

Akademia Górniczo-Hutnicza im. St. Staszica w Krakowie

Geostatystyka – analizy wstępne

Zastosowanie GIS w badaniach przyrodniczych, Ćwiczenie 11

Na podstawie materiałów szkoleniowych ESRI.
Wyłącznie do użytku wewnętrznego AGH.

2021-10-26

Wprowadzenie

Na przełomie lat 50-tych i 60-tych ubiegłego wieku, szwedzki geograf Törsten Hägerstrand wprowadził do geografii analizę ilościową. Analiza danych ilościowych jest jedną z najważniejszych metod pozwalających na zrozumienie i ocenę zjawisk zachodzących w środowisku geograficznym. Mówiąc o metodach ilościowych mamy na myśli głównie różnego rodzaju metody statystyczne z zakresu statystyki opisowej ale także wnioskowania statystycznego. Z metodami ilościowymi mamy do czynienia, gdy własności badanych komponentów i kryteriów analizy są wyrażone za pomocą liczb. Zmienne ilościowe stosowane w analizach geograficznych powinny się cechować następującymi właściwościami: mierzalnością, sumowalnością, niezależnością, zmiennością, relatywnością oraz znanym rozkładem statystycznym (Orłóci 1978; Parysek 1982).

1. Analiza semiwariancji

Podstawowym pojęciem geostatystycznego opisu zmienności parametrów przestrzennych jest zmienna zregionalizowana. Jest ona realizacją pewnej funkcji obrazującej zróżnicowanie analizowanych parametrów w przestrzeni trójwymiarowej. Podstawowym założeniem jest ciągłość tej funkcji w przestrzeni. Główną ideą geostatystyki jest traktowanie zmiennej zregionalizowanej jako zmiennej losowej. W praktyce geologiczno-górnictwej, z której analiza się wywodzi, geostatystyka zajmuje się poszukiwaniem charakterystyk funkcji losowej na podstawie jej pojedynczej realizacji w celu estymacji wartości zmiennej poza punktami opróbowania (Mucha 1994; Namysłowska-Wilczyńska 2006; Urbański 2011; Zawadzki 2011).

Geostatystyczny opis struktury zmienności parametrów realizuje się za pomocą analizy semiwariancji. Wyraża ona wkład średniej wariancji parametru w punktach oddalonych od siebie o odpowiednie przedziały odległości, do całkowitej zmienności populacji próby. Podstawowym estymatorem semiwariogramu jest funkcja Matherona (Wz. 1) (Isaaks & Srivastava 1988, 1989).

$$\gamma_h = \frac{1}{2N_h} \sum_{i=1}^{n_h} (z_{i+h} - z_i)^2 \quad \text{Wz. 1.}$$

gdzie:

z_{i+h}, z_i – wartości analizowanych parametrów w punktach oddległych o h ,
 γ_h – semiwariancja obliczona dla odległości pomiędzy parami punktów równej h ,
 N_h – liczba par punktów oddległych o h ,
 h – odległość (średnia) między parami punktów.

Semiwariancja obliczana na podstawie surowych danych pochodzących opróbowania lub innych charakterystyk komponentów środowiska, nosi nazwę próbkowej lub empirycznej. Wartości semiwariancji próbkowej są funkcją średniej różnicy pomiędzy

wartościami parametrów w punktach oddalonych od siebie o ten sam przedział odległości oraz funkcją kierunku obserwacji. W przypadku kiedy zaniedbujemy ten kierunek i obliczamy uśrednienie wartości semiwariancji pomiędzy punktami z całej badanej przestrzeni mówimy o semiwariancji wielokierunkowej lub izotropowej. Innym sposobem obliczenia semiwariancji jest pomiar zmienności parametrów wzdłuż ściśle określonego kierunku obserwacji. Jego zdefiniowaniu towarzyszy określenie zakresu kąтового, odkładanego od zadanego kierunku obserwacji i wyznaczającego obszar obliczeniowy. Semiwariancje obliczane wzdłuż zadanych kierunków obserwacji noszą nazwę kierunkowych lub anizotropowych.

