

53/2005

Raport Badawczy

RB/36/2005

Research Report

**Program komputerowy
aproksymacji krigingowej**

L. Bogdan, W. Kozłowski

**Instytut Badań Systemowych
Polska Akademia Nauk**

**Systems Research Institute
Polish Academy of Sciences**



POLSKA AKADEMIA NAUK

Instytut Badań Systemowych

ul. Newelska 6

01-447 Warszawa

tel.: (+48) (22) 8373578

fax: (+48) (22) 8372772

Kierownik Pracowni zgłaszający pracę:
Prof. dr hab. inż. Zbigniew Nahorski

Warszawa 2005

Raport Badawczy

RB/36/2005

Research Report

**Program komputerowy
aproxymacji krigingowej**

Lucyna Bogdan, Wojciech Kozłowski

**Instytut Badań Systemowych PAN
Polska Akademia Nauk**

**Systems Research Institute
Polish Academy of Sciences**

Warszawa 2005

Lucyna Bogdan, Wojciech Kozłowski

Program komputerowy aproksymacji krigingowej

Spis treści

Wprowadzenie

1. Co to jest Kriging ?
2. Poprawiony algorytm obliczeniowy
3. Opis programu MrKrig v.2.0
4. Wyniki pomiarowe
5. Wnioski końcowe

Literatura

Wprowadzenie

Celem prowadzonych prac jest opracowanie programów komputerowych umożliwiających aproksymację czasowo – przestrzenną zmiennych meteorologicznych, na przykład temperatury powietrza, wilgotności powietrza, składu chemicznego gleby itp. Takie programy mogą być użytecznym narzędziem przy tworzeniu na przykład map zanieczyszczeń w glebie, wodach powierzchniowych i w atmosferze, co jest istotnym elementem badań w inżynierii i ochronie środowiska. Algorytmy krigingowe są jedną z wielu możliwości aproksymacji zmiennych w czasie i przestrzeni, jednak jest to narzędzie w Polsce stosunkowo mało znane i dlatego podjęto badania dla przetestowania ich użyteczności. W dalszym ciągu planuje się także porównanie skuteczności algorytmów krigingowych i innych metod aproksymacyjnych. Ponieważ dysponujemy w Instytucie Badań Systemowych dostępem do rzeczywistych danych pomiarowych z monitoringu geostatycznego, więc wszelkie obliczenia testowe będą dokonywane na tych danych.

W raporcie w kolejnych punktach przedstawiono opis poprawionego algorytmu komputerowego realizującego metodę krigingową aproksymacji, opis programu obliczeniowego zawierającego ten algorytm oraz wstępne wyniki obliczeń dotyczące aproksymacji opadów deszczowych w zadanych punktach obszaru Polski na podstawie znanych pomiarów opadu deszczu w innych punktach pomiarowych

1. Co to jest Kriging ?

Kriging –metoda geostatyczna – pozwala na szacowanie jakości i wielkości złóż oraz ilustrowanie rozmieszczenia wartości parametrów złożowych za pomocą map izarytm (np. warstwic). Pierwotnie była wykorzystywana do oszacowania złóż złota. Mając kilka próbek (tzw. punktów bazowych) można wygenerować całą mapę złoża (Cressie, 1991; Isaaks, Srivastava, 1989).

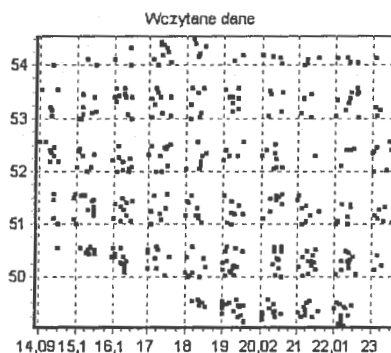
2. Poprawiony algorytm obliczeniowy

Poniżej przedstawiamy kilkanastokrokowy poprawiony algorytm obliczeniowy metody krigingowej. Porzedni algorytm używany w programie MrKrig 1 został uzupełniony o kilka kroków (Kozłowski, Studziński, 2004).

