

# **ACTA SCIENTIARUM POLONORUM**

Czasopismo naukowe założone w 2001 roku przez polskie uczelnie rolnicze

## **Geodesia et Descriptio Terrarum**

Geodezja i Kartografia

7(2) 2008



Bydgoszcz Kraków Lublin Olsztyn  
Poznań Siedlce Szczecin Warszawa Wrocław

**Rada Programowa *Acta Scientiarum Polonorum***

Kazimierz Banasik (Warszawa), Janusz Falkowski (Olsztyn),  
Florian Gambuś (Kraków), Franciszek Kluza (Lublin), Edward Niedźwiecki (Szczecin),  
Janusz Prusiński (Bydgoszcz), Jerzy Sobota (Wrocław) – przewodniczący,  
Stanisław Socha (Siedlce), Waldemar Uchman (Poznań)

**Rada Naukowa serii *Geodesia et Descriptio Terrarum***

Bernard Kontny (Wrocław) – przewodniczący, sekretarz,  
Hieronim Olenderek (Warszawa), Alojzy Wasilewski (Olsztyn), Josef Weigel (Brno),  
Mirosław Żak (Kraków)

Korekta:  
Elżbieta Winiarska-Grabosz  
Janina Szydłowska

Łamanie  
Teresa Alicja Chmura

Projekt okładki  
Daniel Morzyński

ISSN 1644–0668

*Wydanie publikacji dofinansowane ze środków Uniwersytetu Przyrodniczego  
we Wrocławiu*

© Copyright by Wydawnictwo Uniwersytetu Przyrodniczego we Wrocławiu,  
Wrocław 2008

Redaktor Naczelny – prof. dr hab. Andrzej Kotecki  
ul. Sopocka 23, 50–344 Wrocław, tel./fax (071) 328–12–77  
e-mail: [wyd@ozi.ar.wroc.pl](mailto:wyd@ozi.ar.wroc.pl) <http://www.up.wroc.pl>

Nakład 200 + 16 egz. Ark. druk. 3,75  
Druk i oprawa: Wydawnictwo Tekst Sp. z o.o.  
ul. Kossaka 72, 85–307 Bydgoszcz

## **METODY JAKOŚCIOWE I ILOŚCIOWE NA USŁUGACH WYCENY NIERUCHOMOŚCI\***

Anna Barańska

Akademia Górniczo-Hutnicza w Krakowie

**Streszczenie.** W pracy zostanie zaprezentowana propozycja jednoczesnego wykorzystania metod jakościowych i ilościowych w procesie dochodzenia do wartości rynkowej nieruchomości, na przykładzie konkretnego algorytmu wyceny.

W pierwszym etapie analizy, na podstawie informacji rynkowych stanowiących bazę do wyceny, będą estymowane parametry dwóch modeli: addytywnego w postaci liniowej regresji wielorakiej oraz multiplikatywnego, uwzględniającego wykładniczą zmienność cen nieruchomości względem ich atrybutów. Weryfikacja hipotez statystycznych dotyczących wartości estymowanych parametrów modeli będzie podstawą do wnioskowania o ich istotności oraz pozwoli na wybór modelu o lepszej wiarygodności. Na podstawie parametrów wybranego modelu oraz atrybutów nieruchomości wycenianej będzie określana jej wartość rynkowa wraz z pełną analizą wariancji. Tak określona wartość rynkowa nieruchomości, stanowiąca predykcję estymowanego modelu, będzie korygowana poprawką losową, która zostanie określona w drugim etapie analizy.

W drugim etapie – analizie statystycznej będą podlegały odchyłki losowe do jednostkowych cen nieruchomości, wraz z ich macierzą kowariancji, które zostały określone w etapie pierwszym. Do tej analizy zostanie opracowana procedura, bazująca na metodach jakościowych, ustalania stopnia podobieństwa nieruchomości wycenianej i nieruchomości w bazie. W wyniku analizy jakościowej zostanie ustalona podgrupa nieruchomości, która wykazuje najwyższy stopień podobieństwa do nieruchomości wycenianej.

Odchyłki losowe dla jednostkowych cen nieruchomości ustalonej podgrupy oraz ich macierz kowariancji, określone w pierwszym etapie, stanowią podstawę do wyznaczenia poprawki losowej dla wycenianej nieruchomości. Po dodaniu tej poprawki do rynkowej wartości nieruchomości uzyskanej z predykcji modelu, określonej w etapie pierwszym, uzyskuje się bardziej wiarygodną wartość rynkową wycenianej nieruchomości.

**Słowa kluczowe:** modelowanie rynku nieruchomości, czynnik systematyczny, metody jakościowe, metody ilościowe

---

\* Praca finansowana z badań własnych nr 10.10.150.841, prowadzonych w Katedrze Geomatyki, na Wydziale Geodezji Górniczej i Inżynierii AGH w Krakowie.

## WPROWADZENIE

Zaprezentowane rozważania dotyczą sposobu określania bardziej wiarygodnej prognozy rynkowej wartości nieruchomości na podstawie starannie dobranego modelu wyceny, odpowiedniego dla danego lokalnego rynku. W celu urealnienia wykonywanej predykcji zaproponowano dodanie do wartości modelowej czynnika systematycznego. Czynniki systematyczny jest wyznaczany na podstawie odchyłek losowych nieruchomości najbardziej podobnych do wycenianej spośród zgromadzonej bazy danych. Wybór nieruchomości najbardziej podobnych odbywa się przy wykorzystaniu metod jakościowych.

## POSTAĆ TESTOWANYCH MODELI WYCENY

### Model liniowy addytywny

$$c = \sum_{i=1}^m (x_i - \hat{x}_i) \cdot a_i \quad (1)$$

gdzie:

- $c$  – jednostkowa cena nieruchomości,
- $x_i$  – wartość atrybutu  $i$ ,
- $\hat{x}_i$  – przeciętna wartość atrybutu  $i$  w bazie,
- $a_i$  – parametry modelu.

### Model nieliniowy multiplikatywny

$$c = a_1^{x_1 - \hat{x}_1} \cdot a_2^{x_2 - \hat{x}_2} \cdot \dots \cdot a_m^{x_m - \hat{x}_m} \quad (2)$$

gdzie:

- $c$  – jednostkowa cena nieruchomości,
- $x_1, x_2, \dots, x_m$  – atrybuty nieruchomości,
- $\hat{x}_i$  – przeciętna wartość atrybutu  $i$  w bazie,
- $a_j$  – parametry modelu.

## ESTYMACJA PARAMETRÓW MODELI WYCENY

### Model liniowy addytywny

Wyrażenie (1) można zapisać w następującej formie macierzowej:

$$[C] = [X] \cdot [a] + [\delta] \quad (3)$$

gdzie:

$$[C] = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \quad \text{– wektor zmiennej losowej objaśnianej (ceny nieruchomości),}$$

$$[X] = \begin{bmatrix} x_{11} - \hat{x}_1 & \cdots & x_{1m} - \hat{x}_m \\ x_{21} - \hat{x}_1 & \cdots & x_{2m} - \hat{x}_m \\ \vdots & \ddots & \vdots \\ x_{n1} - \hat{x}_1 & \cdots & x_{nm} - \hat{x}_m \end{bmatrix} \quad \text{– macierz zawierająca różnice zmiennych objaśniających (atrybutów) i ich średnich wartości,}$$

$$[a] = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix} \quad \text{– wektor współczynników liniowej regresji wielorakiej,}$$

$$[\delta] = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_n \end{bmatrix} \quad \text{– wektor odchyłek losowych modelu.}$$

Zastosowanie metody najmniejszych kwadratów, czyli wyznaczenie estymatora  $\hat{a}$  wektora współczynników  $a$ , takiego, że:

$$\sum_{i=1}^n (c_i - w_i)^2 = \delta^T \delta = \min \quad (4)$$

gdzie:

$c_i$  – cena zaobserwowana dla  $i$ -tej nieruchomości w bazie,

$w_i$  – modelowa wartość nieruchomości,

$\delta = C - W$  – wektor odchyłek losowych (różnic między cenami zaobserwowanymi  $C$  i wartościami modelowymi  $W$ ),

wiedzie do układu równań normalnych, który jest rozwiązywany za pomocą numerycznych procedur iteracyjnych.

Przez rozwiązanie uogólnionego modelu liniowego rozumie się:

- określenie nieobciążonego estymatora wektora niewiadomych:

$$\hat{a} = (X^T X)^{-1} \cdot X^T C = X^+ \cdot C \quad (5)$$

gdzie:

$X^+$  – pseudoodwrotność macierzy  $X$ ,

- określenie nieobciążonego estymatora wariancji resztowej, określającej niedokładność estymacji parametrów modelu:

$$\hat{\sigma}_0^2 = \frac{C^T C - \hat{a}^T X^T C}{n - m} \quad (6)$$

- wyznaczenie macierzy kowariancji wektora niewiadomych (parametrów modelu):

$$\text{Cov}(\hat{a}) = \hat{\sigma}_0^2 \cdot (X^T X)^{-1} \quad (7)$$

– wyznaczenie macierzy kowariancji wartości modelowych:

$$\text{Cov}(W) = \hat{\sigma}_0^2 \cdot X(X^T X)^{-1} X^T \quad (8)$$

Wartość każdego współczynnika regresji podajemy wraz z poziomem istotności, który jest wskaźnikiem wiarygodności dla wartości otrzymanej na podstawie próby w stosunku do całej populacji.

### Model nieliniowy multiplikatywny

Do estymacji wartości współczynników  $a_i$  funkcję (2) trzeba doprowadzić do postaci liniowej. W tym celu logarytmujemy stronami równanie (2), przy wykorzystaniu logarytmu naturalnego, otrzymując:

$$\ln c = (x_1 - \hat{x}_1) \cdot \ln a_1 + (x_2 - \hat{x}_2) \cdot \ln a_2 + \dots + (x_m - \hat{x}_m) \cdot \ln a_m \quad (9)$$

Układ równań postaci (9) ma cechy modelu probabilistycznego, który w zapisie macierzowym przyjmuje następującą postać:

$$[\ln C] = [X] \cdot [\ln a] + [\ln \delta] \quad (10)$$

gdzie:

$[\ln C]$  – macierz jednokolumnowa,  $[n \times 1]$  zawierająca logarytmy naturalne cen nieruchomości z bazy,

$[X]$  – macierz prostokątna pionowa  $[n \times m]$ , zawierająca różnice wartości atrybutów nieruchomości z bazy i ich średnich wartości,

$[\ln a]$  – macierz jednokolumnowa  $[m \times 1]$ , zawierająca estymowane wartości logarytmów naturalnych współczynników  $a_i$  modelu (2),

$[\ln \delta]$  – macierz jednokolumnowa  $[n \times 1]$ , zawierająca odchyłki losowe do logarytmów naturalnych cen nieruchomości z bazy,

przy czym  $n$  oznacza liczbę rozważanych nieruchomości w bazie, zaś  $m$  – liczbę rozważanych atrybutów.

Po zastosowaniu metody najmniejszych kwadratów otrzymuje się następujące wzory na estymowane parametry modelu nieliniowego:

$$\hat{a} = \exp \left[ (X^T X)^{-1} \cdot X^T \ln C \right] \quad (11)$$

Macierz kowariancji dla estymowanych parametrów wyznacza się według wzoru:

$$\text{Cov}[\hat{a}] = \hat{\sigma}_0^2 (X^T X)^{-1} \quad (12)$$

przy czym  $\sigma_0^2$  oznacza wariancję estymacji nieliniowego modelu multiplikatywnego, której estymator określa się według wzoru:

$$\hat{\sigma}_0^2 = \frac{[\delta]^T \cdot [\delta]}{n - m} \quad (13)$$

gdzie:

$\delta$  – odchyłki losowe modelu (różnice między cenami rynkowymi i ich modelowymi wartościami).

### Weryfikacja modeli wyceny

Podstawowym miernikiem jakości dopasowania modelu do danych jest kwadrat współczynnika korelacji krzywoliniowej, który dla modelu liniowego odpowiada współczynnikowi determinacji i określa udział wariancji wyjaśnionej przez model w stosunku do całkowitego rozproszenia zmiennej objaśnianej wokół jej wartości przeciętnej.

W ramach statystycznej weryfikacji modeli badamy istotność układu uzyskanych parametrów w każdym z modeli oraz istotność każdego z parametrów z osobna. Szczegółowe omówienie procedury weryfikacji modeli zawiera publikacja [Barańska 2005].

## PROGNOZOWANIE JEDNOSTKOWEJ WARTOŚCI RYNKOWEJ NIERUCHOMOŚCI

Wyestymowany model wyceny prowadzi do wyodrębnienia czynnika systematycznego modelu z jednostkowych cen nieruchomości. Są to wartości prognoz cen nieruchomości uzyskane z modelu  $W$ . Różnice rzeczywistych cen rynkowych, zebranych w bazie danych  $C$  i wartości modelowych  $W$  to, jak wspomniano wyżej, odchyłki losowe modelu:

$$[\delta] = [C] - [W] \quad (14)$$

gdzie:

$[W]$  – wektor modelowych cen nieruchomości,

$[C]$  – wektor rzeczywistych cen nieruchomości w bazie danych,

$[\delta]$  – odchyłki losowe modelu wyceny.

Każdemu czynnikowi systematycznemu odpowiada zbiór odchyłek losowych modelu  $\delta$ , który posiada charakterystykę niedokładności zawartą w swojej macierzy kowariancji:

$$[\delta] = [C] - [X] \cdot [X]^+ \cdot [C] = [I - X \cdot X^+] \cdot [C] \quad (15)$$

$$\text{Cov}[\delta] = \sigma_0^2 \cdot \left[ I - X \cdot (X^T \cdot X)^{-1} \cdot X^T \right] \quad (16)$$

gdzie:

$X^+$  – pseudoodwrotność macierzy  $X$ ,

przy czym wariancja resztowa jest szacowana według wzoru (6) lub (13), w zależności od rozpatrywanego modelu.

### Wybór nieruchomości reprezentatywnych przy zastosowaniu metod jakościowych

Stosując analizę jakościową, z zebranej bazy danych wybieramy grupę  $k$  nieruchomości, zawierającą nieruchomość wycenianą i najbardziej do niej podobnych. Dokonując wyboru, postępujemy zgodnie z jedną z metod jakościowych [Czaja i Parzych 2007]:

- analiza porównania względnego,
- analiza szeregowania nieruchomości.

Analiza porównania względnego nieruchomości opiera się na porównaniu jakościowym cech nieruchomości wycenianej z cechami nieruchomości podobnych w taki sposób, że dla każdego atrybutu nieruchomości porównywanej ustala się znak „-1” lub „+1” poprawki korygującej, w zależności od tego czy wartość danego atrybutu jest odpowiednio: większa czy mniejsza od wartości tego atrybutu dla nieruchomości wycenianej. Wydzielenie nieruchomości najbardziej podobnych do wycenianej następuje na podstawie sumy poprawek dla wszystkich atrybutów. Wybieramy wszystkie nieruchomości, dla których suma poprawek wynosi „0” oraz najtańszą z nieruchomości o ujemnej sumie poprawek i najdroższą z nieruchomości o dodatniej sumie poprawek.

Analiza szeregowania nieruchomości polega na nadaniu rang każdemu z atrybutów nieruchomości podobnych oraz nieruchomości wycenianej. Rangi nadaje się w odniesieniu do najniższych lub najwyższych wartości poszczególnych atrybutów w bazie; przy czym uwzględnia się „odległość” wartości danego atrybutu od punktu odniesienia. A zatem rangi mogą przyjmować wartości całkowite  $|r| > 1$ . Podobnie jak w poprzedniej metodzie sumaryczną poprawką dla każdej nieruchomości jest suma rang nadana poszczególnym atrybutom. Oczywiście jest, że gdy za punkt odniesienia przyjmiemy najniższe wartości atrybutów w bazie, to wszystkie rangi będą miały znak ujemny. W sytuacji przeciwnej – dodatni.

### Ostateczna prognoza wartości rynkowej wraz z analizą dokładności

Z wektora składnika losowego wyodrębniamy  $\delta_{i_w}$ , odpowiadające wybranym nieruchomościom, a z macierzy kowariancji – podmacierz  $\text{Cov}[\delta_w]$  o wymiarach  $(k \times k)$ , zawierającą elementy odpowiadające wyodrębnionym odchyłkom.

Z modelu systematycznego szacujemy modelową rynkową wartość nieruchomości wycenianej  $w_M$ :

$$w_M = [X_w] \cdot [\hat{a}] \quad (17)$$

gdzie:

$[X_w] = [x_1 \quad x_2 \quad \dots \quad x_m]$  – wektor wartości atrybutów wycenianej nieruchomości,

$[\hat{a}]$  – wektor wyestymowanych parametrów modelu.