Wartości semiwariancji obliczone dla większej liczby przedziałów odległości są przedstawiane w sposób graficzny w postaci wykresów nazywanych semiwariogramami (Fig. 1). Na osi rzędnych semiwariogramów przedstawiana jest semiwariancja próbkowa, a na osi odciętych – odległości pomiędzy punktami. W typowych sytuacjach, semiwariogramy biorą swój początek od pewnej pierwotnej, losowej wartości zmienności lokalnej nazywanej efektem samorodków – C_0 (ang.: *nugget effect*). Wraz ze wzrostem odległości pomiędzy analizowanymi punktami, wartości semiwariancji parametrów w pewnym zakresie rosną. Dzieje się tak z powodu tzw. autokorelacji wyrażającej większe prawdopodobieństwo wystąpienia podobnych wartości parametrów w domenach położonych blisko siebie niż położonych w większym oddaleniu od siebie (Tobler 1970). Opisywana zmienność ma charakter deterministyczny i wynika wyłącznie z charakteru zmiennej zregionalizowanej. Osiągany maksymalny poziom semiwariancji wyznacza wartość amplitudy semiwariogramu – C (ang. *sill*). Zasięg – a (ang. *range*), nazywany też promieniem autokorelacji, określa odległość, powyżej której wartości parametrów nie wykazują przestrzennej autokorelacji.

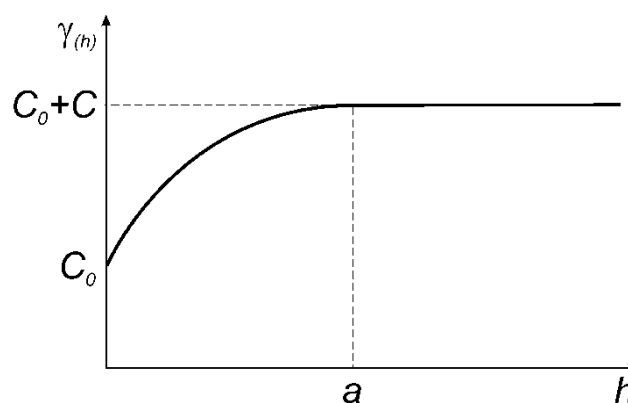


Fig. 1. Typowy semiwariogram. C_0 – zmienność lokalna; $C_0 + C$ – amplituda; a – zasięg

Jednym z podstawowych założeń geostatystycznego opisu zmienności parametrów jest założenie słabej (bądź szerokiej) stacjonarności. Zakłada ono, że wartość oczekiwana zregionalizowanej zmiennej losowej jest stała na całym analizowanym obszarze oraz, że przestrzenna kowariancja zregionalizowanej zmiennej losowej jest jedynie funkcją

odległości pomiędzy punktami opróbowania (Mucha 1994; Zawadzki 2011). W praktyce analiz danych środowiskowych, opisywane założenia są niekiedy trudne do spełnienia. Ocenę struktury zmienności różnorodnych wskaźników krajobrazowych utrudniają anizotropia polegająca na występowaniu trendów zmienności analizowanych parametrów zależnych od kierunków obserwacji oraz koincydencje różnych czynników wpływających na zmienność badanych parametrów. Prezentowane ograniczenia nie wykluczają przydatności metody w jakościowej identyfikacji struktur komponentów krajobrazu (Meisel & Turner 1998).

Krzywe semiwariogramów empirycznych są opisywane za pomocą szeregu dozwolonych modeli – funkcji matematycznych (liniowych, sferycznych, potęgowych, kubicznych i innych) (Mucha 1994). Umożliwiają one identyfikację zasięgu poziomu zmienności lokalnej, amplitudy i zasięgu autokorelacji.

Zbiór semiwariogramów kierunkowych obliczony dla pełnego zakresu obserwacji i przedstawiony na płaszczyźnie w postaci wykresu z centralnym punktem oznaczającym zerową odległość pomiędzy miejscami opróbowania pozwala zbadać czy zmienność analizowanego parametru ma charakter izotropowy czy też anizotropowy. Iloraz zasięgów semiwariogramów jednego typu, obliczonych dla dwóch niezależnych kierunków daje informację na temat siły anizotropii geometrycznej i nosi nazwę współczynnika anizotropii (Wz. 2). Najczęściej definiuje się go jako stosunek zasięgu semiwariogramu dla kierunku minimalnej zmienności badanego parametru do zasięgu semiwariogramu dla kierunku jego maksymalnej zmienności. Przyjmuje się, że współczynnik anizotropii mniejszy od 2 jest uważany za łagodny, natomiast większy od 4 jest uważany za ostry (ciężki). Zwykle kiedy współczynnik anizotropii jest większy od 3 efekt jest wyraźnie zauważalny na mapach (Kokesz & Mucha 1984; Armstrong 1998; Surfer 8. User's Guide 2002).

$$A = \frac{R_1}{R_2} \quad \text{Wz. 2.}$$

A – współczynnik anizotropii,
 R_1, R_2 – zasięgi semiwariogramów wzdłuż dwóch badanych kierunków zmienności.