Algorytm ten został opracowany w postaci programu MrKrig v.2.0 i za pomocą tego programu wykonano wstępne obliczenia aproksymacyjne dla danych pomiarowych dotyczących rocznych opadów deszczu w roku 1980 w wybranych punktach pomiarowych w Polsce.

Kolejne kroki algorytmu obliczeniowego są następujące:

Krok 1. Wyznaczenie zbioru tzw. punktów bazowych, to znaczy punktów pomiarowych badanej zmiennej, dla których są znane ich współrzędne geograficzne (zadane na przykład w postaci szerokości i długości geograficznej wartości N i E, por. rys.1) oraz jest znana wartość danej zmiennej z.



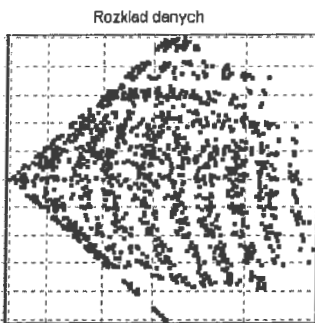
Rys.1. Wybrane punkty bazowe (pomiarowe) pomiarów opadów deszczu w Polsce.

Krok 2. Sprawdzenie, czy istnieją punkty bazowe o identycznym położeniu, a innej wartości z. Oznacza to błąd w pomiarach i należy wówczas zdecydować o usunięciu jednego z punktów.

Krok 3. Na podstawie informacji o punktach pomiarowych utworzenie tablicy par punktów pomiarowych stworzonych na zasadzie: Każdy punkt jest łączony z każdym. Dla N punktów bazowych nowy plik zawiera liczbę elementów L równą liczbie wariacji bez powtórzeń i bez uwzględniania kolejności, to znaczy $L = N! / ((N-2)!2)$. W tablicy zapisane są informacje o

odległości między punktami $h = \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2}$ oraz wartość kąta $\alpha = \arctan\left[\frac{y_i - y_{i+1}}{x_i - x_{i+1}}\right]$, gdzie x i y – współrzędne punktów pomiarowych.

Krok 4. Wyznaczenie kąta α oraz jego tolerancji. Do dalszych obliczeń krigingowych będą używane pary punktów znajdujące się w obszarze tolerancji kąta α



Rys. 2. Pary punktów w znajdujące się w obszarze tolerancji kąta α

Krok 5. Wyznaczenie zakresu maksymalnej odległości między punktami w parach h_{\max}

Krok 6. Wyznaczenie liczby lagów, czyli liczby przedziałów zakresu odległości między punktami w parach

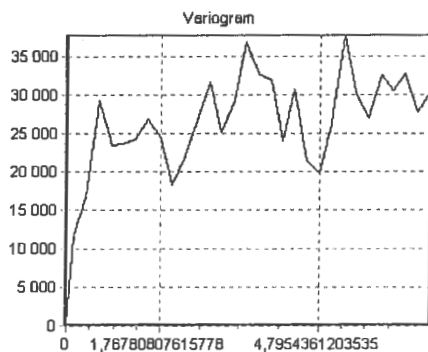
Krok 7. Dla zbioru lagów wyliczana jest tzw. semiwariogram empiryczny ze wzoru:

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^{n_h} (z_{h+i} - z_i)^2$$

gdzie n liczebność pojedynczego laga

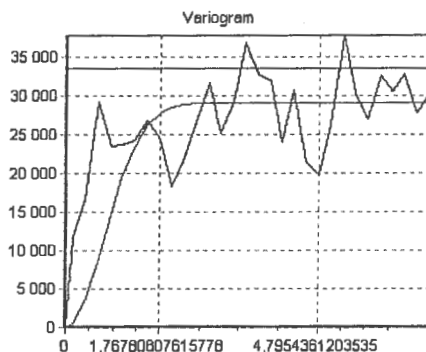
h oznacza odległość między poszczególnymi punktami bazowymi.