Odchylenie standardowe  $\sigma(w_M)$  oszacowania  $w_M$  wyliczamy następująco:

$$V(w_M) = \sigma^2(w_M) = \sigma_0^2 \cdot [X_w] \cdot Cov[\hat{a}] \cdot [X_w]^T \quad (18)$$

Natomiast z modelu losowego szacujemy rynkową wartość czynnika losowego wycenianej nieruchomości:

$$w_L = [\underline{1} \cdot P \cdot \underline{1}^T]^{-1} \cdot [\underline{1} \cdot P] \cdot [\delta_w] \quad (19)$$

gdzie:

$\underline{1} = [1 \ 1 \ \dots \ 1]$  – wektor złożony z jedynek o wymiarach  $(1 \times k)$

przy czym macierz wag stanowi odwrotność macierzy kowariancji:

$$P = Cov^{-1}[\delta_w] \quad (20)$$

Odchylenie standardowe oszacowania czynnika losowego wynosi:

$$\sigma^2(w_L) = \sigma_{0_w}^2 \cdot [\underline{1} \cdot P \cdot \underline{1}^T]^{-1} \quad (23)$$

gdzie wariancja resztowa  $\sigma_{0_w}^2$  jest określona dla grupy  $k$  wybranych nieruchomości:

$$\hat{\sigma}_{0_w}^2 = \frac{\delta_w^T \cdot P \cdot \delta_w - w_L \cdot \underline{1}^T \cdot P \cdot \delta_w}{k - 1} \quad (21)$$

Ostateczna wartość rynkowa wycenianej nieruchomości będzie obliczona następująco:

$$w_{M+L} = w_M + w_L \quad (22)$$

Pomiędzy wariancjami prognozy rynkowej wartości nieruchomości, uzyskanej z modelu  $w_M$  oraz prognozy poprawionej o czynnik losowy  $w_{M+L}$ , zachodzi następująca relacja:

$$V(w_M) = V(w_{M+L}) + V(w_L) \quad (23)$$

Zatem odchylenie standardowe ostatecznej wartości rynkowej nieruchomości wyliczymy ze wzoru:

$$\sigma^2(w_{M+L}) = \sigma^2(w_M) - \sigma^2(w_L) \quad (24)$$

## PRZYKŁAD PRAKTYCZNY

Zaprezentowany powyżej sposób postępowania został zweryfikowany na różnych lokalnych rynkach nieruchomości różnych typów. Za każdym razem doprowadził do poprawienia prognozowanej na podstawie modelu wartości rynkowej wycenianej nieru-

chomości, zmniejszając jej odchylenie standardowe oraz uściślając jej wartość poprzez dodanie poprawki korygującej. Poniżej zamieszczono wyniki wyceny lokalu mieszkalnego z terenu miasta Rzeszowa, na podstawie jednej bazy danych, ale za pomocą dwóch modeli wyceny: addytywnego i multiplikatywnego.

### Wycena według modelu addytywnego

Tabela 1. Macierz kowariancji dla odchyłek dotyczących wybranych nieruchomości  $Cov[\delta_w]$

Table 1. Covariance matrix for remainders concerning selected real estate

41533	4641	-1228	3446	1787	1391	-981	-2179
4641	39834	519	-9664	-2049	-94	1874	-18
-1228	519	45167	1722	-1251	-316	-4880	367
3446	-9664	1722	39023	130	2608	2123	1288
1787	-2049	-1251	130	45285	-2011	-1035	-428
1391	-94	-316	2608	-2011	43289	654	-3766
-981	1874	-4880	2123	-1035	654	46177	907
-2179	-18	367	1288	-428	-3766	907	44291

Tabela 2. Macierz wagowa  $P = Cov^{-1}[\delta_w]$

Table 2. Weight matrix

0,0000249	-3,8E-06	9E-07	-3,2E-06	-1,1E-06	-6E-07	9E-07	1,2E-06
-3,8E-06	2,75E-05	-8E-07	7,3E-06	1,3E-06	-2E-07	-1,6E-06	-4E-07
0,0000009	-8E-07	2,25E-05	-1,4E-06	6E-07	2E-07	2,5E-06	-1E-07
-3,2E-06	7,3E-06	-1,4E-06	0,000028	2E-07	-1,6E-06	-1,8E-06	-1,1E-06
-1,1E-06	1,3E-06	6E-07	2E-07	2,23E-05	1,1E-06	5E-07	2E-07
-6E-07	-2E-07	2E-07	-1,6E-06	1,1E-06	2,34E-05	-3E-07	0,000002
0,0000009	-1,6E-06	2,5E-06	-1,8E-06	5E-07	-3E-07	2,21E-05	-4E-07
0,0000012	-4E-07	-1E-07	-1,1E-06	2E-07	0,000002	-4E-07	2,29E-05

Wartość modelowa nieruchomości:	1909,06 [zł/m <sup>2</sup> ]
Odchylenie standardowe wartości modelowej:	113,88 [zł/m <sup>2</sup> ]
Wartość losowa:	- 84,90 [zł/m <sup>2</sup> ]
Odchylenie standardowe wartości losowej:	75,52 [zł/m <sup>2</sup> ]
Ostateczna prognoza wartości:	1824,17 [zł/m <sup>2</sup> ]
Odchylenie standardowe ostatecznej prognozy:	85,23 [zł/m <sup>2</sup> ]

**Wycena według modelu multiplikatywnego**Tabela 3. Macierz kowariancji dla odchylek dotyczących wybranych nieruchomości  $Cov[\delta_w]$ 

Table 3. Covariance matrix for remainders concerning selected real estates

39902	4459	-1180	3311	1716	1336	-943	-2094
4459	38269	499	-9285	-1969	-90	1801	-18
-1180	499	43394	1654	-1202	-303	-4688	353
3311	-9285	1654	37490	125	2505	2040	1237
1716	-1969	-1202	125	43507	-1932	-995	-411
1336	-90	-303	2505	-1932	41589	629	-3618
-943	1801	-4688	2040	-995	629	44364	871
-2094	-18	353	1237	-411	-3618	871	42552

Tabela 4. Macierz wagowa  $P = Cov^{-1}[\delta_w]$ 

Table 4. Weight matrix

0,000026	-4E-06	9E-07	-3,4E-06	-1,2E-06	-6E-07	9E-07	1,3E-06
-0,000004	2,86E-05	-9E-07	7,6E-06	1,4E-06	-2E-07	-1,6E-06	-4E-07
0,0000009	-9E-07	2,34E-05	-1,5E-06	6E-07	2E-07	2,6E-06	-1E-07
-3,4E-06	7,6E-06	-1,5E-06	2,92E-05	2E-07	-1,7E-06	-1,8E-06	-1,1E-06
-1,2E-06	1,4E-06	6E-07	2E-07	2,32E-05	1,1E-06	5E-07	2E-07
-6E-07	-2E-07	2E-07	-1,7E-06	1,1E-06	2,44E-05	-3E-07	2,1E-06
0,0000009	-1,6E-06	2,6E-06	-1,8E-06	5E-07	-3E-07	0,000023	-4E-07
0,0000013	-4E-07	-1E-07	-1,1E-06	2E-07	2,1E-06	-4E-07	2,38E-05

Wartość modelowa nieruchomości:	1915,33 [zł/m <sup>2</sup> ]
Odchylenie standardowe wartości modelowej:	111,62 [zł/m <sup>2</sup> ]
Wartość losowa:	- 90,10 [zł/m <sup>2</sup> ]
Odchylenie standardowe wartości losowej:	74,18 [zł/m <sup>2</sup> ]
Ostateczna prognoza wartości:	1825,23 [zł/m <sup>2</sup> ]
Odchylenie standardowe ostatecznej prognozy:	83,40 [zł/m <sup>2</sup> ]

**WNIOSKI**

Zaprezentowany dwuetapowy model wyceny nieruchomości ma na celu doprecyzowanie wartości nieruchomości wyznaczonej na podstawie dobrze dopasowanego do lokalnego rynku i statystycznie zweryfikowanego modelu wyceny. Jego zaletą jest wyróżnienie w bazie nieruchomości podobnych, rozpatrywanej w podejściu porównawczym, grupy nieruchomości cechujących się największym podobieństwem do wycenianego obiektu i wykorzystanie ich do „poprawienia” modelowej wartości nieruchomości.

Na podstawie zaprezentowanego przykładu widzimy, iż oba modele dały bardzo zbliżone wyniki wyceny tego samego lokalu. Odchylenie standardowe ostatecznej prognozy wartości lokalu jest na poziomie 4–5% jej wysokości i stanowi zaledwie ok. 75% odchylenia standardowego prognozy wartości nieruchomości uzyskanej wprost z modelu. Uzyskano zatem znacznie dokładniejszy wynik wyceny niż za pomocą wyłącznego wykorzystania modelu funkcyjnego.

## PIŚMIENNICTWO

- Adamczewski Z., 2006. Elementy modelowania matematycznego w wycenie nieruchomości. Podejście porównawcze. Oficyna Wydawnicza Politechniki Warszawskiej. Warszawa.
- Barańska A., 2003. Kryteria stosowania modeli stochastycznych w predykcji rynkowej wartości nieruchomości. Rozprawa doktorska, Akademia Górniczo-Hutnicza, Wydział Geodezji Górniczej i Inżynierii Środowiska. Kraków.
- Barańska A., 2004. Criteria of database quality appraisal and choice stochastic models in prediction of real estate market value. FIG Working Week 2004 and 27<sup>th</sup> General Assembly in Athens, Greece, 22–27 May 2004.
- Barańska A., 2005. Estymacja parametrów nieliniowych modeli funkcyjnych dla potrzeb predykcji rynkowej wartości nieruchomości. Kraków, UWND AGH, Geodezja, t. 2.
- Barańska A., 2006. Estimation of parameters of multiplicative exponential function model for real estate market value prediction. XXIII FIG Congress „Shaping the Change”, Munich, Germany, October 8–13, 2006.
- Barańska A., 2007. Statistical verification of real estate estimation models. FIG Working Week 2007 “Strategic Integration of Surveying Services”, Hong Kong SAR, China, 13–17 May 2007.
- Barańska A., 2007. Dwuetapowy model wyceny nieruchomości. Studia i Materiały Towarzystwa Naukowego Nieruchomości, v. 15, nr 3–4, Olsztyn.
- Czaja J., Parzych P., 2007. Szacowanie rynkowej wartości nieruchomości w aspekcie międzynarodowych standardów wyceny. Wydawnictwo Stowarzyszenia Naukowego im. ST. Staszica, Kraków.
- Czaja J., Preweda E., 2000. Analiza statystyczna zmiennej losowej wielowymiarowej w aspekcie korelacji i predykcji. Kraków, UWND AGH, Geodezja, t. 2.
- Czaja J., 2001. Metody szacowania wartości rynkowej i katastralnej nieruchomości. Kraków.
- Krysicki W., Bartos J., Dyczka W., Królikowska K., Wasilewski M., 1986. Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach. PWN, Warszawa.
- Radhakrishna Rao C. 1982. Modele liniowe statystyki matematycznej. PWN, Warszawa.

## QUALITY AND QUANTITY METHODS FOR ESTIMATING REAL ESTATES

**Abstract.** The paper presents a two-stage method of real estate estimation, basing on the probabilistic model of the variability of real estates prices in relation to their attributes. At the first analysis stage, on the basis of the market information being a base for estimation, a linear model and a multiplicative model of variability of real estate prices in relation to their attributes will be estimated. The verification of statistical hypotheses concerning the differences between the values of the estimated model parameters will be the basis for evaluating their significance and it will permit to choose a more reliable model.

On the basis of model parameters and of the attributes of an estimated real estate, we will determine its market value with a full analysis of variance. A real estate value determined in such a way, being a prediction of estimated model, will be corrected with a random correction, which is going to be determined at the second stage of the analysis.

At the second stage, the statistical analysis will include random deviations, determined at the first stage, relating to the real estate unit prices, together with their covariance matrix. For this analysis, a procedure based on quality methods will be evolved with the aim of determining the degree of similarity between an estimated real estate and a real estate in the database. As result of the quality analysis, a subgroup of real estates showing the highest degree of similarity to the estimated real estate will be established.

Random deviations for real estate unit prices of the established subgroup and their covariance matrix, determined at the first stage, are the basis for determining the random correction for the estimated real estate. This random correction being added to the real estate market value obtained from the model prediction determined at the first stage, we get the most probable market value of the estimated real estate.

**Key words:** modelling of real estate market, systematic factor, quality methods, quantity methods

Zaakceptowano do druku – Accepted for print: 30.06.2008

Do cytowania – For citation: Barańska A., 2008. Metody jakościowe i ilościowe na usługach wyceny nieruchomości. *Acta Sci. Pol. Geodesia et Descriptio Terrarum*, 7(2), 3–13.



## **FILTRACJA DANYCH LOTNICZEGO SKANINGU LASEROWEGO METODĄ RUCHOMYCH POWIERZCHNI WIELOMIANOWYCH – WERYFIKACJA METODY**

Andrzej Borkowski, Grzegorz Józków

Uniwersytet Przyrodniczy we Wrocławiu

**Streszczenie.** Głównym zagadnieniem procesu opracowania danych lotniczego skaningu laserowego, na potrzeby budowy numerycznych modeli terenu, jest identyfikacja punktów będących odbiciami od powierzchni terenu. W pracy przedstawiono metodę hierarchicznej identyfikacji punktów należących do powierzchni terenu, bazującą na aproksymacji danych ruchomą powierzchnią wielomianową. Parametry wielomianu ruchomego wyznaczone są lokalnie z wykorzystaniem estymacji odpornej metodą M-estymatorów. W procesie estymacji wykorzystano funkcję wagową zależną od odległości oraz asymetryczną funkcję tłumienia. Wykonano szereg testów numerycznych dla rzeczywistych danych lotniczego skaningu laserowego, obejmujących piętnaście zestawów testowych z danymi referencyjnymi w postaci zbiorów z poprawnie sklasyfikowanymi punktami terenu i punktami obiektów. Weryfikacja wyników filtracji polegała na porównaniu wyników automatycznej filtracji ze zbiorami referencyjnymi. W wyniku porównania określono procentowe błędy filtracji automatycznej. Całkowity błąd filtracji kształtował się na poziomie od około 1% do około 12%, w zależności od ukształtowania terenu i jego pokrycia. Stwierdzono ponadto, że wprowadzenie dodatkowej informacji *a priori* w postaci punktów należących do powierzchni terenu, w miejscach krytycznych, np. obwałowania rzek, podnosi dokładność filtracji automatycznej.

**Słowa kluczowe:** lotniczy skaningu laserowy, filtracja, wielomianowa powierzchnia ruchoma, estymacja odporna

### **WSTĘP**

Wraz z rozwojem techniki lotniczego skaningu laserowego dane o powierzchni terenu pozyskiwane są z coraz większą dokładnością i rozdzielczością. W dalszym ciągu jednak najbardziej pracochłonnym jest etap opracowania danych. Głównym zagadnie-

niem procesu opracowania danych lotniczego skaningu laserowego jest identyfikacja punktów przynależnych do odpowiednich powierzchni. Zadanie to realizowane jest jako klasyfikacja lub filtracja danych pomiarowych. Pod pojęciem filtracji rozumieć należy pewien (semi)automatyczny sposób eliminacji punktów nie należących do modelowanej powierzchni. W pracy podjęto temat filtracji w celu identyfikacji punktów należących do fizycznej powierzchni terenu. O wadze i aktualności podjętego zagadnienia świadczą może liczba opracowanych metod i liczba publikacji z tego zakresu. Znane dotychczas metody dzielimy na kilka grup:

- bazujące na odpornej predykcji liniowej [Kraus 2000, Kraus i Pfeifer 2001, Briese et al. 2002];
- polegające na iteracyjnym przybliżaniu pewnej, odpowiednio wybranej powierzchni startowej (TIN) do danych pomiarowych [Axelsson 1999, 2000];
- wykorzystujące operatory matematycznej morfologii (kryterium spadków terenu) [Vosselman i Maas 2001, Sithole 2001];
- wykorzystujące analizę skupień [Roggero 2001, Filin i Pfeifer 2006];
- bazujące na minimalizacji energii powierzchni, zależnej od jej nachylenia (spadku) [Elmqvist et al. 2001, Elmqvist 2002] i w postaci uogólnionej [Borkowski 2005].
- realizujące filtrację w dziedzinie częstotliwości: metody polegające na wykorzystaniu techniki FFT [Marmol i Jachimski 2004] lub analizy falkowej (*wavelet*) [Borkowski i Keller 2006].

Generalnie, metody filtracji można zaliczyć do jednej z dwóch grup:

- metody bazujące na danych oryginalnych,
- metody bazujące na danych opartych o regularną siatkę kwadratów, co wymaga dodatkowej interpolacji nieregularnie w płaszczyźnie  $xy$  rozrzuconych punktów (dane oryginalne).