Klasyczne estymatory bezwzględne (wielokierunkowe lub kierunkowe) wyrażane są w jednostkach wariancji badanych parametrów i nie pozwalają na porównywanie poziomów zmienności parametrów mierzonych w różnych jednostkach.

Wymagane oprogramowanie: ArcGIS 10.X for Desktop (ArcView, ArcEditor, lub ArcInfo).

Ćwiczenie 7

Dzięki modułowi geostatystycznemu ArcGIS (ArcGIS Geostatistical Analyst) można łatwo, na podstawie punktowy danych pochodzących np. z opróbowania utworzyć ciągłą mapę powierzchni analizowanej cechy. Punktowe dane mogą być przechowywanych w klasie obiektów o geometrii punktowej lub warstwie rastrowej. Mogą być także z łatwością przypisane do klasy centroid utworzonej dla klasy obiektów powierzchniowych. Danymi próbkowymi mogą być takie cechy jak: elewacja; głębokość warstwy wodonośnej lub poziomu zanieczyszczenia gleb, ale mogą to być także inne cechy o typie zmiennych zregionalizowanych. W połączeniu z ArcMap, Geostatistical Analyst zapewnia kompleksowy zestaw narzędzi do tworzenia powierzchni, które można wykorzystać do wizualizacji, analizy i zrozumienia zjawisk przestrzennych.

Amerykańska Agencja Ochrony Środowiska (*U.S. Environmental Protection Agency*) jest w Kalifornii odpowiedzialna za monitorowanie stężenia ozonu w atmosferze. Stężenie ozonu mierzy się w stacjach monitoringu rozmieszczonych w całym stanie (Fig. 2). Znamy punktowe poziomy stężenie ozonu w obrębie wszystkich stacji, ale chcielibyśmy poznać wartości ozonu także dla innych (nie obserwowanych) miejsc Kalifornii. Ze względu na koszty i kwestie praktyczne, stacje monitorujące nie mogą być umieszczane wszędzie. Geostatystyka jest działem statystyki przestrzennej dostarczającym narzędzi umożliwiających generowanie wiarygodnych prognoz. Generowanie wiarygodnych powierzchni odzwierciedlających zmienność analizowanego parametru (tutaj stężenia ozonu) w dowolnym miejscu badanej przestrzeni jest możliwe dzięki analizie zależności występujących pomiędzy wszystkimi punktami opróbowania. Geostatystyka umożliwia także wykonanie mapy błędów (niepewności) przewidywań i prawdopodobieństw, przekroczenia wartości krytycznych analizowanych parametrów.

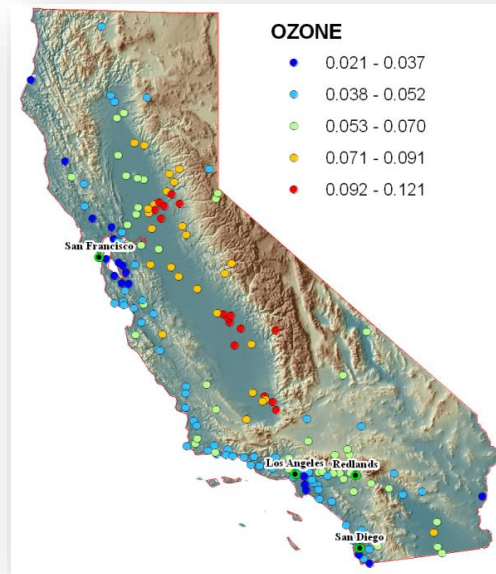


Fig. 2. Położenie stacji monitoringu atmosfery

2. Dane

Dane do ćwiczeń są dostępne w lokalizacji [tutaj](#). Pobierz je, a następnie skopiuj na dysk lokalny komputera, do folderu domowego i podfolderu Geostatystyka. Dane niezbędne do wykonania ćwiczenia znajdują się w geobazie `ca_ozone.gdb`. W ćwiczeniu zostaną użyte następujące zbiory danych:

`ca_outline` – granice stanu Kalifornia,
`O3_Sep06_3pm` – punktowe dane stężenia ozonu [ppm]¹ pochodzące ze stacji monitoringu atmosfery,
`ca_cities` – główne miasta Kaliforni,
`ca_hillshade` – cieniowany relief Kaliforni.