Krok 8. Wykreślenie semiwariogramu empirycznego. Na osi X znajdują się odległości między parami punktów bazowych, na osi Y wartości $\gamma(h)$ dla poszczególnych lagów (por. rys.3).



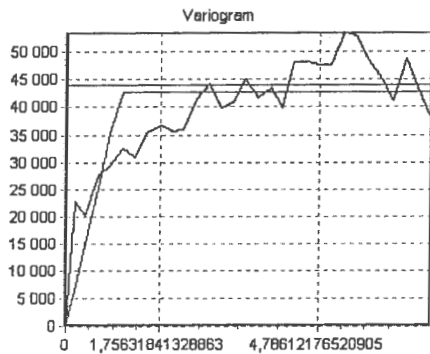
Rys.3. Przykład semiwariogramu empirycznego

Krok 9. Dobranie postaci modelu do semiwariogramu empirycznego ze zbioru zadanych modeli standardowych (por. rys. 4.). Dobór modelu wykonywany jest automatycznie za pomocą gradientowej metody Marquardta. Podstawowe modele standardowe używane w obliczeniach, to model sferyczny, gaussowski i eksponentialny, chociaż można oczywiście stosować również inne postacie modeli aproksymujących obliczone semiwariogramy empiryczne.



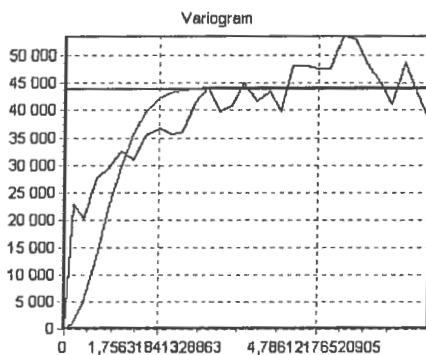
Rys. 4. Aproksymacja semiwariogramu empirycznego za pomocą modelu (funkcji) o wybranej strukturze.

a)



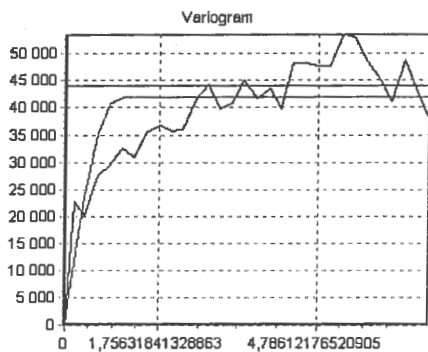
```
//----- sferyczny -----//  
//y = b0*(3/2*h-1/2*h^3)  
// y = b0 <=> h > 1
```

b)



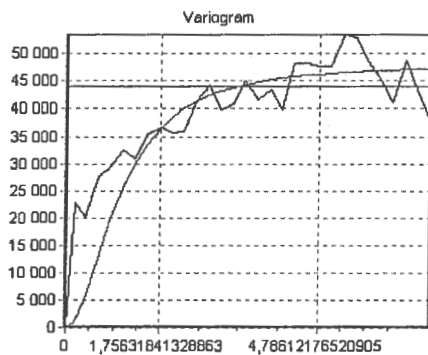
```
//----- gaussowski -----//  
//y = b0*(1-exp(-h*h));
```

c)



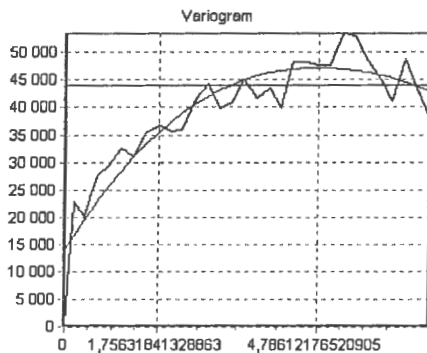
//----- kwadratowy -----//
//y = b0*(2*h - h^2) //

d)