Przegląd metod podany został w pracach: [Borkowski 2004, Sithole i Vosselman 2004]. W drugiej z nich dokonano również weryfikacji wyników filtracji przeprowadzonych różnymi metodami.

Z rozważań przeprowadzonych w pracy [Borkowski i Józków 2006] wynika, że:

- filtracja powinna być prowadzona w miarę możliwości na danych oryginalnych;
- algorytm powinien mieć dobre właściwości dopasowania do lokalnych struktur powierzchni terenu;
- algorytm powinien umożliwiać uwzględnienie dodatkowej informacji *a priori*;
- ze względu na wielkość zbiorów danych skaningu laserowego, rzędu  $10^6$ , nie bez znaczenia jest również złożoność obliczeniowa algorytmu.

W cytowanej pracy przedstawiono algorytm spełniający w znacznym stopniu podane oczekiwania. Bazuje on na, realizowanej hierarchicznie, odpornej aproksymacji danych wielomianową powierzchnią ruchomą.

W niniejszej pracy przedstawiono wyniki badań mających na celu weryfikację skuteczności filtracji z wykorzystaniem proponowanej metody. Wykonano w tym celu szereg testów numerycznych dla danych rzeczywistych skaningu pochodzących z różnych typów terenu, a wyniki filtracji automatycznej porównano z danymi referencyjnymi w postaci prawidłowo sklasyfikowanych punktów terenu i obiektów. Dane testowe i procedurę oceny wyników oparto na pracy [Sithole i Vosselman, 2004]. Otrzymane wskaźniki procentowe poszczególnych typów błędów mogą być zatem odniesione do innych, znanych z literatury, metod filtracji.



## WIELOMIAN RUCHOMY I ODPORNA ESTYMACJA PARAMETRÓW

W przestrzeni  $R^2$  wielomian dowolnego stopnia można zapisać w postaci:

$$z(x, y) = \sum_{i,j} a_{ij} x^i y^j, \quad i, j = 0, 1, 2, \dots \quad (1)$$

Ponieważ korzystne właściwości aproksymacyjne wykazują jedynie wielomiany niskiego stopnia, wykorzystano w prezentowanym algorytmie wielomian drugiego stopnia:

$$z(x, y) = a_{00} + a_{10}x + a_{01}y + a_{11}xy + a_{20}x^2 + a_{02}y^2, \quad (2)$$

gdzie:

$x, y$  – współrzędne płaskie punktu,

$a_{00}, a_{10}, a_{01}, a_{11}, a_{20}, a_{02}$  – parametry wielomianu,

$z(x, y)$  – wartość wielomianu (interpolowana powierzchnia).

Powierzchnia wielomianu dopasowywana jest do najbliższego otoczenia punktu interpolowanego, a nieznane parametry  $a_{ij}$  wielomianu oblicza się lokalnie, oddzielnie w każdym punkcie metodą najmniejszych kwadratów,

$$\sum_{i=1}^n p_i v_i^2 \rightarrow \min. \quad (3)$$

Odchyłki  $v_i$  w punktach pomiarowych o wysokości  $h_i$ ,  $i = 1, 2, \dots, n$  obliczane są z zależności:

$$v_i = a_{00} + a_{10}x + a_{01}y + a_{11}xy + a_{20}x^2 + a_{02}y^2 - h_i \quad (4)$$

natomiast wagi  $p_i$  punktów obliczane są według zależności:

$$p_i = (c / d_i)^r, \quad (5)$$

gdzie:

$c$  – wartość stała, najczęściej równa minimalnej odległości pomiędzy punktami pomiarowymi w zbiorze,

$r$  – wykładnik determinujący wpływ punktów lokalnego otoczenia na wartość interpolowaną,

$d_i$  – odległość pomiędzy punktem interpolowanym a punktem pomiarowym z najbliższego sąsiedztwa punktu interpolowanego.

Wagi punktów maleją jednocześnie ze wzrostem odległości od punktu interpolowanego, a szybkość ich zmniejszania się uzależniona jest głównie od wykładnika  $r$ , który determinuje wpływ punktów pomiarowych na przebieg powierzchni wielomianowej. Im większa wartość  $r$ , tym bardziej lokalny charakter wielomianu. Ponieważ punkty pewne

*a priori* są punktami terenowymi, zatem interpolowana powierzchnia wielomianu powinna przebiegać dokładnie lub w bardzo bliskiej odległości punktu pewnego, dlatego też dla punktów pewnych wagi  $p_i$  przyjmują zawsze stałą wartość, która jest zdecydowanie większa od wag pozostałych punktów z lokalnego otoczenia.

Parametry wielomianu (2) otrzymuje się z rozwiązania układu obserwacyjnego (4) zestawionego dla  $n$  punktów pomiarowych. Rozwiązanie w zapisie macierzowym przyjmuje postać:

$$\hat{x} = (A^T P A)^{-1} A^T P h, \quad (6)$$

gdzie:

$\hat{x} = [a_{00} \ a_{10} \ a_{01} \ a_{11} \ a_{20} \ a_{02}]^T$  – macierz parametrów wielomianu,

$$A = \begin{bmatrix} 1 & x_1 & y_1 & x_1 y_1 & x_1^2 & y_1^2 \\ 1 & x_2 & y_2 & x_2 y_2 & x_2^2 & y_2^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & y_n & x_n y_n & x_n^2 & y_n^2 \end{bmatrix} \quad \text{– macierz współczynników,}$$

$P = \text{diag} \{p_1 \ p_2 \ \dots \ p_n\}$  – macierz wag,

$h = [h_1 \ h_2 \ \dots \ h_n]^T$  – macierz pomierzonych wysokości punktów.

Estymacja parametrów wielomianu metodą najmniejszych kwadratów (8) oznacza, że na wynik znaczny wpływ mają również błędy grube, którymi w omawianym procesie filtracji są wszystkie punkty nie będące odbiciami od powierzchni terenu. Wpływ błędów grubych może być wyeliminowany dzięki estymacji odpornej. W metodzie tej parametry otrzymuje się w wyniku rozwiązania zadania minimalizacyjnego:

$$\sum_{i=1}^n p_i \rho(v_i) \rightarrow \min, \quad (7)$$

Wprowadzona tutaj funkcja jest tzw. funkcją straty, która może być interpretowana jako odchylenie estymatora od poszukiwanej optymalnej wartości tego estymatora. Celowym jest zatem te odchylenia minimalizować. Rozwiązując zadanie minimalizacyjne (7), otrzymuje się układ równań normalnych, który odpowiada iteracyjnemu rozwiązaniu zadania najmniejszych kwadratów:

$$\sum_{i=1}^n p_i w(v_i)^{(k-1)} v_i^2 \rightarrow \min. \quad (8)$$

Wprowadzona dodatkowa funkcja wagowa  $w(v_i)$  związana jest z funkcją straty i obliczana iteracyjnie w  $k$ -tym kroku iteracji na podstawie poprawek z poprzedniego kroku ( $k-1$ ). Funkcja ta nazywana jest często funkcją tłumienia, która modyfikuje dotychczasowe wagi obserwacji:

$$\bar{p}_i = p_i w(v_i). \quad (9)$$

Uwzględniając zmodyfikowaną macierz wag, estymowane parametry wyznacza się w  $k$ -tym kroku iteracji z zależności [Koch 1996]:

$$\hat{x}^k = (\mathbf{A}^T \bar{\mathbf{P}}^{(k-1)} \mathbf{A})^{-1} \mathbf{A}^T \bar{\mathbf{P}}^{(k-1)} \mathbf{h} . \quad (10)$$

Proces iteracyjny kończy się po uzyskaniu zgodności estymatora z poprzednim krokiem iteracji, na założonym poziomie dokładności.

Kluczowym zagadnieniem w estymacji odpornej jest wybór odpowiedniej funkcji tłumienia. Eksperymenty numeryczne przeprowadzone w pracy [Borkowski i Józków, 2006] pokazały, że spośród różnych funkcji tłumienia najlepsze rezultaty filtracji prezentowaną metodą otrzymuje się z wykorzystaniem funkcji Krausa [Kraus 2000]:

$$w(v) = \begin{cases} \frac{1}{1 + (\alpha|v - \delta|)^\beta}, & |v| > \sigma \\ 1, & |v| \leq \sigma \end{cases}, \quad (11)$$

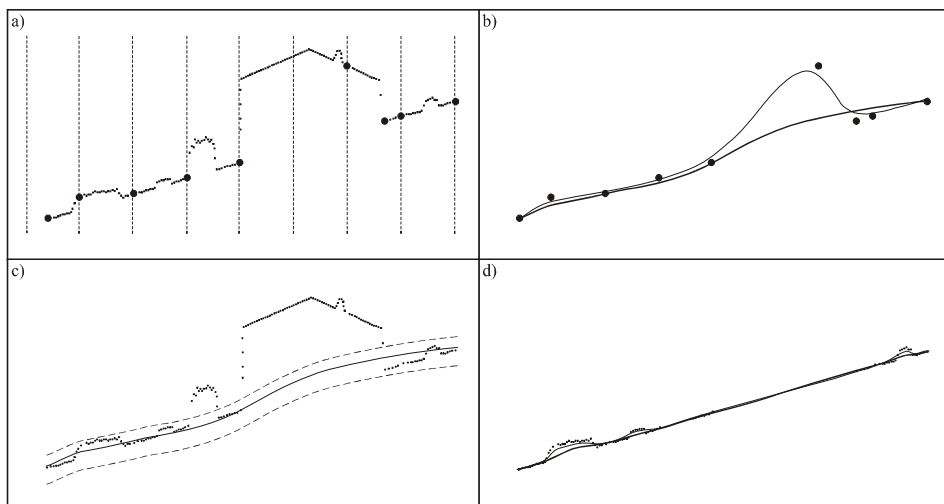
gdzie  $\delta$  jest dobieranym empirycznie parametrem determinującym zakres błędów przypadkowych. Dodatkowe parametry  $\alpha$  i  $\beta$ , również dobierane są empirycznie.

## FILTRACJA HIERARCHICZNA

W algorytmie wykorzystano ideę filtracji hierarchicznej zaproponowanej w pracy [Briese et al. 2002]. Rysunek 1 przedstawia w sposób schematyczny filtrację hierarchiczną dla profilu terenowego. Filtracja ta realizowana jest według następującego schematu:

- podział obszaru opracowania na mniejsze podobszary, dla których wybierany jest punkt reprezentatywny, najczęściej punkt najniższy (rys. 1a);
- odporna estymacja parametrów wielomianu ruchomego w punktach reprezentatywnych – stworzony zostaje model trendu nie uwzględniający lokalnych form ukształtowania terenu (rys. 1b);
- usunięcie ze zbioru punktów spoza ustalonego buforu trendu terenu (rys. 1c);
- odporna estymacja wielomianu ruchomego w punktach nie wyeliminowanych w poprzednim etapie (rys. 1d).

W algorytmie użytym w pracy – model hierarchiczny został rozszerzony do wieloetapowej filtracji hierarchicznej, w której opisane wyżej pierwsze trzy kroki następują po sobie kilkakrotnie, a wielkość bufora zmniejszana jest na każdym etapie.



Rys. 1. Schemat filtracji hierarchicznej [Briese et al. 2002]; objaśnienia w tekście  
 Fig. 1. Steps of the hierarchic filtration [Briese et al. 2002]; explanation in the text

Wieloetapowość pozwala na eliminację punktów obiektów znajdujących się coraz bliżej powierzchni terenu. Jak wykazały testy numeryczne, wadą jednoetapowej (najprostszy przypadek) filtracji hierarchicznej jest pozostawianie zbyt dużej liczby „punktów wysokich” lub eliminacja poprawnych punktów terenu w zależności od wielkości wybranego bufora dla trendu powierzchni terenu. Ostatni krok filtracji hierarchicznej (rys. 1d) ma na celu identyfikację punktów najniższych obiektów, np. niskiej roślinności.

## DANE TESTOWE ORAZ PROCEDURA WERYFIKACJI

Do oceny dokładności filtracji prezentowaną metodą wykorzystano oryginalne zbiory danych lotniczego skaningu laserowego [<http://www.itc.nl/isprswgIII-3/filtertest/Reference.zip>]. Jest to zestaw piętnastu specjalnie wyselekcjonowanych zbiorów będących fragmentami skaningu wykonanego skanerem Optech ALTM, gdzie rejestrowane były oba (pierwsze i ostatnie) odbicia. Każdy zestaw posiadał dane referencyjne w postaci zbiorów punktów terenu i punktów nie należących do powierzchni terenu (punktów obiektów). Obszary zestawów testowych zostały wybrane tak, by sprawdzić w jaki sposób automatyczne algorytmy wykonują filtrację danych zawierających różne obiekty, np. samochody, mosty, budynki, drzewa, rampy; oraz pozyskanych dla terenów o szczególnym ukształtowaniu, np. stromych zboczy, ostrych grzbietów, nieciągłych obszarów, przy których może nastąpić błędna klasyfikacja punktów. Szczegółowy opis próbek testowych oraz uzasadnienie ich wyboru podano w pracy [Stihole i Vosselman 2004].

Procedura weryfikacji polegała na porównaniu wyników filtracji metodą wielomianowych powierzchni ruchomych z danymi referencyjnymi i określeniu parametrów ilościowych tego porównania. W wyniku filtracji metodą wielomianowych powierzchni ruchomych otrzymano dla każdej z próbek testowych dwa zbiory punktów:

- punkty sklasyfikowane jako punkty terenu – zbiór A,
- punkty sklasyfikowane jako punkty obiektów – zbiór B.

Zbiór danych referencyjnych również podzielony został na:

- punkty terenu – zbiór C,
- punkty obiektów – zbiór D.

Wykonując działania na powyższych zbiorach, każdy z punktów klasyfikowany jest do jednego z następujących zbiorów:

- zbiór E – część wspólna zbiorów C i A ( $C \cap A$ ), są to punkty terenu zidentyfikowane prawidłowo (rys. 2, rys. 3, rys. 4 a – punkty jasnoszare);
- zbiór F – różnica zbiorów C i A ( $C \setminus A$ ), są to punkty terenu zidentyfikowane przez algorytm jako punkty obiektów (błąd filtracji typu pierwszego), (rys. 2 – ciemnoszare krzyże, rys. 3, rys. 4 a – punkty niebieskie);
- zbiór G – różnica zbiorów D i B ( $D \setminus B$ ) są to punkty obiektów zidentyfikowane przez algorytm jako punkty terenu (błąd filtracji typu drugiego), (rys. 2 – ciemnoszare trójkąty, rys. 3, rys. 4 a – punkty czerwone);
- zbiór H – część wspólna zbiorów D i B ( $D \cap B$ ), są to punkty obiektów zidentyfikowane prawidłowo (rys. 2 – punkty czarne, rys. 3, rys. 4a – punkty ciemnoszare).

Jakość filtracji wyraża procentowy udział błędnie sklasyfikowanych przez algorytm punktów w całkowitej ilości punktów. Wówczas błędy filtracji typu pierwszego  $\sigma_1$ , typu drugiego  $\sigma_2$  oraz całkowity  $\sigma$  obliczane są według zależności:

$$\sigma_1 = \frac{f}{e + f}, \quad (12)$$

$$\sigma_2 = \frac{g}{g + h}, \quad (13)$$

$$\sigma = \frac{f + g}{e + f + g + h}, \quad (14)$$

gdzie:

$e, f, g, h$  – liczebność odpowiednio zbiorów E, F, G, H.

## ANALIZA UZYSKANYCH WYNIKÓW

Uzyskane wyniki dla zestawów testowych zestawiono w tabeli 1.

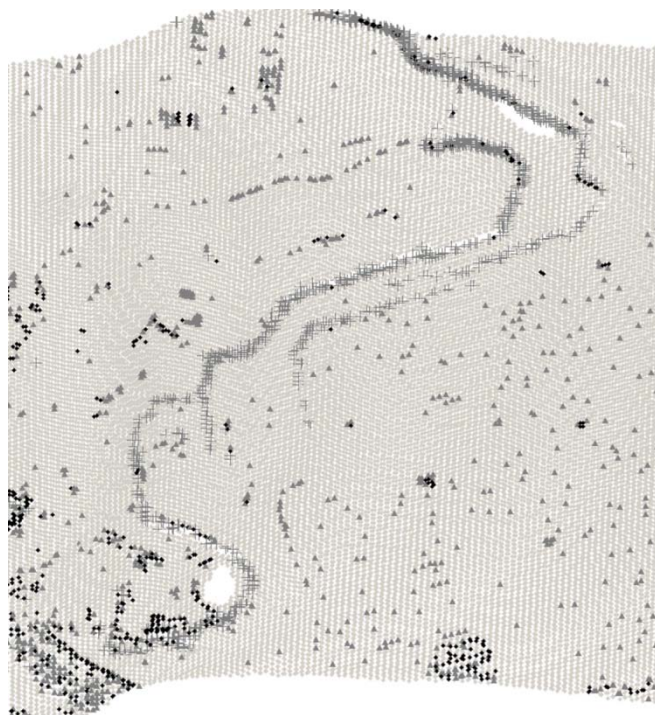
Poza jednym przypadkiem uzyskano skuteczność filtracji ponad 92%. W porównaniu z przeprowadzonymi testami ośmiu innych algorytmów [Sithole Vosselman 2004] we wszystkich przypadkach błędy całkowite filtracji metodą wielomianowych powierzchni ruchomych zbliżone są do błędu najlepszego z algorytmów, a niejednokrotnie są od niego mniejsze. Dla zestawu „Samp11” (teren zurbanizowany na stromym zboczu) uzyskano największy błąd typu pierwszego i całkowity. Spowodowane jest to trudnościami z poprawnym odróżnieniem przez algorytm punktów stromej zbocza od punktów obiektów.