Zbiór danych pomiarów zawartości ozonu w atmosferze (`O3_Sep06_3pm`) został udostępniony dzięki uprzejmości *California Air Resources Board* i stanowi stężenie ozonu zmierzone w dniu 6 września 2007 r., w godzinach od 3:00-16:00, w częściach na milion (ppm).

Na podstawie pomiarów próbek zawartości ozonu w atmosferze otrzymamy dwie ciągle powierzchnie (dwie mapy) przewidujące zawartości stężenia ozonu w każdej lokalizacji w stanie Kalifornia. Pierwsza z utworzonych map zostanie wykonana z użyciem wszystkich domyślnych opcji. Pozwoli nam to zapoznać się z procesem generowania powierzchni na podstawie punktowych danych próbkowych. Druga mapa, którą otworzymy, umożliwi włączenie do analizy większej liczby powiązań przestrzennych

¹ ppm – liczba części na milion (ang. *parts per million*) – 10^{-6} .

pomiędzy danymi próbkowymi. Podczas tworzenia tej mapy użyjemy narzędzi do analizy danych przestrzennych (ESDA). Wprowadzimy też niektóre dodatkowe opcje analizy geostatystycznej, takie jak usuwanie trendów i modelowanie autokorelacji przestrzennej, które można wykorzystać do tworzenia powierzchni. Wykorzystanie narzędzi ESDA i dodatkowych parametrów analizy geostatystycznej, pozwoli na utworzenie bardziej dokładnych powierzchni. Często w analizach nie poszukujemy rzeczywistych wartości parametrów ale raczej informacji czy wartości są powyżej pewnego poziomu toksycznego, czy też nie. Jeśli wartości parametrów uległy przekroczeniu, należy podjąć natychmiastowe działania. Trzecia powierzchnia, którą utworzymy, oceni prawdopodobieństwo zdarzenia, że krytyczna wartość progowa stężenia ozonu została przekroczona. Jako wartość krytyczną zdefiniujemy zawartość ozonu przekraczającą 0,09 ppm (norma jakości powietrza w Kalifornii dla pomiarów godzinowych). Miejsca przekroczenia norm powinno być poddane stałemu monitoringowi. Użyjemy narzędzi analizy geostatystycznej do przewidywania prawdopodobieństwa, że wartości ozonu na obszarze Kalifornii przestały spełniać ten standard w dniu 6 września 2007 r., w godzinach od 3:00 do 16:00.

Ćwiczenie zostało podzielone na pięć etapów (w kolejnych dokumentach pdf):

- ✓ Zaznajomienie się z etapami tworzenia modeli interpolacyjnych, wprowadzenie w tematykę Geostatistical Analyst i proces tworzenia map stężenia ozonu przy użyciu domyślnych wartości parametrów.
- ✓ Wprowadzenie w proces wyszukiwania danych przed utworzeniem powierzchni interpolacyjnej w celu wykrycia nietypowych danych (próby huraganowe) i proces rozpoznawania trendów.
- ✓ Utworzenie drugiej powierzchni obejmującej więcej powiązań przestrzennych, odkrytych w etapie 2 i poprawa powierzchni utworzonej podczas etapu 1. Etap ten wprowadza także w niektóre podstawowe pojęcia geostatystyki.
- ✓ Porównanie wyników interpolacji dwóch powierzchni utworzonych w ćwiczeniach 1 i 3 oraz sposobu określania, która powierzchnia zapewnia lepszą estymację parametru.
- ✓ Proces odwzorowania prawdopodobieństwa, zdarzenia polegającego na przekroczeniu krytycznej wartości progowej stężenia ozonu, utworzenie trzeciej powierzchni.

Tworzenie powierzchni interpolacyjnych przy użyciu parametrów domyślnych

Ta część ćwiczenia wprowadza w zagadnienia rozszerzenia *Geostatistical Analyst*. W celu wygenerowania powierzchni stężenia ozonu, przy użyciu wartości domyślnych narzędzi geostatystycznych utworzymy prosty model interpolacyjny.

3. Uruchom ArcMap i włącz Geostatistical Analyst

Na początek uruchom ArcMap i włącz rozszerzenie *Geostatistical Analyst*.