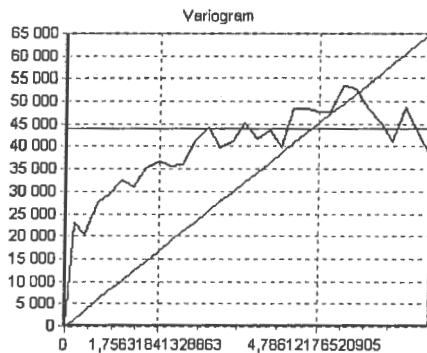
//----- r kwadratowy -----//
//y = b0*(h^2/(1+h^2)) //

c)



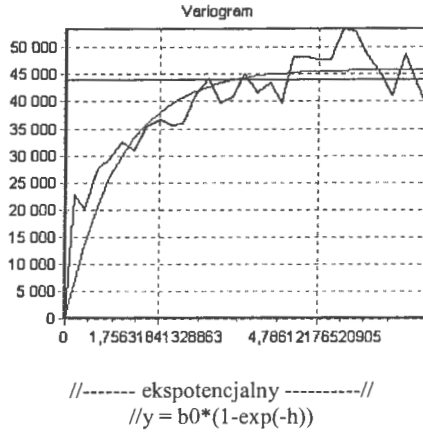
//----- wielomianowy -----//
//y= b0+ (b1 + (b2 + (b3*h))*h)*h

f)



//----- liniowy -----//
// y = b0*h

g)



Rys. 5. Zadane struktury modeli do aproksymacji semiwariogramów empirycznych: a) sferyczny, b) gaussowski, c) kwadratowy, d) r kwadratowy, e) wielomianowy, f) liniowy, g) eksponentialny

Krok 10. Tworzenie tabeli wartości Z dla poszczególnych punktów bazowych

	x_1	x_2	x_3
z_i	z_1	z_2	z_3

Krok 11. Tworzenie symetrycznej macierz Q odległości między punktami bazowymi.

	x_1	x_2	x_3
x_1	$h_{11} = 0$	h_{12}	h_{13}
x_2	h_{21}	$h_{22} = 0$	h_{23}
x_3	h_{31}	h_{32}	$h_{33} = 0$

Krok 12. Macierz symetryczną Q zostanie transformowana na Q_i , czyli zostanie wyliczana wartość $\gamma(h_{ij})$, z wykorzystaniem najlepiej dobranej modelu oraz zostanie dopisany wiersz i kolumna *jedynek*.

	x_1	x_2	x_3	
x_1	$\gamma(h_{11})$	$\gamma(h_{12})$	$\gamma(h_{13})$	1
x_2	$\gamma(h_{31})$	$\gamma(h_{32})$	$\gamma(h_{33})$	1
x_3	$\gamma(h_{31})$	$\gamma(h_{32})$	$\gamma(h_{33})$	1
	1	1	1	0

Krok 13. Transformowana macierz Q_i zostanie odwrócona na $Q_i^{(-1)}$

Krok 14. Tworzenie macierzy S odległości punktu x_0 od punktów bazowych.

	x_0
x_1	h_{10}
x_2	h_{20}
x_3	h_{30}

Krok 15. Transformacja macierzy S na S_r , czyli wyliczenie wartości $\gamma(h_{i0})$ oraz dopisanie kolumnę *jedynek*.

	x_0
x_1	$\gamma(h_{10})$
x_2	$\gamma(h_{20})$
x_3	$\gamma(h_{30})$
	1

Krok 16. Obliczenie wagi krigingu λ_i oraz współczynnik Lagrange'a z równania macierzowego:

$$\begin{bmatrix} \gamma(x_i, x_j) & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_i \\ \nu \end{bmatrix} = \begin{bmatrix} \gamma(x_i, x_0) \\ 1 \end{bmatrix}$$

lub

$$[Q_i^{(-1)}] \cdot [S_i] = \begin{bmatrix} \lambda_i \\ \nu \end{bmatrix}$$

Krok 17. Wyliczenie szukanej wartości z_0 punktu x_0 , wykorzystując dane z tabeli Z oraz z Kroku 11.

$$z_0 = \sum_{i=1}^n \lambda_i \cdot z_i$$

3. Opis programu MrKrig 2.0

Program został napisany w języku C++ na platformie Borland C++ Enterprise 6.0.