Tabela 1. Wartości procentowych błędów filtracji;  $\sigma_{\min}, \sigma_{\max}$  – całkowite błędy procentowe odpowiednio najlepszego i najgorszego algorytmu [Sithole i Vosselman 2004]

Table 1. Percentage values of errors of filtering;  $\sigma_{\min}, \sigma_{\max}$  – total percentage errors of best and worst algorithm [Sithole i Vosselman 2004]

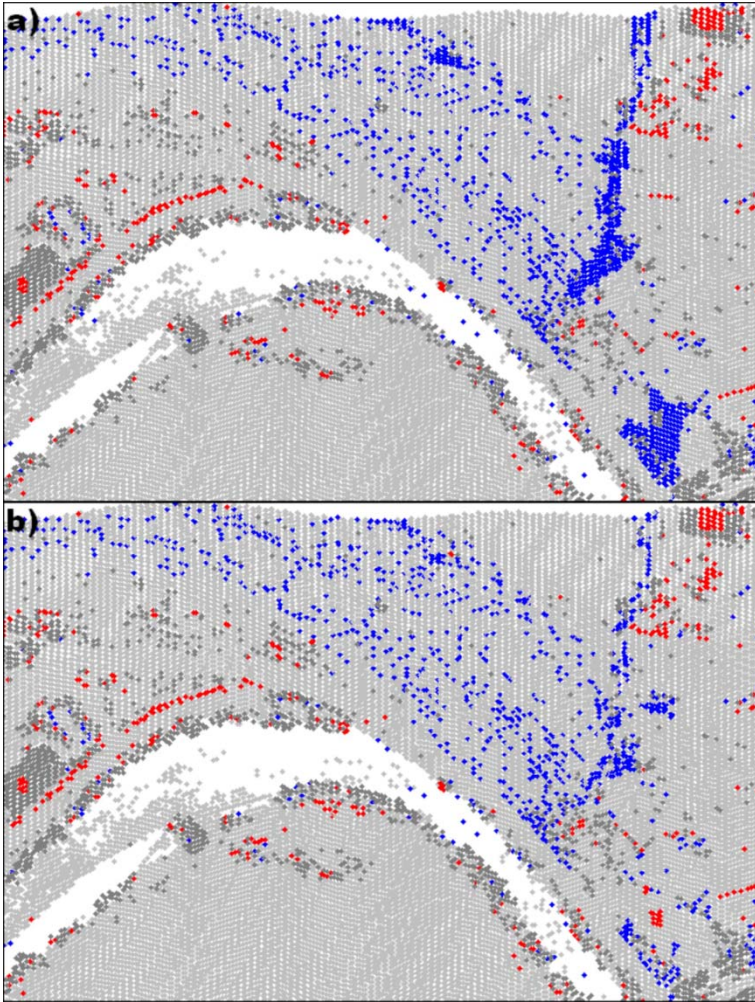
Zestaw testowy Testing sample	Liczba punktów Number of points	$\sigma_1$ [%]	$\sigma_2$ [%]	$\sigma$ [%]	$\sigma_{\min}$ [%]	$\sigma_{\max}$ [%]
Samp11	38010	14,01	8,83	11,80	10,76	36,96
Samp12	52119	6,32	1,53	3,99	3,25	16,28
Samp21	12960	0,75	3,30	1,32	2,57	9,84
Samp22	32706	4,83	8,16	5,87	3,63	23,78
Samp23	25095	5,58	5,86	5,71	4,00	27,80
Samp24	7492	6,02	7,73	6,49	4,42	36,06
Samp31	28862	0,93	1,64	1,26	1,80	12,92
Samp41	11231	4,30	3,04	3,67	8,76	23,67
Samp42	42470	2,96	2,07	2,33	1,62	6,38
Samp51	17845	2,32	3,57	2,59	2,72	22,81
Samp52	22474	6,35	17,36	7,51	3,07	45,56
Samp53	34378	5,07	26,21	5,92	8,91	52,81
Samp54	8608	7,85	3,39	5,31	3,23	23,89
Samp61	35060	1,81	4,31	1,90	2,08	35,87
Samp71	15645	1,25	5,59	1,74	1,63	34,98
Samp12 (3 punkty pewne) (3 fixed points)	52119	6,09	1,54	3,87	–	–
Samp52 (18 punktów pewnych) (18 fixed points)	22474	4,73	17,57	6,08	–	–

Większość znanych algorytmów skonstruowana jest w taki sposób, aby eliminować jak największą liczbę punktów obiektów, co niejednokrotnie powoduje również błędną klasyfikację sporej liczby punktów terenu. Tak jest również w przypadku testowanych w pracy [Sithole i Vosselman 2004] algorytmach, gdzie zazwyczaj procentowy błąd typu pierwszego jest większy niż typu drugiego. Dla metody wielomianowej okazuje się, że częściej procentowy błąd typu drugiego jest większy niż typu pierwszego. Uwzględniając doświadczenia z zakresu komercyjnego opracowania danych lotniczego skaningu laserowego, wydaje się być to cechą pozytywną, gdyż po filtracji automatycznej następuje jej ręczna korekta i łatwiej jest zauważyć punkty obiektów nie usunięte w procesie filtracji niż wyeliminowane niepotrzebnie ze zbioru punkty terenu. Duże wartości błędów typu drugiego dla niektórych testów często są wynikiem dysproporcji pomiędzy liczbą punktów terenu i obiektów. Dla zestawu „Samp53” prawie 96% punktów to punkty terenowe i jedynie 1398 punktów to punkty obiektów. Zatem błąd procentowy typu drugiego 26,21% to jedynie 364 błędnie sklasyfikowane punkty obiektów, a błąd typu pierwszego 5,07% to 1672 błędnie zidentyfikowane punkty terenu (rys. 2).



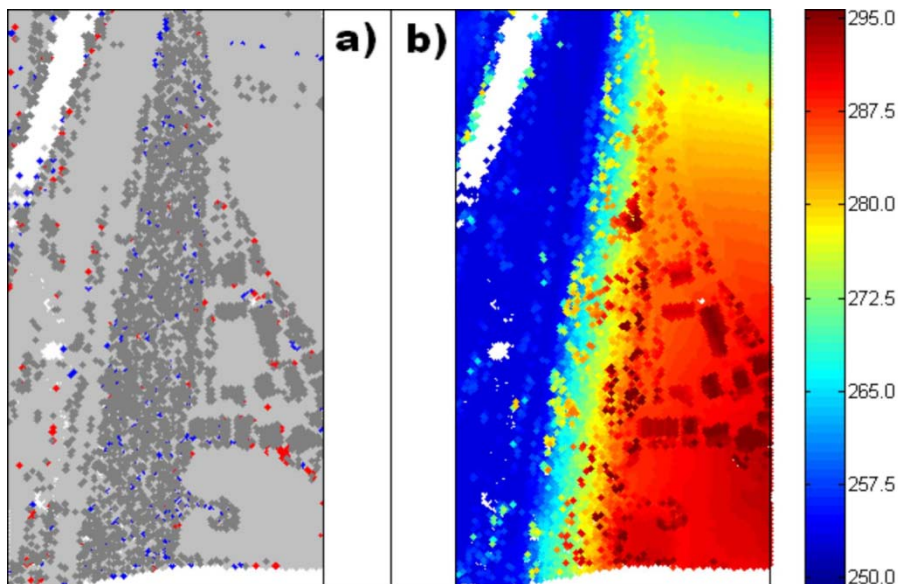
Rys. 2. Błędy filtracji: zestaw „Samp53”  
 Fig. 2. Filtering errors of sample „Samp53”

Przedstawiony algorytm pozwala na uwzględnienie w obliczeniach dodatkowej informacji *a priori* w postaci punktów pewnych terenu. Wybór punktów pewnych nie jest przypadkowy. Uzasadniony wydaje się wybór jedynie w tych miejscach, gdzie możliwe jest błędne sklasyfikowanie przez algorytm większych grup punktów terenu, np. wzdłuż wałów czy w miejscach o dużej przewadze liczbowej punktów niskich obiektów, które mogą nie zostać usunięte na etapie filtracji hierarchicznej. Jak wynika z przeprowadzonych testów, nawet niewielka liczba odpowiednio wybranych punktów pewnych powoduje zwiększenie skuteczności filtracji. Dla zestawu „Samp12” wybrano jedynie 3 punkty, które spowodowały poprawę klasyfikacji dodatkowo 62 punktów będących wcześniej błędami typu 1. Dla zestawu „Samp52” wybrano 18 punktów pewnych rozmieszczonych przede wszystkim na szczycie ostrego grzbietu. W tym przykładzie aż 326 dodatkowych punktów zostało poprawnie sklasyfikowanych jako punkty terenu (rys. 3). Z reguły, błędnie sklasyfikowane przez algorytm punkty porzucane są pojedynczo lub w małych grupach na całym obszarze, wówczas trudno uzyskać poprawę filtracji poprzez uwzględnienie punktów pewnych, jednak w pewnych przypadkach tylko kilka punktów pewnych może spowodować wzrost skuteczności filtracji nawet o 1%, co przy uzyskanych wynikach na poziomie 90% jest bardzo dobrym rezultatem. Na rysunku 4a przedstawiono rozmieszczenie błędów filtracji dla zestawu „Samp51” posiadającego różnorodne obiekty, np. drzewa, budynki, dużą skarpcę oraz przerwę (*gap*) pomiędzy sąsiednimi skanami. Rysunek 4b przedstawia natomiast oryginalne dane tego zestawu.



Rys. 3. Błędy filtracji: zestaw „Samp52”; a) bez punktów pewnych, b) z punktami pewnymi  
Fig. 3. Filtering errors of sample „Samp52”; a) without fixed points, b) with fixed points





Rys. 4. Zestaw „Samp51”; a) błędy filtracji, b) dane oryginalne (wysokości kodowane kolorem)  
 Fig. 4. Sample „Samp51”; a) filtering errors, b) original data (heights color coded)

## WNIOSKI

W pracy przedstawiono wyniki oceny poprawności filtracji danych lotniczego skaningu laserowego metodą wielomianowych powierzchni ruchomych. Na podstawie danych referencyjnych zostały określone procentowe błędy filtracji typu pierwszego, drugiego oraz całkowity. Porównując wyniki otrzymane z wykorzystaniem prezentowanej metody filtracji z innymi algorytmami [Sithole i Vosselman 2004], otrzymujemy pozytywną ocenę metodą wielomianowych powierzchni ruchomych. Zaletą metody jest również możliwość wykonania filtracji w podejściu hierarchicznym, a swobodny wybór parametrów, szczególnie wykładnika  $r$  funkcji wagowej (5) oraz parametrów  $\alpha$  i  $\beta$  funkcji tłumienia (11) pozwala na dobre dopasowanie modelowanej powierzchni do lokalnych struktur terenu. Algorytm ma możliwość uwzględnienia dodatkowej informacji *a priori* w postaci punktów, które są na pewno punktami terenowymi. Odpowiedni wybór punktów pewnych na niektórych formach ukształtowania terenu, np. wałach, powoduje znaczne podniesienie dokładności filtracji. Nie jest wymagana interpolacja danych pomiarowych, gdyż algorytm bazuje na danych oryginalnych. Algorytmu jest stosunkowo prosty, jednak konieczność wykonywania wielu kroków iteracji w odpornej estymacji parametrów wielomianu w każdym punkcie pomiarowym powoduje dużą ilość obliczeń numerycznych, co w przypadku opracowania większych zbiorów danych niesie z sobą konieczność użycia dużej mocy obliczeniowej.

**PIŚMIENNICTWO**

- Axelsson P., 1999. Processing of laser scanner data – algorithms and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2), 138–147.
- Axelsson P., 2000. DEM generation from laser scanner data using adaptive TIN models. *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIII-1/B4, 110–117.
- Borkowski A., 2004. Modellierung von Oberflächen mit Diskontinuitäten. *Deutsche Geodätische Kommission, Reihe C, Heft Nr 575*.
- Borkowski A., 2005. Filtracja danych lotniczego skaningu laserowego z wykorzystaniem metody aktywnych powierzchni. *Roczniki Geomatyki, tom III, zeszyt 4, PTIP*, 35–42.
- Borkowski A., Józków G., 2006. Wykorzystanie wielomianowych powierzchni ruchomych w procesie filtracji danych pochodzących z lotniczego skaningu laserowego. *Archiwum Fotogrametrii, Kartografii i Teledetekcji*, Vol. XVI, 2006, 63–73.
- Borkowski A., Keller W., 2006. An Attempt to ALS-data filtering in wavelet domain. 8th Bilateral Geodetic Meeting Poland-Italy. Wrocław, 22–24 June. <http://www.geo.ar.wroc.pl/8bgmpi/>
- Briese C., Pfeifer N., Dorninger P., 2002. Applications of the robust interpolation for DTM determination. Symposium ISPRS Commission III, Photogrammetric Computer Vision, Graz, 9–13 September 2002. *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIV / 3A, 55–61.
- Elmqvist M., Jungert E., Persson A. und Soderman U., 2001. Terrain modelling and analysis using laser scanner data. *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIV-3/W4, Annapolis, Maryland, 22–24 October, 219–227.
- Elmqvist M., 2002. Ground surface estimation from airborne laser scanner data using active shape models. ISPRS, Commission III, Symposium Photogrammetric Computer Vision, September 9–13, Graz, 114–118.
- Filin S., Pfeifer N., 2006. Segmentation of airborne laser scanning data using a slope adaptive neighborhood. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60, 71–80.
- Kraus K., 2000. *Photogrammetrie. Band 3. Topographische Informationssysteme*. Dümmler, Köln.
- Kraus K., Pfeifer N., 2001. Advanced DTM generating from LIDAR data. *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIV-3/W4, Annapolis, Maryland, 22–24 October, 23–30.
- Koch K.-R. 1996. Robuste Parameterschätzung. *AVN*, 103, 1–18.
- Marmol U., Jachimski J., 2004. A FFT based method of filtering airborne laser scanner data. ISPRS Congress, 12–23 July, Istanbul, Turkey, Commission 3. <http://www.isprs.org/commission3/wg3>
- Roggero M., 2001. Airborne laser scanning: Clustering in row data. *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIV-3/W4, Annapolis, Maryland, 22–24 October, 227–232.
- Sithole G., 2001. Filtering of laser altimetry data using a slope adaptive filter. *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIV-3/W4, Annapolis, Maryland, 22–24 October, 203–210.
- Sithole G., Vosselman G., 2004. Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vo. 59, 85–101.
- Vosselman G., Maas, H.-G., 2001. Adjustment and filtering of raw laser altimetry data. OEEPE-Workshop on Airborne Laserscanning and Interferometric SAR for Detailed Digital Elevation Models, Stockholm, 1–3 march.

## **AIRBORNE LASER SCANNING DATA FILTERING USING MOVING POLYNOMIAL SURFACE MODEL – THE VERIFICATION OF THE METHOD**

**Abstract.** For the digital terrain modelling from airborne laser scanning data the identification of points that are reflections from bare earth is the main issue of process of elaboration ALS data. In this work the hierarchic method of identification of points belonging to terrain surface was presented. This method is based upon the approximation of data using moving polynomial surface. Moving polynomial parameters are estimated locally based upon M-estimators of robust estimation method. In the estimation process the depended on the distance weighting function and asymmetrical damping function were used. A lot of numeric tests on the real airborne laser scanning data were executed. This data had a form of 15 testing samples contained referenced data as correctly classified terrain and objects' points. The method was verified based upon the comparison of data after automatic filtration with referenced sets of points. In the result of comparison the percentage errors of automatic filtering were determined. Total percentage filtering error was evaluated on the level from about 1% to about 12%. These values depended mainly on the terrain form and terrain coverage. It has been noticed that in the critical places e. g. along the dykes the including to the algorithm additional information a-priori as correct terrain points make the automatic filtering more accurate.