- 3.1. Uruchom program *ArcMap*.
- 3.2. W oknie dialogowym *ArcMap – Rozpoczęcie Pracy* (*ArcMap – Getting Started*) kliknij przycisk *Anuluj* (*Cancel*).
- 3.3. W menu głównym kliknij polecenie *Dostosuj* (*Customize*) > *Rozszerzenia* (*Extensions*).
- 3.4. Zaznacz pole wyboru *Geostatistical Analyst*.
- 3.5. Kliknij przycisk *Zamknij* (*Close*).

4. Dodanie paska narzędzi Geostatistical Analyst


- 4.1. W menu głównym kliknij polecenie *Dostosuj* (*Customize*) > *Paski Narzędzi* (*Toolbars*) > *Geostatistical Analyst*.

Pasek narzędzi *Geostatistical Analyst* został dodany do sesji ArcMap.

Rozszerzenie i pasek narzędzi muszą być włączone i dodawane tylko raz; będą one aktywne i pojawiają się następnym razem, gdy włączysz ArcMap.

5. Dodanie danych do sesji ArcMap

Dodamy teraz dane i zmienimy ich symbolikę.

- 5.1. Kliknij przycisk *Dodaj dane* (*Add data*)  znajdujący się na standardowym pasku narzędzi.
- 5.2. Przejdź do folderu, do którego zostały skopiowane dane do ćwiczenia np.:

D:\MNwNP\nazwisko_imie\Geostatystyka.

- 5.3. Aby przejrzeć zawartość geobazy *ca_ozone.gbd*, dwukrotnie kliknij ją.
- 5.4. Przytrzymaj klawisz *CTRL* i wybierz zestawy danych *O3_Sep06_3pm* i *ca_outline*.
- 5.5. Kliknij przycisk *Dodaj* (*Add*).

- 5.6. Kliknij w tabeli zawartości ppm legendę warstwy *ca_outline* (pole poniżej nazwy warstwy).
- 5.7. Kliknij przycisk *Bez koloru* (No color) (Fig. 3).

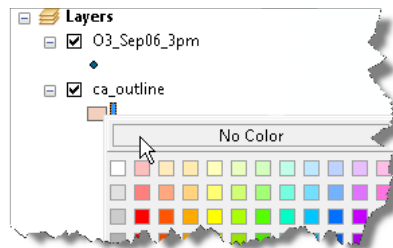


Fig. 3. Wybór stylu wypełnienia warstwy *ca_outline*

Zostanie wyświetlony tylko zarys stanu Kalifornia. Pozwala to zobaczyć warstwy, które zostaną utworzone w ćwiczeniu pod tą warstwą.

- 5.8. W tabeli zawartości kliknij dwukrotnie nazwę warstwy *O3_Sep06_3pm*.
- 5.9. W oknie dialogowym *Właściwości Warstwy* (Layer Properties) kliknij zakładkę *Symbolika* (Symbology).
- 5.10. W polu *Pokaż* (Show) kliknij pozycję *Ilość* (Quantities), a następnie *Stopniowanie kolorów* (Graduated colors).
- 5.11. W polu dialogowym *Pola* (Fields) ustaw wartość atrybutu *OZONE*.
- 5.12. Wybierz skalę kolorów (Color Ramp) od białej do czarnej (Fig. 4).

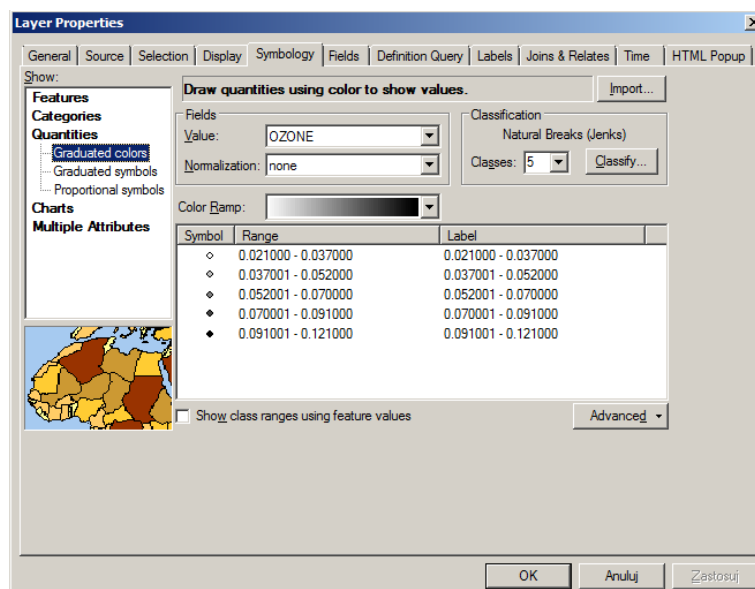


Fig. 4. Okno dialogowe *Właściwości Warstwy*, zakładka *Symbolika* z symbolizacją warstwy stacji monitoringu powietrza

- 5.13. Kliknij przycisk *OK*.