Program znajduje się w pliku o nazwie *kriging.exe*, wczytuje dane pomiarowe znajdujące się w pliku tekstowym *dane.txt*. Te dwa pliki powinny być w jednym katalogu.



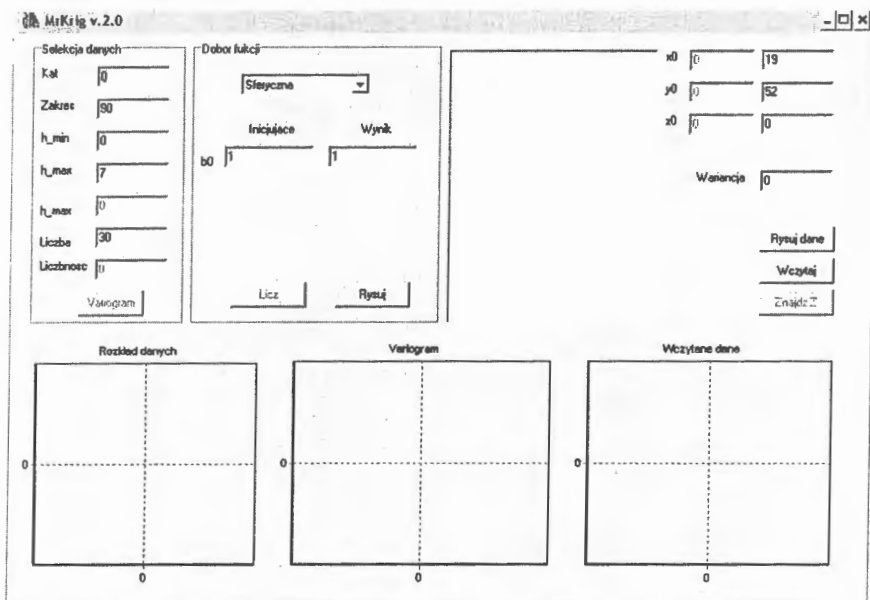
Rys. 6. Pliki programu MrKrig v.2.0

Dane pomiarowe muszą być uporządkowane w trzech kolumnach. Pierwsza kolumna zawiera współrzędną x , druga współrzędną y , trzecia wartość z . Jeżeli dane są wartościami ułamkowymi, to część ułamkowa musi być oddzielona od części całkowitej „kropką”, nie „przecinkiem”.

Plik	Edycja	Format	Widok	Pomoc
15.13	53	627		
22.02	54.15	752		
23.53	51.03	797		
23.51	52.42	855		
23.45	52.57	852		
23.37	52.04	585		
23.33	51.33	623		
23.3	51.07	700		
23.26	50.23	887		
23.17	51.29	571		
23.17	53.18	822		
23.16	51.11	711		

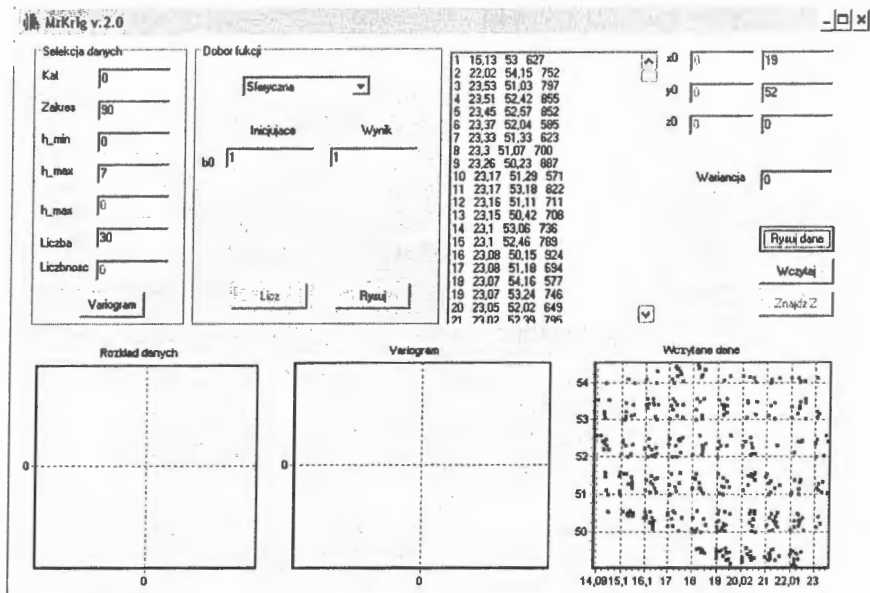
Rys. 7. Plik *dane.txt*.