**Key words:** airborne laser scanning, filtering, moving polynomial surface, robust estimation

Zaakceptowano do druku – Accepted for print: 30.06.2008

Do cytowania – For citation: Barańska A., Józków G., 2008. Filtracja danych lotniczego skaningu laserowego metodą ruchomych powierzchni wielomianowych – weryfikacja metody. *Acta Sci. Pol. Geodesia et Descriptio Terrarum*, 7(2), 15–27.



## **NUMERYCZNY SPOSÓB WYZNACZANIA CAŁKOWITYCH WARTOŚCI NIEOZNACZONOŚCI POMIARU FAZOWEGO**

Sławomir Cellmer, Jacek Rapiński

Uniwersytet Warmińsko-Mazurski w Olsztynie

**Streszczenie.** W artykule przedstawiony opracowany przez autorów sposób poszukiwania całkowitych wartości nieoznaczoności pomiaru fazowego. Sposób ten jest pewną numeryczną techniką poszukiwania optymalnego rozwiązania w teorii estymacji wartości całkowitych metodą najmniejszych kwadratów (theory of integer least squares). Zaprezentowano podstawy teoretyczne oraz przykłady numeryczne ilustrujące działanie metody.

**Słowa kluczowe:** pomiar fazowy GPS, ambiguity

### **WPROWADZENIE**

Wyznaczenie poprawnych, całkowitych wartości nieoznaczoności (ambiguity) jest kluczowym elementem opracowania pomiarów fazowych GPS [Han, Rizos 1997]. Po wyznaczeniu ambiguity dalszy etap opracowania jest stosunkowo prosty, zbliżony do klasycznego zadania przestrzennego, liniowego wcięcia wstecz. Różne ośrodki naukowo-badawcze proponowały własne metody wyznaczenia całkowitych wartości nieoznaczoności. W pracy [Hatch 2000] została przedstawiona klasyfikacja metod wyznaczenia całkowitych wartości ambiguity, oparta na rozróżnieniu rodzajów informacji wykorzystywanych w tym procesie. W tej klasyfikacji rozróżniono trzy grupy metod:

1. Wyznaczenie ambiguity w obszarze danych obserwacyjnych
2. Techniki poszukiwawcze (search techniques) w obszarze współrzędnych
3. Techniki poszukiwawcze w obszarze ambiguity.

Metody z pierwszej grupy oparte są na analizie kombinacji liniowych różnych typów obserwacji (pseudoodległości i pomiary fazowe na różnych częstotliwościach).

Do drugiej grupy należy Metoda Funkcji Ambiguity (Ambiguity Function Method – AFM), poszukująca ostatecznego rozwiązania-współrzędnych punktu, przy jednoczesnym ograniczeniu obszaru poszukiwań (dzięki pewnej funkcji matematycznej)

do takich rozwiązań, dla których ambiguity przyjmują wartości całkowite [Remondi 1991].

Trzecia grupa opiera się na teorii estymacji całkowitych wartości przy spełnieniu kryterium minimalizacji kwadratów poprawek (integer least squares) przedstawionej w pracy: [Teunissen 1993].

W metodach z trzeciej grupy proces estymacji składa się z trzech etapów:

1. Klasyczna estymacja metodą najmniejszych kwadratów bez spełnienia warunku całkowitych wartości dla ambiguity (float solution)
2. Estymacja całkowitych wartości ambiguity
3. Estymacja metodą najmniejszych kwadratów przy stałych wartościach ambiguity uzyskanych w drugim etapie (fixed solution).

Obecnie najbardziej znaną i uznawaną za najlepszą jest metoda LAMBDA (The Least-Squares AMBIGuity Decorrelation Adjustment) [Teunissen 1993, 1998].

W metodzie tej wykorzystywane są przybliżone wartości ambiguity oraz ich macierz kowariancji, uzyskane z pierwszego etapu estymacji (float solution). Poszukiwania rozwiązania dotyczą obszaru ograniczonego elipsoidą (hiperelipsoidą) jednakowego (przyjętego) prawdopodobieństwa (poziomu ufności). W celu zwiększenia efektywności procesu poszukiwania rozwiązania wykonuje się transformację powodującą zmianę kształtu obszaru poszukiwań (zmniejszenie dysproporcji między długościami osi elipsoidy ufności). Transformacja ta musi spełniać warunek otrzymania całkowitych wartości rozwiązania przy całkowitych wartościach elementów transformowanych.

Przedmiotem tego artykułu jest etap poszukiwania rozwiązania ambiguity przy spełnieniu ogólnych założeń teorii estymacji całkowitych wartości metodą najmniejszych kwadratów. W odróżnieniu od metody LAMBDA, w przedstawionym niżej podejściu, nie przewiduje się ograniczania obszaru poszukiwań do elipsoidy ufności ani wykonywania transformacji.

## PRZESTRZEŃ ROZWIĄZAŃ AMBIGUITY

W teorii estymacji całkowitych wartości metodą najmniejszych kwadratów rozwiązanie jest wektorem liczb całkowitych, minimalizującym następującą formę kwadrata:

$$\Phi = (a - a_{\text{float}})^T Q_a (a - a_{\text{float}}), \quad (1)$$

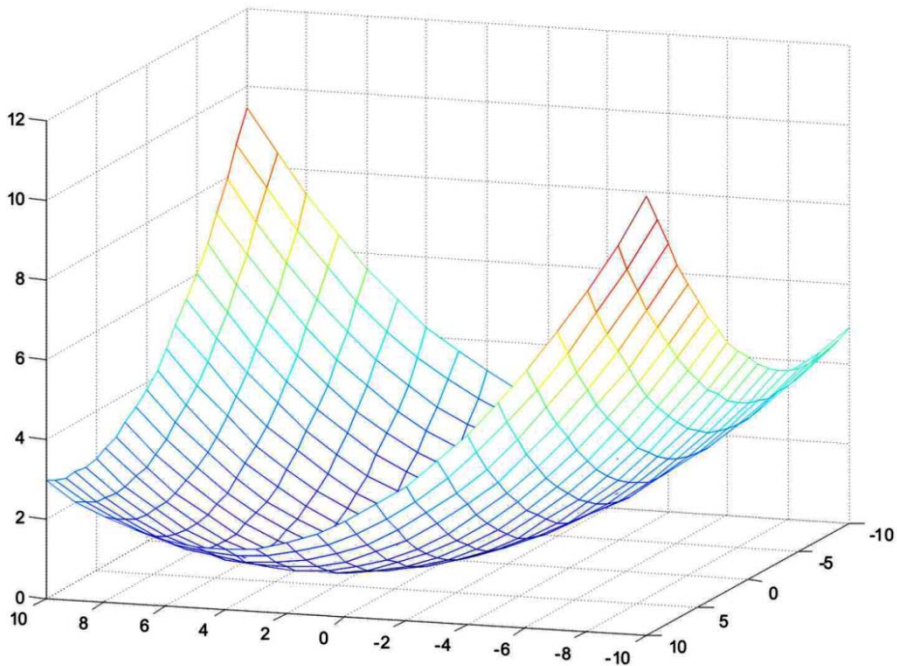
gdzie:

$a$  – poszukiwane rozwiązanie całkowite,

$a_{\text{float}}$  – wektor nieoznaczoności w postaci liczb rzeczywistych z pierwszego równania,

$Q_a$  – macierz kowariancji nieoznaczoności po pierwszym równaniu.

Zbiór możliwych rozwiązań możemy zinterpretować jako węzły  $n$ -wymiarowej sieci o jednostkowych oczkach. Na rys. 1 jest przedstawiony poglądowy wykres zbioru rozwiązań nieoznaczoności pomiaru fazowego dla przypadku dwóch wartości niewiadomych.



Rys. 1. Graficzna interpretacja zbioru rozwiązań ambiguitę dla przypadku wektora dwuelementowego

Fig. 1. Graphic interpretation of solution set for case of two component vector

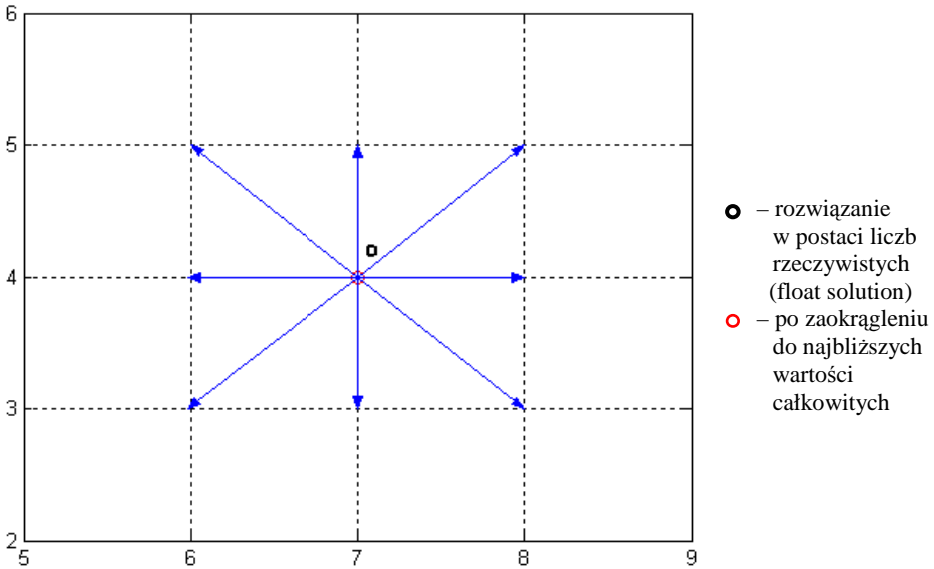
Na osiach poziomych oznaczone są wartości nieoznaczoności pomiaru fazowego, a na osi pionowej odpowiadające im wartości formy kwadratowej (1).

## NUMERYCZNA METODA POSZUKIWANIA ROZWIĄZANIA

Proponowana metoda polega na sprawdzaniu wartości formy kwadratowej (1) w sąsiednich węzłach sieci, wokół aktualnej pozycji. Następnie jako aktualną pozycję przyjmuje się węzeł, w którym wartość  $\Phi$  jest najmniejsza i wykonywana jest następna iteracja.

Pozycję wyjściową określa się przez zaokrąglenie nieoznaczoności uzyskanych z pierwszego etapu opracowania (float solution) do liczb całkowitych (najbliższego węzła sieci).

Rysunek 2 ilustruje sytuację na początku procesu poszukiwania rozwiązania dla przypadku dwuelementowego wektora nieoznaczoności pomiaru fazowego.



Rys. 2. Pierwszy krok procesu poszukiwania całkowitych wartości nieznaczoneści pomiaru fazowego

Fig. 2. First step of search process of integer value of phase's measurement ambiguity

Niebieskie strzałki wskazują miejsca, w których poszukiwane jest rozwiązanie w jednym kroku iteracyjnym. Wartości ambiguity w sąsiednich punktach sieci możemy określić za pomocą sumy wektorów:

$$a_{ij} = a_0 + a_j, \quad (2)$$

gdzie:

- $a_0$  – wektor wartości ambiguity po zaokrągleniu do najbliższych liczb całkowitych,
- $a_j$  – wektor wartości ambiguity po przesunięciu do  $j$ -tego sąsiedniego węzła sieci.

W przedstawionym na rys. 2 przykładzie wektor  $a_0$  ma postać:  $a_0 = [7 \ 4]^T$ , a wektory  $a_j$  są kolumnami macierzy:

$$D_{2\text{dim}} = \begin{bmatrix} -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 \end{bmatrix} \quad (3)$$

Dla przypadku trzelementowego wektora nieznaczoneści, macierz  $D_{3\text{dim}}$  której kolumny reprezentują wektory translacji do sąsiednich węzłów sieci, ma następującą postać:



$$D_{3dim} = \begin{bmatrix} -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 \\ -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 \end{bmatrix} \quad (4)$$

W analogiczny sposób można zbudować macierz utworzoną z wektorów translacji dla dowolnego wymiaru wektora nieoznaczoności:

$$D_{ndim} = \begin{bmatrix} -\underline{I} & \underline{0} & \underline{I} \\ D_{(n-1)dim} & D_{(n-1)dim} & D_{(n-1)dim} \end{bmatrix}, \quad (5)$$

gdzie:

$n$  – wymiar wektora nieoznaczoności,

$\underline{I}$  – wierszowy wektor jednostkowy o wymiarze  $3^{(n-1)}$ ,

$\underline{0}$  – wierszowy wektor zerowy o wymiarze  $3^{(n-1)}$ .

Macierz  $D_{ndim}$  jest tworzona w sposób rekurencyjny wg formuły (5) z pierwszą macierzą pustą dla przypadku zerowego wymiaru:  $D_{0dim}=[]$ .

Proces obliczeniowy przebiega wg następującej formuły:

1. Określenie punktu startowego przez zaokrąglenie wartości wektora  $a_{float}$  do najbliższych wartości całkowitych:  $a_0 = \text{round}(a_{float})$
2. Iteracyjne poszukiwanie wartości całkowitych nieoznaczoności, minimalizujących formę kwadratową (1):

$$a_i = a_{i-1} + a_j(\Phi_{min}), \quad (6)$$

gdzie:

$i$  – nr iteracji

$a_i$  – wektor całkowitych nieoznaczoności w  $i$ -tej iteracji

$a_j(\Phi_{min})$  – wektor translacji (kolumna macierzy  $D_{ndim}$ ), dla którego suma (6) minimalizuje wartość formy kwadratowej  $\Phi(a_i)$

### PRZYKŁADY

Dane do pierwszego przykładu zostały zaczerpnięte z książki [Teunissen 1993]:

$$a_{float} = \begin{bmatrix} 2.97 \\ 3.10 \\ 5.45 \end{bmatrix} \quad Q_a = \begin{bmatrix} 6.288 & 2.340 & 0.544 \\ 2.340 & 6.292 & 5.978 \\ 0.544 & 5.978 & 6.290 \end{bmatrix}$$

Na podstawie powyższych danych można określić punkt startowy:  $a_0 = [3 \ 3 \ 5]^T$ .

W tabeli przedstawione są wartości nieoznaczoności dla węzłów sieci sąsiadujących z  $a_0$  wraz z odpowiadającymi im wartościami  $\Phi$ :

$a_{ij}$	2	2	2	2	2	2	2	2	2	3	3	3	3	3
	2	2	2	3	3	3	4	4	4	2	2	2	3	3
	4	5	6	4	5	6	4	5	6	4	5	6	4	5
$\Phi$	3.98	2.23	22.73	30.24	4.78	1.57	82.14	32.96	6.04	1.03	6.18	33.58	19.81	1.25

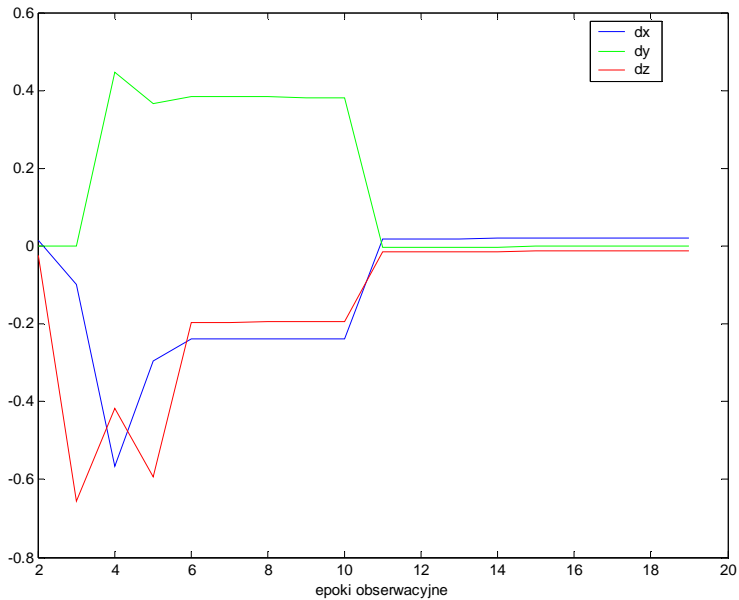
  

$a_{ij}$	3	3	3	3	4	4	4	4	4	4	4	4	4	4
	3	4	4	4	2	2	2	3	3	3	4	4	4	4
	6	4	5	6	4	5	6	4	4	6	4	5	6	6
$\Phi$	4.94	64.22	21.94	1.92	0.59	12.64	46.94	11.89	0.22	10.81	48.81	13.43	0.31	0.31

Wyróżnionemu wektorowi odpowiada najmniejsza wartość  $\Phi$ . Kolejna iteracja nie przyniosła poprawy wyników. Ostatecznym rozwiązaniem jest więc wektor  $a_i = [4 \ 3 \ 5]^T$ . Otrzymany wynik jest zbliżony z rozwiązaniem uzyskanym przez Teunissen za pomocą metody LAMBDA.