Warto zauważyć, że 6 września 2007 r. najwyższe zawartości ozonu występowały w Dolinie Centralnej (Fig. 5), a najniższe – wzdłuż wybrzeża.

Mapowanie danych są pierwszymi krokami do ich eksploracji i zrozumienia zjawiska, które zamierzamy modelować.



Fig. 5. Położenie Doliny Centralnej w Kalifornii (na żółto)

6. Zachowanie dokumentu mapy

Zaleca się możliwie częste zapisanie dokumentu mapy w trakcie każdego ćwiczenia oraz po zakończeniu każdego etapu ćwiczenia.

- 6.1. W menu głównym kliknij polecenie *Plik (File) > Zapisz (Save)*.
- 6.2. Przejdź do folderu roboczego D:\Mnwnp\nazwisko_imie\Geostatystyka.
- 6.3. W polu tekstowym *Nazwa pliku* wpisz *Ozone Prediction Map.mxd*.
- 6.4. Kliknij przycisk *Zapisz*.

7. Tworzenie powierzchni interpolacyjnej za pomocą opcji domyślnych

Utworzymy teraz (wyinterpolujemy) powierzchnię stężenia ozonu. Będziemy używać ustawień domyślnych narzędzi *Geostatistical Analyst*. Jako zbioru danych wejściowych użyjemy zbioru punktowych pomiarów stężenia ozonu w stacjach kontroli powietrza *O3_Sep06_3pm*. Do utworzenia powierzchni wykorzystamy metodę krigingu. Aby zaakceptować ustawienia domyślne, w wielu polach dialogowych będziemy klikać przycisk *Dalej*.

W tym etapie ćwiczenia nie musisz koncentrować się na szczegółach okien dialogowych. Szerzej zapoznamy się z nimi w późniejszych etapach ćwiczenia.

- 7.1. Na pasku narzędzi *Geostatistical Analyst* kliknij listę *Geostatistical Analyst* i kliknij przycisk *Kreator Geostatystyczny (Geostatistical Wizard)* (Fig. 6).

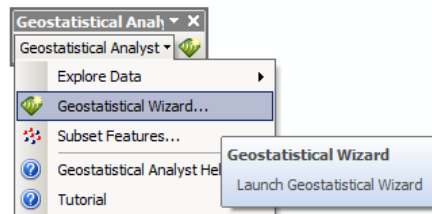


Fig. 6. Zestaw narzędzi *Geostatistical Analyst*

Zostanie wyświetlone okno dialogowe *Geostatistical Wizard* (Fig. 7).

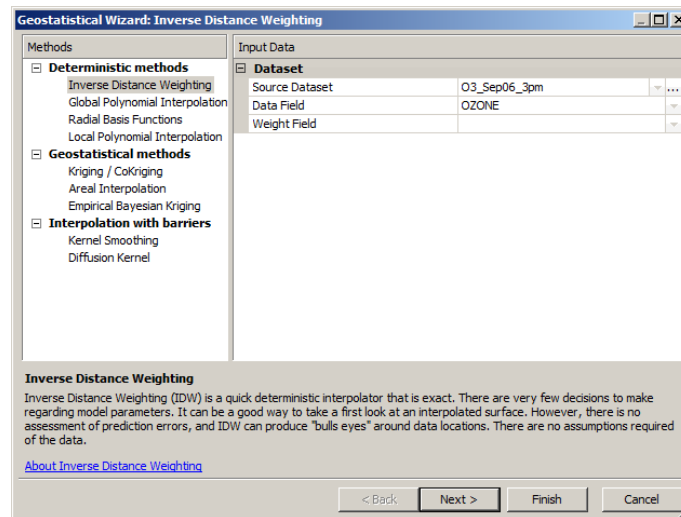


Fig. 7. Okno dialogowe *Geostatistical Wizard*

- 7.2. W polu listy *Metody (Methods)* Kliknij *Kriging/CoKriging*.
- 7.3. Z listy rozwijanej *Źródła Danych (Source Dataset)* wybierz zbiór O3_Sep06_3pm.
- 7.4. Z listy rozwijanej *Pole Danych (Data Field)* wybierz atrybut OZONE (Fig. 8).