Program MrKrig v.2.0 uruchamiany jest, gdy zostanie otwarty plik *kriging.exe* (por. rys. 8).



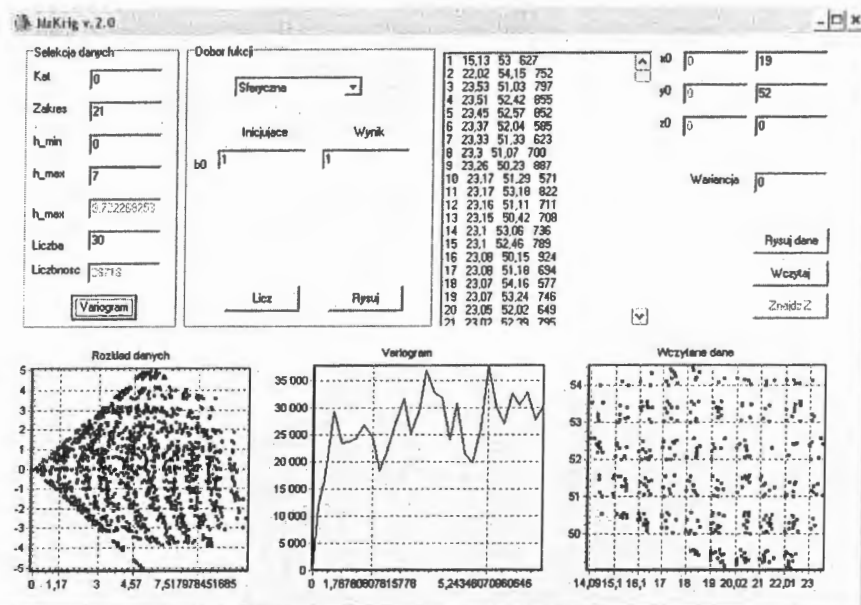
Rys. 8. Program MrKrig v.2.0

Następnie należy wczytać dane, naciskając przycisk **Wczytaj**. Dane można wyświetlić na prawym wykresie *Wczytane dane*, gdy zostanie naciśnięty przycisk **Rysuj dane**. Dodatkowo na wykresie (na czerwono) jest zaznaczony punkt szukany, którego współrzędne x , y znajdują się na prawym wykresie w prawej kolumnie. Na środku znajduje się okno dialogowe zawierające wszystkie punkty bazowe (por. rys. 9). Wybór współrzędnych punku szukanego można zrealizować na dwa sposoby. Pierwszy - należy wpisać wartości współrzędnych x i y w oknie dialogowym x_0 i y_0 . Drugi – należy wybierać z listy punktów bazowych punkt, którego wartość z_0 zostanie wyliczona.



