowało znaczącą poprawę miar diagnostycznych (Tab. 2) i rozmieszczenia przestrzennego błędów CV, ale przede wszystkim korektę przestrzennego rozkładu pola temperatury, zwłaszcza w przypadku z 3.01.2002 (Rys. 3).

Kriging resztowy dla modelu globalnego jest rozwiązaniem znacząco poprawiającym jakość estymowanego pola pod względem miar statystycznych oraz wiarygodnym pod względem skutków oddziaływania procesów fizycznych, także w obszarze ekstrapo-

Tab. 2. Miary diagnostyczne błędów oceny krzyżowej wybranych metod interpolacji.

	22.05.2001				3.01.2002			
	MLR	RK	GWR	GWRK	MLR	RK	GWR	GWRK
BIAS	0.00	0.00	0.06	0.00	0.00	0.00	0.02	0.00
MAE	0.78	0.53	0.73	0.53	0.90	0.44	0.78	0.45
RMSE	0.99	0.68	0.93	0.68	1.15	0.54	0.96	0.56

lacji. W przypadku niestacjonarności procesu, a z takim mamy do czynienia w analizowanych przypadkach model regresji powinien być jednak określony lokalnie (GWR).

W początkowym etapie wyboru optymalnej macierzy sąsiedztwa dla analizowanych przypadków UHI porównano miary σ^2 i AICc dla wszystkich możliwych wielkości macierzy w modelu lokalnym i porównano z modelem globalnym (Tab. 3). Dla obydwu zastosowanych miar mniejsze wielkości oznaczają lepsze dopasowanie modelu.

Tab. 3. Miary dopasowania modeli MLR i GWR o zmiennej liczbie punktów w macierzy sąsiedztwa (n – liczba punktów).

	22.05.2001				3.01.2002			
	MLR	GWR	GWR	GWR	MLR	GWR	GWR	GWR
n	206	195	80	15	206	195	67	15
σ^2	0.96	0.91	0.81	0.54	1.30	1.21	0.91	0.41
AICc	580.3	577.6	557.9	505.7	642.6	635.7	569.1	446.1
R ²	0.77	0.78	0.81	0.87	0.67	0.68	0.78	0.90

Stwierdzono, iż zmniejszanie liczby punktów, z których kalibrowany model prowadzi do polepszenia jakości jego dopasowania do danych obserwacyjnych. Jednak zgodnie z założeniami metodycznymi regresji, liczba punktów, w oparciu o którą budowany jest model lokalny, musi być znacząco większa od liczby zmiennych niezależnych. Dodatkowo, przy zbyt małej liczbie punktów część lokalnych modeli może nie być istotna statystycznie. Zastosowane oprogramowanie daje możliwość automatycznego doboru macierzy sąsiedztwa na podstawie minimalizacji błędów CV lub AICc, jednak przeprowadzone próby wskazały także na trzeci czynnik, który powinien być uwzględniony przy ustalaniu wielkości zmiennego sąsiedztwa. Im mniejsza liczba punktów jest użyta do kalibracji modelu, tym bardziej rośnie zmienność estymowanych parametrów regresji (β), aż do momentu, w którym dla jednej bądź kilku ze zmiennych niezależnych obserwujemy zmianę znaku na przeciwny (przejście przez zero). W takim przypadku nie ma możliwość fizycznej interpretacji modelu, która leży u podstaw deterministycznego podejścia do analizy regresji. W związku z tym zdecydowano, iż optymalnym będzie zastosowanie możliwie najmniejszej liczby punktów, dla których model jest istotny statystycznie i fizycznie interpretowalny. Warunek ten był spełniony dla sąsiedztwa 80-punktowego w przypadku 22.05.2001 i 67-punktowego w przypadku 3.01.2002. Błędy oceny krzyżowej GWR są mniejsze niż dla MLR (Tab. 2). Zauważalna jest zmiana w rozkładzie przestrzennym błędów, jednak ciągle widoczne są obszary z systematyczną tendencją do przeszacowania lub niedoszacowania temperatury powietrza (Rys. 2, Rys. 3). Próby przeprowadzone dla mniejszej liczby punktów wykazały, iż takie modele lepiej estymują proces zewnętrzny, który nie został wyjaśniony przez zmienne niezależne, jednak odbywa się to kosztem utraty możliwości fizycznej interpretacji. W związku z tym, podobnie jak dla MLR przeprowadzono procedurę krigingu reszt (GWRK), który dla analizowanych przypadków można uznać za optymalną metodę interpolacji przestrzennej. Rezultaty oceny krzyżowej dla RK i GWRK są zbliżone, jeśli chodzi o statystyki błędów (Tab. 2).

W rozkładzie przestrzennym UHI można odnaleźć różnice, zwłaszcza w obszarze ekstrapolacji, jednak ilościowe stwierdzenie, w którym przypadku rezultat jest bardziej wiarygodny jest praktycznie niemożliwe ze względu na brak danych pomiarowych.