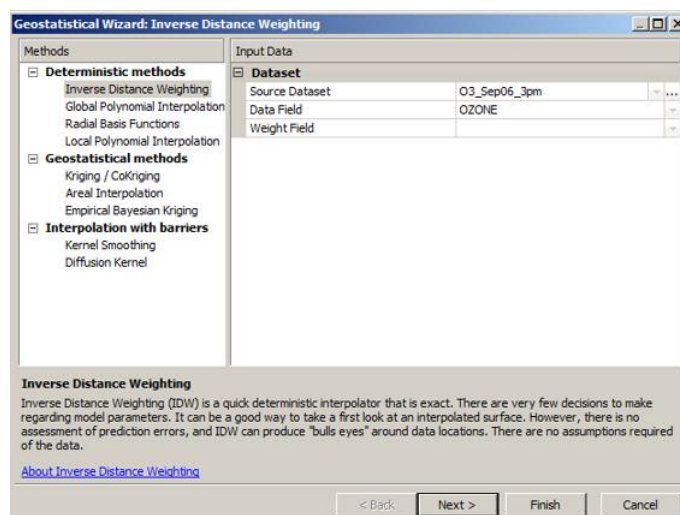


Fig. 8. Okno dialogowe *Geostatistical Wizard* z wybranym zbiorem danych i atrybutem zawierającym wartości zmiennej zregionalizowanej

- 7.5. Kliknij *Dalej (Next)*.
- 7.6. Jako rodzaj kriginu wybieramy *Kriging zwyczajny (Ordinary)*, zaś jako rodzaj powierzchni wynikowej *Mapa predykcji (Prediction)*.
- 7.7. Ponieważ metoda mapowania powierzchni stężenia ozonu jest zaznaczona, aby utworzyć powierzchnię przy użyciu domyślnych parametrów, możnaby kliknąć przycisk *Zakończ (Fig. 9)*.

To jednak nie koniec. Pozostają do wykonania kroki od 6 do 10.

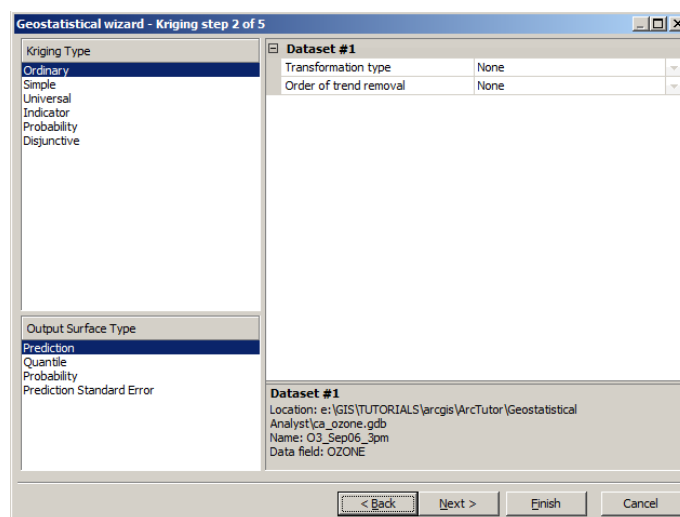


Fig. 9. Okno dialogowe *Geostatistical Wizard* z określonymi: metodą interpolacji oraz rodzajem powierzchni wynikowej

Należy zauważyć, że w prawym dolnym rogu Kreatora Geostatystycznego (Fig. 9) znajduje się pole zawierające krótki opis podświetlonej metody lub parametru. Na tym etapie pole to pokazuje nazwę zbioru danych i nazwę wybranego atrybutu.

- 7.8. Kliknij *Dalej (Next)*.