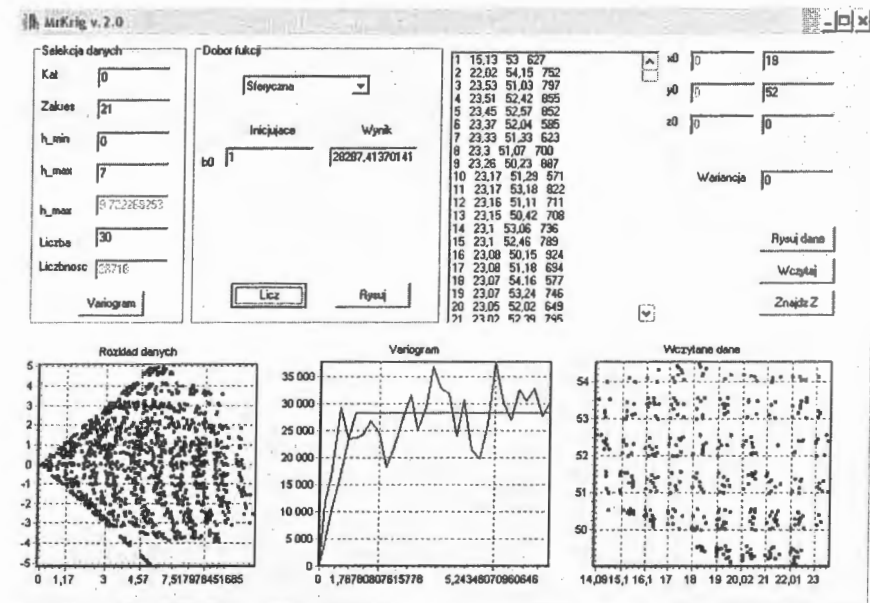
Rys. 9. Wczytanie punktów bazowych i wybór współrzędnych x i y punktu szukanego

Następnie w części *Selekcja danych* należy wybrać kąt α (okno dialogowe **Kąt**) oraz tolerancję kąta (okno dialogowe **Zakres**), odległość minimalną między punktami w parze (okno dialogowe **h_min**), odległość maksymalną między punktami w parze (okno dialogowe **h_max**) oraz liczbę podziałów wartości h , czyli liczbę lagów (okno dialogowe **Liczba**). Po naciśnięciu przycisku **Variogram** program wyświetla na lewym rysunku *Rozkład danych* wszystkie pary z wybranego wcześniej przedziału oraz Variogram na środkowym rysunku, wyliczony ze zbioru wybranych par. W oknie dialogowym *Liczbnosc* wyświetlana jest liczebność par punktów z wybranego wcześniej przedziału (por. rys. 10).



Rys. 10. Rysowanie variogramu empirycznego

Następnie w części *Dobór funkcji* należy wybrać model i wartości inicjujące parametrów b . W programie zaimplementowano siedem modeli: sferyczny, gaussowski, eksponentialny, kwadratowy, r kwadratowy, liniowy i wielomianowy (por. rys. 5). Modele mają od jednego do czterech parametrów b . Po naciśnięciu przyciska **Licz** zostają automatycznie dobrane najlepsze parametry b , przy wykorzystaniu metody Marquardt i wyświetlone w oknie dialogowym *Wynik*. Również po naciśnięciu przyciska **Licz** na rysunku środkowym pojawia się wykres modelu funkcji. Program umożliwia zmianę wyliczonych parametrów b (można to wykonać ręcznie w oknie dialogowym *Wynik*) i narysowania zmienionego modelu funkcji za pomocą przycisku **Rysuj** (por. rys. 11).



Rys. 11. Wybór i rysowanie modelu funkcji.