5. PODSUMOWANIE

Przeprowadzona w oparciu o dwa typowe przypadki miejskiej wyspy ciepła we Wrocławiu analiza wyników przestrzennej interpolacji wskazuje na regresję ważoną geograficznie jako na zasadną metodę interpolacji w warunkach niestacjonarności procesu przestrzennego i nieregularnego rozmieszczenia punktów. Pomijając kontrowersje metodyczne stosowania modelu globalnego przy braku stacjonarności, porównanie wyników oceny krzyżowej dla MLR i GWR wskazuje na lepsze dopasowanie do danych modelu lokalnego. W obydwu przypadkach (MLR i GWR) stosowano te same zmienne objaśniające dobrane uprzednio metodą regresji krokowej. Zastosowanie GWR wiąże się z przyjęciem pewnych założeń i znalezieniem optymalnych parametrów modelu. Zróżnicowana gęstość przestrzenna punktów powoduje konieczność zastosowania zmiennego sąsiedztwa, w obrębie którego specyfikowany jest model. Pomimo, iż zmniejszanie liczby punktów prowadzi do statystycznie lepszych rezultatów interpolacji, to zachowanie możliwości fizycznej interpretacji modelu wymusza przyjęcie takiej liczby punktów, dla której nie obserwuje się zmiany znaków lokalnych współczynników regresji. Optymalna liczba punktów, z których budowany jest model lokalny, a przyjęte założenia metodyczne są zachowane, została ustalona dla przypadku 22.05.2001 na 80 punktów, a dla 3.01.2002 na 67 punktów. Zastosowanie kręgu reszt, zarówno dla modelu globalnego, jak i lokalnego, prowadzi do znaczącej poprawy jakości interpolacji, jednak statystyki błędów CV dla RK i GWRK w analizowanych przypadkach nie różnią się istotnie, co nie pozwala na jednoznaczne wskazanie lepszej metody. Decydujące o wyborze algorytmu jest w tym wypadku zachowanie założeń metodycznych, wskazujących na konieczność stosowania modeli lokalnych w warunkach niestacjonarności procesu przestrzennego. Można zatem uznać, iż spośród analizowanych algorytmów interpolacji przestrzennej optymalną metodą jest kręgu resztowy dla regresji ważonej geograficznie (GWRK).

Praca naukowa finansowana ze środków na naukę w latach 2010-2012 jako projekt badawczy nr N N306 155038.

6. LITERATURA

Daly C., 2006. Guidelines for assessing the suitability of spatial climate data sets. *International Journal of Climatology*, 26, s. 707-721.

Dobesch H., Dumolard P., Dyras I., 2007. *Spatial interpolation for climate data, the use of GIS in climatology and meteorology*, ISTE, London.

Fotheringham A.S, Brunsdon C., Charlton C., 2002. *Geographically weighted regression*. John Wiley & Sons, Chichester.

Hurvich C.M., Simonoff J.S., Tsai C.L., 1998. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of Royal Statistical Society*, 60, s. 271-293.

Kulczycki M., Ligas M., 2007. Regresja ważona geograficznie jako narzędzie analizy rynku nieruchomości. *Geomatics and Environmental Engineering*, 1, s. 59-68.

Lloyd C.D., 2007. *Local models for spatial analysis*. CRC Press, Boca Raton.

Oke T.R., 1976. The distinction between canopy and boundary layer urban heat islands. *Atmosphere*, 14, s. 268–277.

Oke T.R., 1987. *Boundary layer climates*, 2nd edn. Methuen, London.

Szymanowski M., 2004. Miejska wyspa ciepła we Wrocławiu. *Studia Geograficzne*, 77, Wydawnictwo Uniwersytetu Wrocławskiego, Wrocław.

Szymanowski M., 2005. Interactions between thermal advection in frontal zones and the urban heat island of Wrocław, Poland. *Theoretical and Applied Climatology*, 82, s. 207-224.

Szymanowski M., Kryza M., 2008. GIS application for the spatial interpolation of the urban heat island in Wrocław, Poland, [w:] Kłysik K., Wibig J., Fortuniak K. (red.), *Klimat i bioklimat miast*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź, s. 21-34.

Szymanowski M., Kryza M., 2009. GIS-based techniques for urban heat island spatialization. *Climate Research*, 38, s. 171-187.

Tveito O.E., Wegehenkel M., van der Wel F., Dobesch H. (red.), 2008. *The use of geographic information systems in climatology and meteorology, Final Report, COST Action 719*. Office for Official Publications of the European Communities, Luxembourg.

APPLICATION OF GEOGRAPHICALLY WEIGHTED REGRESSION FOR URBAN HEAT ISLAND MODELLING IN WROCLAW

KEY WORDS: geographically weighted regression, spatial interpolation, GIS, urban heat island, Wrocław

SUMMARY: Implementation of geographically weighted regression (GWR) for urban heat island (UHI) modelling in Wrocław is presented. The results of spatial interpolation using 4 methods are evaluated and compared. The methods are: multiple linear regression (MLR) – global model, GWR – local model and residual kriging for both regression techniques (RK and GWRK). The analysis was performed based on 2 examples of UHI. Air temperature data were gathered in 206 irregularly distributed points. Additional explanatory variables were developed based mainly on the land-use map and Landsat TM images. The cross-validation (CV) technique was used to compare results obtained with the different algorithms together with the evaluation of errors (BIAS, MAE, RMSE) and their spatial distribution. The results confirmed the usefulness of GWR

in the case of non-stationarity of the spatial process. Calibration of the local models was performed using adaptive kernel, taking into account the possibility of physical interpretation of the model. Kriging of MLR and GWR residuals significantly improved the spatial interpolation results in terms of cross-validation errors. The most accurate results of the UHI spatialization were obtained with the GWRK technique

Mariusz Szymanowski
szymanowski@geogr.uni.wroc.pl
telefon: +48 71 3752830
fax: +48 71 3435184

dr Maciej Kryza
kryzam@meteo.uni.wroc.pl
telefon: +48 71 3729497
fax: +48 71 3729498