Drugi przykład ilustruje wyniki opracowania krótkiego (8 km) wektora GPS, pomierzonego metodą statyczną, w którym do określenia całkowitych wartości nieoznaczoności wykorzystano przedstawioną wyżej metodę. Dane obserwacyjne z 50-minutowej sesji statycznej zostały opracowane za pomocą programu GPPS firmy Ashtech. Interwał między kolejnymi epokami wynosił 20 sekund. Następnie ten sam zbiór obserwacyjny został opracowany za pomocą oprogramowania stworzonego przez autorów [Cellmer 2006], w którym do określania całkowitych wartości nieoznaczoności podwójnie zróżnicowanego pomiaru fazowego z jednej częstotliwości wykorzystano proponowaną metodę. Opracowanie wykonano w kilkunastu wariantach różniących się wielkością zbioru obserwacyjnego. Pierwszy wariant zawierał obserwacje z dwóch pierwszych epok. W każdym kolejnym wariantcie zbiór obserwacyjny był powiększany o obserwacje z kolejnej epoki. Wyniki z tego opracowania porównywano z wynikami otrzymanymi z opracowania całego zbioru obserwacyjnego za pomocą programu GPPS. Wielkości różnic z obu opracowań przedstawione są na rysunku 3.

Na podstawie wykresu można stwierdzić, że wyniki ustabilizowały się po 12 epoce obserwacyjnej, a więc po 4 min obserwacji.



Rys. 3. Różnice współrzędnych z dwu opracowań (wartości na osi pionowej są przedstawione w metrach)

Fig. 3. Coordinate differences from two solutions (values of vertical axis in metres)

## PODSUMOWANIE

Przedstawiona w artykule metoda stanowi proste i efektywne narzędzie do określania całkowitych wartości nieoznaczoności pomiaru fazowego.

## PIŚMIENNICTWO

- Cellmer S., 2006. Algorithm of GPS baseline processing in real time mode for engineering surveys. Reports on Geodesy, Nr 3 (74), 211–318.
- Han S., Rizos C., 1997. Comparing GPS Ambiguity Resolution Techniques, GPS World, 8, p. 54.
- Hatch R.R., 2000. GPS Carrier-Phase Ambiguity Resolution, Nav Com Technology Inc. A John Deere Company.
- Remondi B.W., 1991. Pseudo-Kinematic GPS Results Using the Ambiguity Function Method, Navigation, Spring.
- Teunissen P.J.G., 1993. Least Squares Estimation of the integer GPS ambiguities, Invited Lecture, Section IV Theory and Methodology, IAG General Meeting, Beijing.
- Teunissen P.J.G., Kleusberg A., 1998. GPS for Geodesy, 2nd edition Springer, Berlin, Heidelberg, New York.

## NUMERIC APPROACH OF ESTIMATION OF INTEGER VALUE OF PHASE'S MEASUREMENTS

**Abstract.** The self developed by authors approach of integer value of ambiguity search for phase's measurements is presented in the paper. This approach is one of the numeric techniques of optimal solution search in theory of integer least squares. Theoretical bases as well as numeric examples illustrating of method are presented.

**Key words:** GPS phase's measurements, ambiguity

Zaakceptowano do druku – Accepted for print: 30.06.2008

Do cytowania – For citation: Cellmer S., Rapiński J., 2008. Numeryczny sposób wyznaczania całkowitych wartości nieoznaczoności pomiaru fazowego. *Acta Sci. Pol. Geodesia et Descriptio Terrarum*, 7(2), 29–36.

## **BADANIE DOKŁADNOŚCI NMT INTERPOLOWANEGO NA PODSTAWIE DANYCH LOTNICZEGO SKANINGU LASEROWEGO SYSTEMU SCALARS\***

Piotr Gołuch, Andrzej Borkowski, Grzegorz Józków

Uniwersytet Przyrodniczy we Wrocławiu

**Streszczenie.** Dokładność Numerycznego Modelu Terenu (NMT), interpolowanego na podstawie danych lotniczego skaningu laserowego, zależy od wielu czynników, m.in. od ukształtowania terenu, pokrycia terenu, stabilności nalotu fotogrametrycznego, jakości danych nawigacyjnych i dokładności kalibracji, terenowej wielkości śladu plamki promienia lasera (wysokości lotu i zbieżności wiązki), gęstości pozyskanych punktów, zastosowanej metody filtracji danych. W pracy przedstawiono ocenę dokładności NMT zrealizowanego dla 20-kilometrowego odcinka doliny rzeki Widawy na potrzeby modelowania hydrodynamicznego. Skaningu laserowego wykonano został prototypowym skanerem ScaLARS, skonstruowanym w Instytucie Nawigacji Uniwersytetu w Stuttgarcie. Do rejestracji sygnału INS i GPS wykorzystano system Applanix POS/AV 510. Nalot wykonano samolotem AN-2, z wysokości 550 m. Terenowa wielkość śladu plamki lasera to około 0.6 m. Kalibrację systemu wykonano semi-automatycznie, uzyskując błąd bezwzględny w odniesieniu do obszarów kontrolnych, pomierzonych techniką GPS na poziomie 0.3 m wzdłuż i w poprzek do kierunku lotu oraz błąd wysokości 0.1 m. Badanie dokładności zbudowanego NMT przeprowadzono w oparciu o dane pozyskane z pomiaru terenowego technikami GPS i tachimetryczną. Wykonano pomiar na czterech reprezentatywnych obszarach obiektu badawczego. Filtrację danych przeprowadzono automatycznie z wykorzystaniem własnych algorytmów, bazujących na odpornej aproksymacji danych ruchomą powierzchnią wielomianową. W zależności od ukształtowania i pokrycia terenu uzyskano dokładności wysokościowe NMT od 0.17 m do 0.46 m.

**Słowa kluczowe:** lotniczy skaningu laserowy, NMT, ScaLARS

---

\* Praca naukowa finansowana ze środków na naukę w latach 2005–2007 jako projekt badawczy nr 4T12E01729. Obliczenia wykonano w systemie MATLAB, licencja nr 101979, grant obliczeniowy Wrocławskiego Centrum Sieciowo-Superkomputerowego.

## WSTĘP

Dominującym źródłem danych do budowy Numerycznych Modeli Terenu (NMT) staje się w ostatnich latach lotniczy skaningu laserowy. Technika ta zapewnia zarówno wysoką dokładność, jak i szczegółowość numerycznego opisu powierzchni terenu, zwłaszcza na terenach zalesionych, gdzie pozyskiwanie danych innymi technikami jest praktycznie niemożliwe. Dokładność produktu finalnego (NMT), rozumiana jako różnica wysokości pomierzonej bezpośrednio w terenie techniką klasyczną a wysokości wyinterpolowanej z NMT, zależy od trzech czynników: dokładności danych źródłowych skaningu, zastosowanej metody filtracji i metody interpolacji NMT. Głównym czynnikiem determinującym dokładność NMT jest oczywiście dokładność danych źródłowych skaningu, na którą wpływa wiele czynników, m.in. stabilność nalotu fotograficznego, jakość danych nawigacyjnych, dokładność kalibracji, wielkość śladu plamki promienia lasera (wysokości lotu i zbieżności wiązki) na powierzchni terenu, ukształtowanie terenu oraz pokrycie terenu [Gołuch et al. 2007]. Udział poszczególnych czynników trudno jest ocenić i dlatego należy rozpatrywać ich ogólny wpływ na dokładność produktu końcowego.

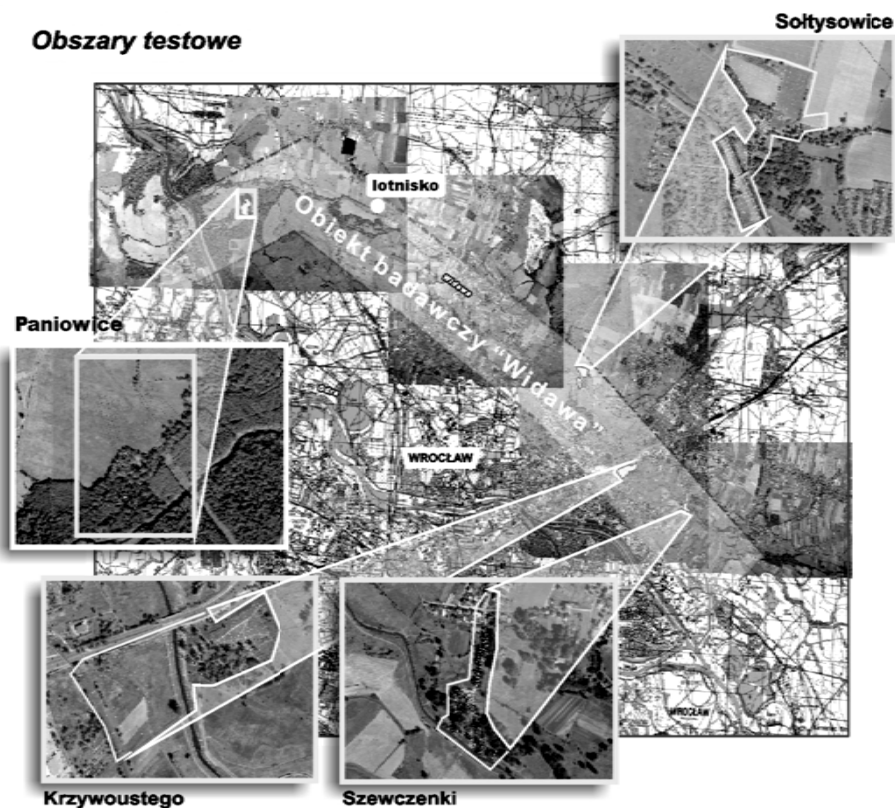
W publikacji [Gołuch i in. 2007] przedstawiono ocenę dokładności danych lotniczego skaningu laserowego systemu ScaLARS [Borkowski et al. 2006], w niniejszej pracy poddano ocenie dokładnościowej produkt końcowy w postaci NMT. Dane skaningu pozyskano w ramach realizacji projektu badawczego, dotyczącego budowy NMT doliny rzeki Widawy na potrzeby modelowania hydrodynamicznego. Dolina rzeki Widawy jest w znacznej części użytkowana rolniczo, natomiast jej ujściowy odcinek to obszar zalesiony. Występują tu liczne zagłębienia i wzniesienia powierzchni terenu. Filtrację danych przeprowadzono automatycznie z wykorzystaniem własnych algorytmów, bazujących na odpornej aproksymacji danych ruchomą powierzchnią wielomianową [Borkowski i Józków 2006]. Oprócz ilościowej oceny dokładności NMT przedstawiono w końcowej części pracy wyniki weryfikacji zbudowanego NMT w modelu hydrodynamicznym.

## OBIEKT BADAWCZY WIDAWA

Badania związane z oceną dokładności NMT interpolowanego na podstawie danych lotniczego skaningu laserowego ScaLARS były prowadzone na obiekcie badawczym „Widawa” [Borkowski et al. 2006]. Obiekt położony jest na północ od Wrocławia i swoim zasięgiem obejmuje ok. 40 km<sup>2</sup> doliny rzeki Widawy (rys. 1).

Skanowanie laserowe terenu zostało przeprowadzone w listopadzie 2005 r. z pokładu samolotu AN-2, z wysokości 550 m, przy użyciu systemu ScaLARS. Z uwagi na fakt, że obiekt badawczy (dolina rzeki) posiada szerokość ok. 2 km, skanowanie wykonano w 11 pasach. Szerokość jednego pasa skanowania w terenie wyniosła ok. 280 m, natomiast wzajemne pokrycie punktów pozyskanych z dwóch sąsiednich pasów skanowania zaplanowano na poziomie 30%.

Jednym z kryterium oceny poprawności zrealizowanego skanowania jest brak pustych obszarów (gaps) pomiędzy poszczególnymi pasami skanowania. Świadczy to o tym, że skanowanie zostało zrealizowane zgodnie z poprawnie wykonanym projektem lotu, oraz o stabilności i precyzji lotu. Czynniki te mają znaczny wpływ na ostateczną dokładność wyników skanowania.



Rys. 1. Objekt badawczy „Widawa” z zaznaczonymi obszarami testowymi [Gołuch et al. 2007]  
 Fig. 1. Research area „Widawa” with marked testing areas [Gołuch et al. 2007]

Pierwsze obliczenia, związane z kalibracją wyników skanowania laserowego, przeprowadzone zostały przy użyciu oprogramowania LASCAL [Schiele et al. 2005]. Kalibracja została wykonana w oparciu o pola kontrolne, które pomierzono w terenie techniką GPS. Parametry kalibracji wewnętrznej skanera oraz parametry kalibracji wyników pomiaru w układzie lokalnym w odniesieniu do pól kontrolnych podano w pracy [Borkowski i Jóźków 2006]. W tabeli 1 zamieszczono ostateczne wartości charakteryzujące dokładność kalibracji, tj. obliczone odchyłki wzdłuż i w poprzek do kierunku lotu oraz obliczone odchyłki wysokościowe, jak również średnią wartość błędu bezwzględnego wzdłuż i w poprzek do kierunku lotu oraz składową wysokościową.

W celu oszacowania dokładności NMT interpolowanego na podstawie danych pozyskanych metodą lotniczego skaningu laserowego wytypowano na obiekcie badawczym cztery reprezentatywne obszary testowe o zróżnicowanym pokryciu terenu (rys. 1): Paniowice, Sołtysowice, Krzywoustego i Szewczenki. Nazwy obszarom badawczym nadano od nazw osiedli bądź ulic, w okolicach których są one zlokalizowane. Wytypowane obszary testowe znajdują się na krańcach obiektu badawczego oraz w jego części środkowej. Taki dobór położenia tych obszarów był podyktowany tym, aby ich lokali-

zacja była w znacznej odległości od płyty lotniska, względem której wykonano kalibrację systemu. Łącznie na obszarach testowych pomierzono 1728 punktów. Pomiar przeprowadzono bezpośrednio w terenie, wykorzystując technikę pomiaru GPS-RTK i metodę tachimetryczną. Ze względu na fakt, że dolina rzeczna charakteryzuje się dużą różnorodnością ukształtowania i pokrycia terenu, analizę przeprowadzono dla czterech kategorii terenu:

- las i zadrzewienia – teren o bardzo urozmaiconym ukształtowaniu, z licznymi lokalnymi obniżeniami i oczkami wodnymi;
- zarośla i wysoka trawa – obszar wzdłuż koryta rzeki, poza lasami, porośnięty zaroślami i bardzo wysoką trawą, z licznymi skarpami;
- łąki i pola orne – teren użytkowany rolniczo, generalnie obszar płaski; do tej kategorii włączono również pastwiska i nieużytki ze średnio wysoką trawą (wysokość ok. 40–60 cm);
- drogi – są to drogi gruntowe, brukowe i asfaltowe, wzdłuż tych dróg często występują drzewa i zarośla, także są rowy i nieduże skarpy.

Tabela 1 Zestawienie parametrów dokładności kalibracji względem pól kontrolnych  
Table 1. Parameters of calibration accuracy for the sake of control fields

	Wzdłuż kierunku lotu Along flight direction	Poprzecznie do kierunku lotu Across flight direction	Składowa wysokościowa Height component
Obliczone odchyłki Calculated deviations	± 0.6 m	± 0.4 m	± 0.15 m
Błąd bezwzględny (średnia wartość) Absolute error (mean value)	0.3 m	0.3 m	0.10 m

## POMIARY KONTROLNE I OCENA DOKŁADNOŚCI

Punkty pomierzone bezpośrednio w terenie zostały przyjęte jako bezbłędne dane referencyjne, gdyż błędy określenia ich wysokości są o rząd wielkości mniejsze niż błędy określenia wysokości punktów skaningu. Do pomiarów terenowych wykorzystano geodezyjne odbiorniki GPS Ashtech Z-Xtreme oraz elektroniczny tachimetr Leica TCR 407. Pomiar terenowy polegał na określeniu położenia punktów równomiernie rozmieszczonych na obszarach testowych z uwzględnieniem charakterystycznych punktów ukształtowania terenu. Większość punktów pomierzona została tachimetrycznie w oparciu o zamarkowane w terenie własne punkty osnowy. Współrzędne punktów osnowy wyznaczono metodą GPS. W celu zapewnienia dużej dokładności pomiarów GPS – na każdym punkcie osnowy uwzględniane były poprawki referencyjne ze stacji WROC permanentnych obserwacji GPS/GLONASS. Poprawki uzyskiwano bezpośrednio w terenie poprzez mobilny Internet i połączenie odbiornika GPS z laptopem. Niewielka odległość pól testowych od referencyjnej stacji GPS pozwoliła na wyznaczenie współrzędnych punktów osnowy (w szczególności wysokości) z dokładnością rzędu 1 cm.