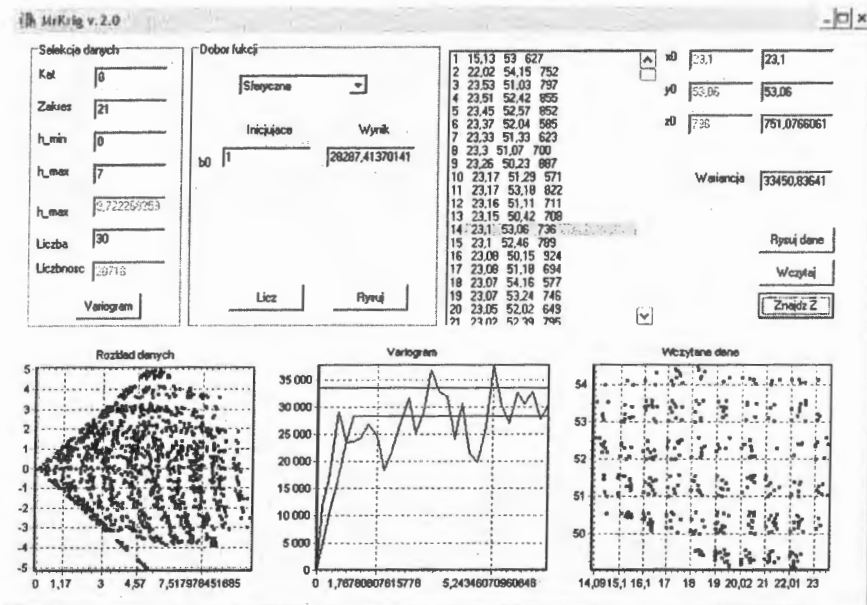
Ostatnią czynnością jest wyliczenie wartości z_0 za pomocą przycisku **Znajdź Z**. Program umożliwia wybór innego szukanego punktu i wyliczenie wartości z_0 bez ponownego powtarzania całej procedury. Wystarczy wybrać z listy jeden punkt lub ręcznie wpisać współrzędne x i y . Oczywiście wybrany szukany punkt przestaje należeć do zbioru punktów bazowych (por. rys. 12)

Program również wylicza wariancję ze wzoru:

$$\text{var} = \frac{\sum_{i=1}^N z_i^2}{N} - \left(\frac{\sum_{i=1}^N z_i}{N} \right)^2$$

gdzie : N – liczba punktów w wybranym przedziale

Wariancja pojawia się na środkowym wykresie jako niebieska linia. (por. rys. 12)



Rys. 12. Wyliczenie szukanej wartości z_0 , narysowanie i podanie wartości wariancji.

4. Wyniki pomiarowe

Za pomocą opracowanego programu wykonano wstępne obliczenia aproksymacyjne dla danych pomiarowych dotyczących rocznych opadów deszczowych w jednym roku w wybranych punktach Polski. Badaniami objęto 399 stacji pomiarowych rozmieszczonych w miarę równomiernie na terenie kraju (por. rys. 1). Pomiarzy deszczu dotyczą 1980 roku. W obliczeniach stosowano do aproksymacji siedem typów modeli: sferyczny, gaussowski, eksponentialny, kwadratowy, r kwadratowy, liniowy i wielomianowy (por. rys. 5).

Badania aproksymacji dotyczyły jednego punktu pomiarowego, a pozostałe 398 punktów potraktowano jako punkty bazowe.

Wykonano kilka wariantów badań, w których zmieniano

- kąt α ,
- tolerancję kąta α ,
- zakres odległości między punktami h_{max} ,
- liczbę lagów

Stacja pomiarowa nr 14

$x - 23,10$

$y - 53,06$

$z - 736$

Wariant a)

Selekcja danych:

Kąt	0
Zakres	21
h_{\min}	0
h_{\max}	7
Liczba lagów	30

Modele:

Wzorcowa	736
Sferyczny	751,0766061
Gaussowski	882,1476817
Kwadratowy	737,9408207
Kwadratowy R	747,0115396
Wielomianowy	1015,902025
Liniowy	659,2534709
Ekspotencjalny	718,6588219

Wariant b)

Selekcja danych:

Kąt	0
Zakres	10
h_{\min}	0
h_{\max}	7
Liczba lagów	30

Modele:

Wzorcowa	736
Sferyczny	830,50542658
Gaussowski	856,5768293
Kwadratowy	807,17862180
Kwadratowy R	832,0268262
Wielomianowy	877,7818467
Liniowy	815,5029802
Ekspotencjalny	813,3414277

Wariant c)

Selekcja danych:

Kąt	10
Zakres	21
h_{\min}	0
h_{\max}	7
Liczba lagów	30