Ponieważ znaczną część obszarów testowych stanowią tereny zadrzewione i zakrzewione, dlatego pomiary terenowe zostały wykonane przede wszystkim metodą tachimetryczną. Jedynie na obszarze testowym Paniowice dodatkowo wykonano na



terenie odkrytym pomiary GPS-RTK, gdzie umieszczony na punkcie osnowy pomiarowej odbiornik stał się stacją referencyjną transmitującą poprawki poprzez radiomodem.

W celu zapewnienia jednolitego układu współrzędnych dla pomiarów terenowych i skaningu laserowego współrzędne punktów osnowy przeliczone zostały do tego samego układu 1992 (współrzędne  $x, y$ ) i Kronsztad 1986 (wysokości  $h$ ), w którym wyrażone były dane skaningu. Przeliczenie współrzędnych punktów do układów PUWG 1992 i Kronsztad 1986 wykonane zostało za pomocą programu Transpol i własnych procedur napisanych w środowisku MATLAB.

W tabeli 2 przedstawiono liczbę punktów pomierzonych na każdym z obszarów testowych.

Tabela 2. Liczba punktów pomierzonych bezpośrednio w terenie  
Table 2. Number of points measured using GPS and total-station

Obszar testowy Testing area	Krzywoustego	Sołtysowice	Paniowice	Szewczenki
Liczba punktów osnowy Number of control points	2	4	7	7
Liczba punktów tachimetrycznych Number of total-station points	467	760	115	514
Liczba punktów GPS-RTK Number of GPS-RTK points	–	–	231	–

Numeryczny model terenu generowany jest na podstawie tych punktów zbioru danych skaningu laserowego, które są odbiciami od powierzchni terenu. Wszystkie punkty odbite od obiektów nieterenowych, muszą być wcześniej usunięte z chmury punktów. Zadanie to realizowane jest w znacznym stopniu automatycznie poprzez oprogramowanie służące do filtracji danych skaningu laserowego. W pracy wykorzystano własne algorytmy filtracji opracowane w środowisku MATLAB. Algorytmy bazują na odpornej aproksymacji powierzchni terenu wielomianem drugiego stopnia w każdym punkcie pomiarowym [Borkowski i Józków 2006]. Punkty będące odbiciami wiązki lasera od obiektów nieterenowych traktuje się w tym podejściu jako błędy grube. Odporna estymacja parametrów wielomianu lokalnego polega na modyfikacji wag punktów pomiarowych w zależności od residuów pomiędzy wysokością interpolowaną z wielomianu a pomierzoną. Jako funkcję tłumienia przyjęto funkcję Krausa [Kraus 2000]. W algorytmie wykorzystano również ideę filtracji hierarchicznej [Briese et. al. 2002] polegającą na stopniowym eliminowaniu punktów obiektów, gdzie w każdym kolejnym kroku eliminowane są punkty leżące coraz bliżej powierzchni terenu, jednak wciąż nie będące punktami terenowymi. Zbiory punktów wydzielone dzięki automatycznemu procesowi filtracji zostały manualnie sprawdzone i skorygowane. W tym celu wykorzystano stację fotogrametryczną ImageStation oraz zdjęcia lotnicze w skali 1:26000. Ręczna weryfikacja polegała na wyeliminowaniu błędów filtracji:

- typu pierwszego (punkty terenu sklasyfikowane przez algorytm jako punkty obiektów),
- typu drugiego (punkty obiektów sklasyfikowane przez algorytm jako punkty terenu).

W wyniku przeprowadzonej weryfikacji automatycznego procesu filtracji danych lotniczego skaningu laserowego otrzymano dla każdego obszaru testowego zbiory punktów terenowych, które posłużyły do zbudowania NMT.

Numeryczny model terenu wygenerowany został na stacji fotogrametrycznej ImageStation w postaci siatki TIN opartej na punktach terenowych skaningu. NMT przedstawiony w postaci siatki TIN reprezentuje najprostszą metodę interpolacji liniowej. Mimo że metoda ta nie odtwarza wiernie powierzchni terenu, jednak z racji dużej gęstości punktów skaningu (ok. 3 punkty na 1 m<sup>2</sup>) utworzony model terenu można uznać za bardzo szczegółowy.

Ze względu na różną dokładność danych źródłowych skaningu dla zróżnicowanych kategorii terenu [Gołuch i in. 2007], dokładność wygenerowanego NMT określono oddzielnie w każdej kategorii terenu. Zbiór punktów pomierzonych bezpośrednio w terenie (łącznie z punktami osnowy pomiarowej) podzielony został na podzbiory reprezentujące każdą kategorię terenu. Liczba punktów terenowych w każdym z obszarów testowych w rozbiciu na poszczególne kategorie terenu została przedstawiona w tabeli 3.

Tabela 3. Liczba punktów pomierzonych bezpośrednio w terenie w rozbiciu na kategorie terenu  
Table 3. Number of points measured using GPS and total-station in each type of terrain coverage

Rodzaj pokrycia terenu Type of land cover	Liczba punktów Number of points				
	Razem Overall	Obszary testowe Testing areas			
		Krzywoustego	Sołtysowice	Paniowice	Szewczenki
Las i zadrzewienia Wood, forestry terrain	730	65	34	110	521
Zarośla i wysoka trawa Thicket and high grass terrain	753	213	540	–	–
Łąki i pola orne Meadows, arable fields	574	159	172	243	–
Drogi Roads	50	32	18	–	–
Ogółem Total	2107	469	764	353	521

W każdym punkcie pomierzonym techniką klasyczną obliczona została różnica wysokości pomiędzy wartością wyinterpolowaną z NMT,  $Z_{NMT}$  a wartością pomierzoną w terenie,  $Z_{Ter}$ . Jako dokładność NMT przyjęto błąd średni  $m_{NMT}$  obliczony dla wszystkich wartości różnic wysokości  $dz = Z_{NMT} - Z_{Ter}$ , w każdej z kategorii terenu oddzielnie:

$$m_{NMT} = \sqrt{\frac{\sum dz^2}{n}}, \quad (1)$$

gdzie  $n$  jest liczbą punktów kontrolnych.

Zestawienie uzyskanych dokładności przedstawiono w tabeli 4. W dolnym wierszu tabeli podano [Gołuch i in. 2007] dokładności danych skaningu laserowego dla poszczególnych kategorii pokrycia terenu. Dokładności te zostały oszacowane na

podstawie bezpośrednich pomiarów terenowych, techniką GPS i tachimetryczną, punktów skaningu laserowego. Z porównania dwóch ostatnich wierszy tabeli wynika, że dokładność NMT jest nieznacznie gorsza od dokładności danych skaningu laserowego. Dokładność tych ostatnich danych jest czynnikiem decydującym o dokładności NMT. Podane w tabeli 4 wartości mediany i średniej wartości  $dz$  wskazują na występowanie pewnego błędu systematycznego. Obliczony na podstawie danych skaningowych NMT „leży” powyżej rzeczywistej powierzchni terenu od kilkunastu do kilkudziesięciu centymetrów. Wielkość ta wynika głównie z szorstkości powierzchni terenu.

Tabela 4. Dokładności NMT w zależności od pokrycia terenu

Table 4. DTM accuracies depended on the landcover

	Rodzaj pokrycia terenu Type of landcover			
	Las i zadrzewienia Wood, forestry terrain	Zarośla i wysoka trawa Thicket and high grass terrain	Łąki i pola orne Meadows, arable fields	Drogi Roads
Liczebność n Quantity n	730	753	574	50
Minimalna wartość $dz$ Minimal value of $dz$	[m] -1.20	-0.99	-0.31	-0.21
Maksymalna wartość $dz$ Maximum value of $dz$	[m] 1.06	0.99	0.50	0.30
Mediana $dz$ Median $dz$	[m] 0.12	0.32	0.23	0.14
Średnia wartość $dz$ Mean value of $dz$	[m] 0.07	0.29	0.22	0.12
Średnia wartość bezwzględna $dz$ Absolute mean value of $dz$	[m] 0.37	0.37	0.23	0.15
Dokładność NMT $m_{NMT}$ Accuracy of DTM $m_{NMT}$	[m] 0.46	0.44	0.27	0.17
Dokładność danych skaningu Accuracy of LIDAR data	[m] 0.33	0.33	0.23	0.16

## WERYFIKACJA W MODELU HYDRODYNAMICZNYM

W odniesieniu do modelowania hydrodynamicznego ważnym parametrem jest dokładność względna (bądź wewnętrzna) NMT, charakteryzująca dokładność wzajemnego usytuowania elementów powierzchni terenu. Pewien błąd systematyczny nie powinien mieć znaczenia – ważniejsze jest zachowanie spadków terenu. Dla zweryfikowania przydatności zbudowanego NMT przeprowadzono modelowanie hydrodynamiczne dla fragmentu doliny rzeki Widawy.

Dwuwymiarowe modelowanie hydrodynamiczne przeprowadzono w środowisku SMS (Surface-water Modeling System) na podstawie siatki dyskretyzacyjnej, utworzonej w module Mesh. W projekcie dane z pozyskanego wysokodokładnego NMT były bezpośrednio importowane do modelu hydrodynamicznego. W module Mesh, na podstawie importowanych punktów, zbudowano sieć elementów metodą triangulacji. Następnie otrzymaną siatkę poddano weryfikacji, korzystając z funkcji oprogramowania

(np. w celu łączenia trójkątów, zamiany krawędzi podziału elementu czworokątnego na dwa elementy trójkątne czy wręcz usunięcia elementów). Na rysunku 2 przedstawiono, utworzoną w module Mesh programu SMS, siatkę elementów skończonych. Siatka ta zbudowana została na podstawie 34 573 punktów NMT, składa się z 47 332 punktów obliczeniowych (nodes) i posiada 23 317 elementów trójkątnych lub czworokątnych. Kiedy siatka dyskretyzacyjna budowana jest na podstawie NMT w postaci TIN, wówczas co prawda wymagany jest większy wkład pracy przy jej weryfikacji, ale w NMT o takiej strukturze mamy już uwzględnione elementy morfologiczne terenu, bardzo ważne przy modelowaniu przepływu wody (np. wały przeciwpowodziowe). Istotny jest również fakt, że liczba punktów w modelu terenu o strukturze TIN jest znacznie mniejsza niż w NMT o strukturze GRID (szczególnie w terenach płaskich) – jest to ważne zwłaszcza w dwuwymiarowym modelowaniu hydrodynamicznym, gdzie czas uzyskania rozwiązania jest proporcjonalny do kwadratu liczby punktów siatki.

W projekcie zamodelowano przepływ  $Q=130 \text{ m}^3/\text{s}$ , na który była zaprojektowana przez Niemców przepustowość doliny rzeki Widawy. Rysunek 3 ilustruje rozkład strug wody uzyskany w wyniku przeprowadzonego modelowania hydrodynamicznego. Otrzymany rozkład strug wody odpowiada, w ocenie hydrologa, wynikom otrzymanym z modelowań alternatywnych oraz z pomiarów wykonanych podczas powodzi w 1997 r. [Parzonka 2003]. Również inne parametry związane z modelowaniem hydrodynamicznym potwierdziły przydatność zbudowanego NMT. Szersze przedstawienie problematyki związanej z modelowaniem hydrodynamicznym wykracza poza ramy tego artykułu.



Rys. 2. Siatka dyskretyzacyjna utworzona na podstawie punktów NMT  
Fig. 2. Discrete mesh based upon DTM points



Rys. 3. Układ strug wody przy modelowanym przepływie  $Q=130 \text{ m}^3/\text{s}$

Fig. 3. Arrangement of flow trace in modelled flow of  $Q=130 \text{ m}^3/\text{s}$

## PODSUMOWANIE

W pracy przedstawiono wyniki oceny dokładności numerycznego modelu terenu interpolowanego na podstawie danych lotniczego skaningu laserowego systemu ScaLARS. Wysokości uzyskane z NMT zostały porównane z przyjętymi jako bezbłędne wysokościami punktów pomierzonych bezpośrednio w terenie. Na tej podstawie obliczono dokładność NMT, która jest różna dla zróżnicowanych kategorii terenu. Dla terenów płaskich, na których pokrycie (szczególnie roślinnością) nie występuje (np. drogi) dokładność NMT wynosi 17 cm i jest zbliżona do dokładności danych źródłowych skaningu, która wynosi 16 cm [Gołuch i in. 2007]. Natomiast jednocześnie ze wzrostem wysokości oraz gęstości roślinności pokrywającej teren dokładność NMT zmniejsza się w stosunku do dokładności danych źródłowych. Spowodowane jest to głównie mniejszą gęstością punktów terenowych skaningu, przez co model terenu w większym uproszczeniu opisuje lokalne struktury terenu. Fakt ten ma szczególne znaczenie dla terenów o mocno zróżnicowanej powierzchni (zagłębienia terenu, skarpy). Najmniejszą dokładność NMT uzyskano dla terenów leśnych oraz zadrzewionych; wyniosła ona 46 cm. Dla tej kategorii terenu również dokładność danych źródłowych jest najgorsza i wynosi 33 cm. W tym przypadku proces filtracji danych skaningu spowodował zmniejszenie gęstości punktów w stosunku do terenów nie porośniętych roślinnością, przez co dla terenów leśnych numeryczny model terenu oparty na siatce TIN ma mniejszą rozdzielczość, a w efekcie również dokładność.

W kontekście uzyskanych wartości różnic pomiędzy dokładnościami NMT i danych źródłowych skaningu należy uznać, że proces filtracji przeprowadzony został prawidłowo i wraz z wybraną metodą interpolacji NMT nie ma znaczącego wpływu na dokładność numerycznego modelu terenu interpolowanego na podstawie danych

lotniczego skaningu laserowego. Poza dokładnością danych źródłowych skaningu największy wpływ na dokładność NMT ma rodzaj pokrycia terenu.

## PIŚMIENNICTWO

- Borkowski A., Gołuch P., Wehr O., Schiele M., Thomas M., 2006. Airborne laser scanning for the purpose of hydrodynamic modelling of Widawa river valley. Reports on Geodesy, No 2(77), pp. 85–94.
- Borkowski A., Józków G., 2006. Wykorzystanie wielomianowych powierzchni ruchomych w procesie filtracji danych pochodzących z lotniczego skaningu laserowego. Archiwum Fotogrametrii, Kartografii i Teledetekcji, Vol. XVI, 63–73.
- Briese C., Pfeifer, N., Dorninger P., 2002. Applications of the robust interpolation for DTM de-termination. Symposium ISPRS Commission III, Photogrammetric Computer Vision, Graz, 9–13 September 2002. International Archives of Photogrammetry and Remote Sensing, Vol. XXXIV / 3A, 55–61.
- Gołuch P., Borkowski A., Józków G., 2007. Ocena dokładności danych lotniczego skaningu laserowego systemu ScaLARS. Archiwum Fotogrametrii, Kartografii i Teledetekcji, Vol. XVII, s. 261–270.
- Kraus K., 2000. Photogrammetrie. Band 3. Topographische Informatonssysteme. Dümmler, Köln.
- Parzonka W., 2003. Ocena przepustowości doliny Widawy dla przeprowadzenia przepływów powodziowych Odry. Praca zbiorowa pod red. W. Parzonki, Zesz. Nauk. AR Wrocław nr 454.
- Schiele O., Wehr A., Kleusberg A., 2005. Operational Calibration of Airborne Laserscanners by Using LASCAL. Proc. Optical 3-D Measurement Techniques, vol. 1, Vienna Oct. 3–5, pp. 81–89.

## STUDY OF ACCURACY OF DTM INERPOLATED FROM AIRBORNE LASER SCANNING DATA OF SCALARS SYSTEM

**Abstract.** Accuracy of Digital Terrain Model (DTM) generated from airborne laser scanning data depends on many factors, e.g. terrain structures, landcover, stability of photogrammetric flight, quality of navigation data, accuracy of calibration, size of laser footprint on terrain (height of flight and convergence of laser beam), density of captured points, method of raw ALS data filtering. In this work the accuracy determination of DTM generated for 20-kilometer part of valley of Widawa river was presented. This DTM was used in hydrodynamic modelling. Airborne laser scanning was carried out using prototypic ScaLARS scanner (developed in Institute of Navigation of Stuttgart University). INS and GPS signals were registered by Applanix POS/AV 510 system. Photogrammetric flight using AN-2 aeroplane was made from height of 550 m. Footprint of laser beam had on the terrain size of about 0.6 m. Calibration of system was carried out semi-automatically. In the reference of GPS measured control fields relative error was estimated on the level about 0.3 m (along and across the flight direction) and error of height was about 0.1 m. Research of accuracy determination of generated DTM was carried out based upon fields measurements using GPS and tacheometric techniques. The measure-

ments were made for four representative fields of study area. Data filtering was carried out using own algorithms based upon robust estimation of moving polynomial surface to scanning data. Depending on the terrain landscape and landcover DTM accuracy was evaluated from value 0.17 m to 0.46 m.

**Key words:** airborne laser scanning, DTM, ScaLARS

Zaakceptowano do druku – Accepted for print: 30.06.2008

Do cytowania – For citation: Gołuch P., Borkowski A., Józków G., 2008. Badanie dokładności NMT interpolowanego na podstawie danych lotniczego skaningu laserowego systemu ScaLARS. *Acta Sci. Pol. Geodesia et Descriptio Terrarum*, 7(2), 37–47.





## DANE W PLANOWANIU PRZESTRZENNYM

Anna Kowalczyk, Kamil Kowalczyk

Uniwersytet Warmińsko-Mazurski w Olsztynie

**Streszczenie.** W ciągu ostatnich lat ze względu na postęp techniczny (Internet, dostępność danych GPS, informatyka) geodezja rozszerzyła swoje działanie. Wcześniejsze zarządzanie terenów rolnych przekształciło się także w planowanie terenów miejskich, wiejskich oraz stref przejściowych. W referacie zaprezentowano zakres danych, jaki jest niezbędny do właściwego planowania przestrzeni, zaprezentowano istotę tematu oraz ściśle powiązanie planowania z systemami informacji o terenie (SIT) i systemami informacji geograficznej (GIS). Dokonano podziału danych, pozyskiwanych do wykorzystywania w procesach planistycznych, ze względu na ich specyfikację. Przedstawiono zakresy oraz źródła danych do celów planistycznych.

Pokazano także jak wiele informacji, w tym również geodezyjnych, ma wpływ na racjonalne wykorzystanie przestrzeni.

**Słowa kluczowe:** planowanie przestrzenne, SIT, GIS, dane przestrzenne

### WSTĘP

Właściwe zaplanowanie przestrzeni wymaga od planisty dużej wiedzy o terenie, na którym prowadzone są prace projektowe. Im więcej informacji, tym zaplanowana przestrzeń będzie bardziej funkcjonalna. Dane do planowania przestrzennego pozyskiwane są z różnych źródeł: od historycznych po ankiety wśród mieszkańców. Dotyczą one takich informacji, jak: istniejąca infrastruktura techniczna, drogi, domy, położenie historyczne i turystyczne, bliskość jezior, rzek, odległość od miasta, wysokość zabudowy istniejącej, uwarunkowania ochrony środowiska, topografia terenu, potrzeby ludzi itd. Bez odpowiedniej wiedzy na temat przestrzeni nie można kompetentnie wykonać zadanych opracowań.

Aby prace planistyczne przebiegały sprawnie, najlepiej te dane sklasyfikować i uporządkować, innymi słowy – stworzyć system informacji. Jak w każdym systemie informacji pozyskanie danych ma pierwszorzędne znaczenie a liczba danych jest niezmiernie duża.

## ŹRÓDŁA DANYCH

Pomimo wielu definicji DANYCH nie istnieje uniwersalna formuła. W informatyce dane to zbiory liczb i tekstów o różnych formach. Są one wykorzystywane do obliczeń, prezentowane i przetwarzane cyfrowo. Tematyczne zbiory danych (informacji) to BAZY DANYCH. W fizyce i matematyce dane to wartości znane w rozwiązywaniu problemów, np. zadań. W języku potocznym dane to otrzymane informacje lub wiadomości używane do wyciągania jakichś wniosków [2].

Meta-teoria TOGA mówi, że "to wszystko co jest/może być przetwarzane umysłowo lub komputerowo". Dane mogą przyjmować różną postać, jak np.: znaki, słowa, wykresy, sygnały.

Do celów planistycznych dane można pozyskać z następujących źródeł:

- mapy zasadnicze (istniejąca infrastruktura techniczna, drogi, budynki, ukształtowanie terenu),
- mapy topograficzne (odległość od miasta, położenie jezior, rzek i lasów),
- mapy glebowe (dane o stabilności gruntu),
- mapy tematyczne: środowiskowe, historyczne, turystyczne, itp. (ewentualne tereny zagrożone zalaniem, położenie rezerwatów i terenów chronionych, miejsca kultu, punkty widokowe),
- bezpośrednie pomiary terenowe (kontrola danych),
- zdjęcia podczas pomiaru terenowego (charakter istniejącej zabudowy),
- zdjęcia lotnicze i satelitarne (podobnie jak mapy topograficzne i tematyczne),
- ankiety wśród mieszkańców (potrzeby lokalnej społeczności),
- strategie rozwoju miasta i województwa (potrzeby miasta i mieszkańców),
- Systemy Informacji Przestrzennej (dane o terenie),
- Ewidencja Gruntów i Budynków (dane o stanie prawnym działek),
- Ewidencja Sieci Uzbrojenia Terenu (dane o istniejącej i projektowanej infrastrukturze technicznej).

## WYMAGANIA WZGLĘDEM DANYCH Z SYSTEMÓW INFORMACYJNYCH

Dużą rolę w pozyskaniu danych do celów planowania przestrzennego odgrywają wszelkiego rodzaju systemy informacyjne. Dotyczą one przestrzeni, ewidencji gruntów i budynków, ewidencji sieci uzbrojenia terenu oraz wszelkich innych klasyfikacji tematycznych. Aby taki system mógł być wykorzystany przez planistów (i nie tylko), musi spełniać pewne warunki. I tak najważniejszymi cechami danych w SIP są:

- dokładność (bliskość wartości prawdziwej),
- precyzja (zdolność dokładnego przedstawienia wielkości),
- powtarzalność (wzajemna zgodność powtarzanych danych),
- rozdzielczość (zdolność wykrywania lub rozróżniania wielkości),
- zmienność (atrybut: czas po jakim następuje zmiana),
- aktualność (cecha ta związana jest z odstępami w czasie),
- wiarygodność (zgodność pomiędzy informacjami),
- dostępność (łatwość i szybkość uzyskiwania danych),
- kompletność (to co się ma, a to co powinno być),
- odpowiedniość (stopień zaspokojenia potrzeb),

- koszt (koszt uzyskania, przetworzenia i dostarczenia),
- wartość (korzyść bądź strata z wykorzystania danych).

## POSTAĆ DANYCH

Podczas prac planistycznych materiał, z którego pozyskiwane są dane, może mieć różną formę. Dzieli się go na dwie kategorie: analogowa i cyfrowa.

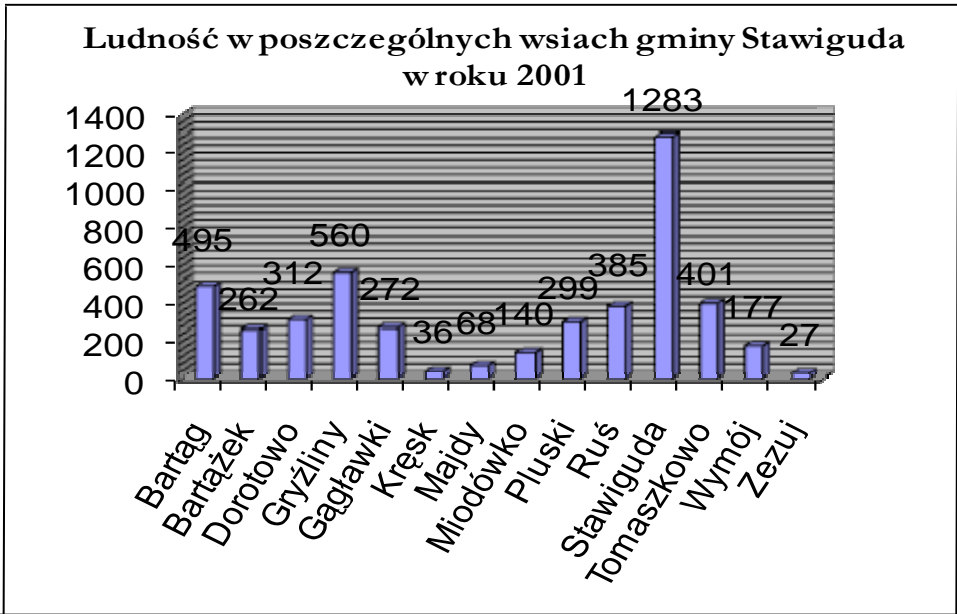
Forma analogowa jest to graficzna postać rysunkowa lub obrazowa, są to przede wszystkim mapy wykonywane w technice papierowej (rys. 1), opracowania branżowe, zapisane informacje ustne itd.



Rys. 1. Przykład analogowej formy źródła danych

Fig. 1. Example of analog form of data source

Forma cyfrowa jest to zapis informacji na nośnikach cyfrowych (dyskietka, dysk, płyta CD, płyta DVD, serwer danych). Postać danych cyfrowych może być numeryczna, literowa bądź alfanumeryczna [1]. Dane cyfrowe mogą być w formie tabel (np. Excel, OpenOffice), wykresów (rys. 2), map cyfrowych i numerycznych [Kowalczyk 2007], a także baz danych mniej lub bardziej rozbudowanych.



Rys. 2. Przykład przedstawienia danych w formie wykresu stworzonego w programie Excel  
 Fig. 2. The example of the numerical data, created in the Excel program

## DANE W PLANOWANIU PRZESTRZENNYM

Ze względu na przeznaczenie danych odnoszących się do przestrzeni, można je klasyfikować w następujący sposób:

- dane do celów planowania przestrzennego (np. dla planistów, np. EGİB)
- dane planistyczne dla zainteresowanych (np. dla osób fizycznych, np. MPZP)

Stworzenie opracowań planistycznych wymaga posiadania szerokiego spektrum informacji o przestrzeni. Do tego celu niezbędne są kompleksowe informacje o tym, co pod ziemią, na ziemi jak i nad ziemią. Innymi słowy niezbędna jest wiedza wykorzystująca dane GIS (Geographical Information System).



Rys. 3. System Informacji Geograficznej  
Fig. 3. Geographical Information System

Dane przestrzenne oraz informacje w nich zawarte mogą pochodzić z różnych źródeł (tab. 1).

Tabela 1. Źródła danych przestrzennych  
Table 1. Sources of spatial data

Źródło danych	Rodzaj źródła	Postać danych	Zapis danych
Teren	Rzeczywistość	Cyfrowa Analogowa	Wektorowy Rysunkowy
Zdjęcia i dane fotogrametryczne i teledetekcyjne	Zarejestrowany model	Cyfrowa Analogowa	Obrazowy Rastrowy
Mapa graficzna	Przetworzony model	Analogowa	Rysunkowy Obrazowy
Mapa numeryczna	Przetworzony model	Cyfrowa	Wektorowy Rastrowy

Źródło: Gaździcki 1990, „System Informacji Przestrzennej” Państwowe Przedsiębiorstwo Wydawnictw Kartograficznych, Warszawa–Wrocław 1990

Źródła danych niezbędnych w planowaniu przestrzennym to przede wszystkim EGiB (ewidencja gruntów i budynków), KW (księgi wieczyste), PPM (projekty planów miejscowych), analizy własne, mapy oraz dokumentacja prawna.

Niezbędne w pozyskiwaniu danych jest zdefiniowanie tego, co się chce pozyskać. Ze względu na ilość istotnych na potrzeby planistyczne danych można je podzielić ze względu na specyfikę i wyróżnić następujące dane: [4]:

- I. formalno-prawne,
- II. administracyjne dotyczące obszaru objętego opracowywanym planem,
- III. strukturalno-przestrzenne dotyczące obszaru objętego opracowywanym planem,
- IV. o wyposażeniu w infrastrukturę techniczną,
- V. rynkowe.

Dane te można również sklasyfikować, biorąc pod uwagę ich zakres. Mogą one bowiem dotyczyć pewnego obszaru: województwa, powiatu, jednostki ewidencyjnej, obrębu ewidencyjnego, konkretnej działki itd. Poszukując niezbędnych danych, należy wiedzieć, z jakich miejsc możemy pozyskać konkretne dane. Przykładowe zestawienie danych oraz miejsc ich pozyskania przedstawia tab. 2.

Tabela 2. Dane, ich źródła i miejsca pozyskania dla potrzeb gospodarki przestrzennej w gminie (Kryszk, 2006)

Table 2. Data, their sources and places of data acquisition for needs of spatial economy in commune (Kryszk, 2006)

Dane	Źródła i miejsca ich pozyskania
Ludność	Urząd Statystyczny, Odpowiednie referaty Urzędu Gminy, Spisy Ludności
Administracyjno-prawne	EGiB, Dokumentacja klasyfikacji gruntów, mapy glebowo-rolnicze, Miejskowy Plan Zagospodarowania Przestrzennego, KW
Funkcje gospodarcze	Miejskowe Plany Zagospodarowania Przestrzennego, odpowiednie referaty Urzędu Gminy
Infrastruktura społeczna i techniczna	Dokumentacja Urzędu Gminy, dokumentacja jednostek branżowych, Urząd Statystyczny
Środowisko	Wydział Gospodarki Przestrzennej i Ochrony Środowiska Urzędu Wojewódzkiego, Wydział Infrastruktury Technicznej i Komunikacji Urzędu Wojewódzkiego, Wojewódzki Inspektorat Ochrony Środowiska, Stacja Sanitarno-Epidemiologiczna

## PODSUMOWANIE

Podsumowując powyższe rozważania dotyczące danych w planowaniu przestrzennym, można dokonać podziału „istnienia” danych na trzy podstawowe etapy.

Pierwszym etapem jest pozyskiwanie danych do celów planowania przestrzennego. Tutaj należy dokonać wnikliwej analizy, jakie dane są potrzebne. Następnie trzeba znaleźć źródło tych danych, czyli wiedzieć, skąd je wziąć. Ostatnim krokiem na tym etapie jest wybór formy pozyskiwania danych.

Drugi etap stanowi gromadzenie danych przestrzennych. Jest wiele sposobów oraz form gromadzenia danych. Mogą to być elektroniczne bazy danych, Ewidencje Gruntów i Budynków, Księgi Wieczyste czy też Miejscowe Plany Zagospodarowania Przestrzennego, itd.

Ostatnim etapem jest udostępnianie danych planistycznym zainteresowanym. Poszukując niezbędnych danych, zainteresowany powinien wiedzieć, z jakich miejsc możemy pozyskać konkretne dane. Ważna też jest forma, w jakiej wydawane są dane, np. wypis czy wyrys lub wyciąg.

## PIŚMIENNICTWO

[http://home.agh.edu.pl/~hycner/files/ppp\\_wykl5\\_2007.pdf](http://home.agh.edu.pl/~hycner/files/ppp_wykl5_2007.pdf)

<http://pl.wikipedia.org/wiki/Dane>

Gaździcki J., 1990. System Informacji Przestrzennej, PPWK, Warszawa–Wrocław 1990.

Kowalczyk K., 2007, Wybrane zagadnienia z rysunku map, UWM, Olsztyn

Kryszk H., 2006. Opracowanie metody wykorzystania informacji przestrzennych do prognozowania skutków finansowych miejscowych planów zagospodarowania przestrzennego, rozprawa doktorska, Uniwersytet Warmińsko-Mazurski w Olsztynie 2006.

## DATA IN SPATIAL PLANNING

**Abstract.** Geodesy widened one's working in last years in view of technical progress (Internet, GPS, informatics). Earlier rural development reformed also into planning of municipal areas, country as well as transitory zones. Range of data what is indispensable to proper planning of space is presented in this paper. Essence of subject as well as connection of planning with LIS and GIS is presented. Division of data was executed for planning processes in view of their specification. Ranges and sources of data for spatial planning are presented. It was shew also as many information, in this also geodetic, has influence onto rational use of space.

**Key words:** spatial planning, LIS, GIS, spatial data

Zaakceptowano do druku – Accepted for print: 30.06.2008

Do cytowania – For citation: Kowalczyk A., Kowalczyk K., 2008. Dane w planowaniu przestrzennym. *Acta Sci. Pol. Geodesia et Descriptio Terrarum*, 7(2), 49–55.





## SPIS TREŚCI CONTENTS

### **Anna Barańska**

- Metody jakościowe i ilościowe na usługach wyceny nieruchomości..... 3  
Quality and quantity methods for estimating real estates

### **Andrzej Borkowski, Grzegorz Jóźków**

- Filtracja danych lotniczego skaningu laserowego  
metodą ruchomych powierzchni wielomianowych – weryfikacja metody..... 15  
Airborne laser scanning data filtering using moving  
polynomial surface model – the verification of the method

### **Sławomir Cellmer, Jacek Rapiński**

- Numeryczny sposób wyznaczania całkowitych wartości  
nieoznaczoności pomiaru fazowego..... 29  
Numeric approach of estimation of integer value  
of phase's measurements

### **Piotr Gołuch, Andrzej Borkowski, Grzegorz Jóźków**

- Badanie dokładności NMT interpolowanego  
na podstawie danych lotniczego skaningu laserowego systemu ScaLARS..... 37  
Study of accuracy of DTM interpolated from airborne laser  
scanning data of ScaLARS system

### **Anna Kowalczyk, Kamil Kowalczyk**

- Dane w planowaniu przestrzennym..... 49  
Data in spatial planning