Zostanie wyświetlony model semiwariogramu/kowariancji, umożliwiający zbadanie relacji przestrzennych zachodzących pomiędzy danymi sieci opróbowania (w tym przypadku wartościami ozonu pomierzonymi przez stacje monitoringu atmosfery). Pierwsze prawo geografii Toblera (Tobler 1970), które definiuje relacje podobieństwa pomiędzy obiektami (ich cechami) w przestrzeni mówi, że w geografii wszystko jest związane ze wszystkim innym, przy czym rzeczy bliższe są bardziej ze sobą związane niż rzeczy bardziej odległe. Nazywane jest to autokorelacją przestrzenną. Semiwarogram pozwala na zbadanie tego założenia (Fig. 10). Proces dopasowania modelu teoretycznego do semiwariogramu empirycznego, w celu uchwycenia relacji przestrzennych zawartych w danych jest nazywany wariografią.

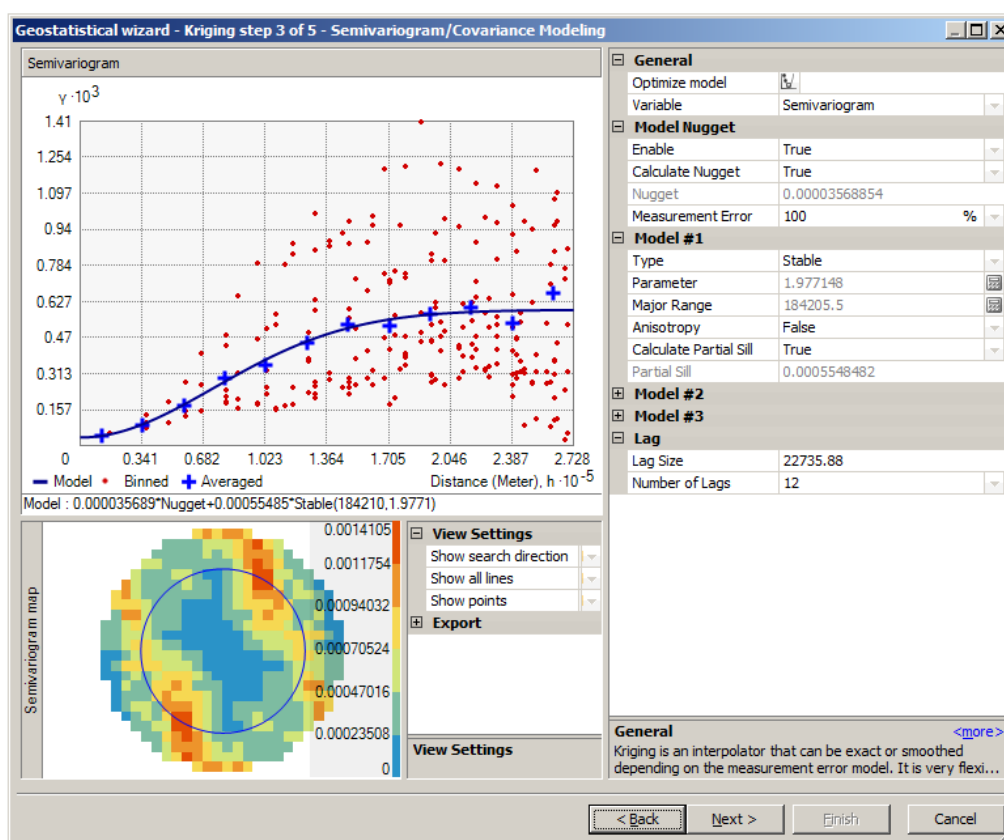


Fig. 10. Okno dialogowe z wyświetlonym semiwariogramem empirycznym (jako modelem relacji przestrzennych zachodzących pomiędzy danymi) oraz mapą semiwariogramów kierunkowych

7.9. Kliknij *Dalej (Next)*.

Krzyż wskazuje lokalizację, która nie ma pomierzonej wartości próbkowej (Fig. 11). Aby przewidzieć wartość analizowanego parametru (w tym przypadku stężenia ozonu) w miejscu krzyża, musimy użyć znanych wartości w punktach pomiaru (w tym przypadku w stacjach monitoringu). Zgodnie z zasadą autokorelacji Toblera, wartość analizowanego parametru w lokalizacji pozbawionej punktu pomiarowego jest najbardziej zbliżona do znanych wartości parametru z punktów pomiaru położonych od niego najbliższej. Czerwone punkty na obrazku poniżej będą ważone (lub będą miały wpływ na nieznaną wartość)

większą niż zielone punkty, ponieważ znajdują się bliżej przewidywanej lokalizacji. Używając punktów otaczających i modelu semiwariogramu/kowariancji, można przewidzieć wartości dla każdej niezmierzonej lokalizacji.

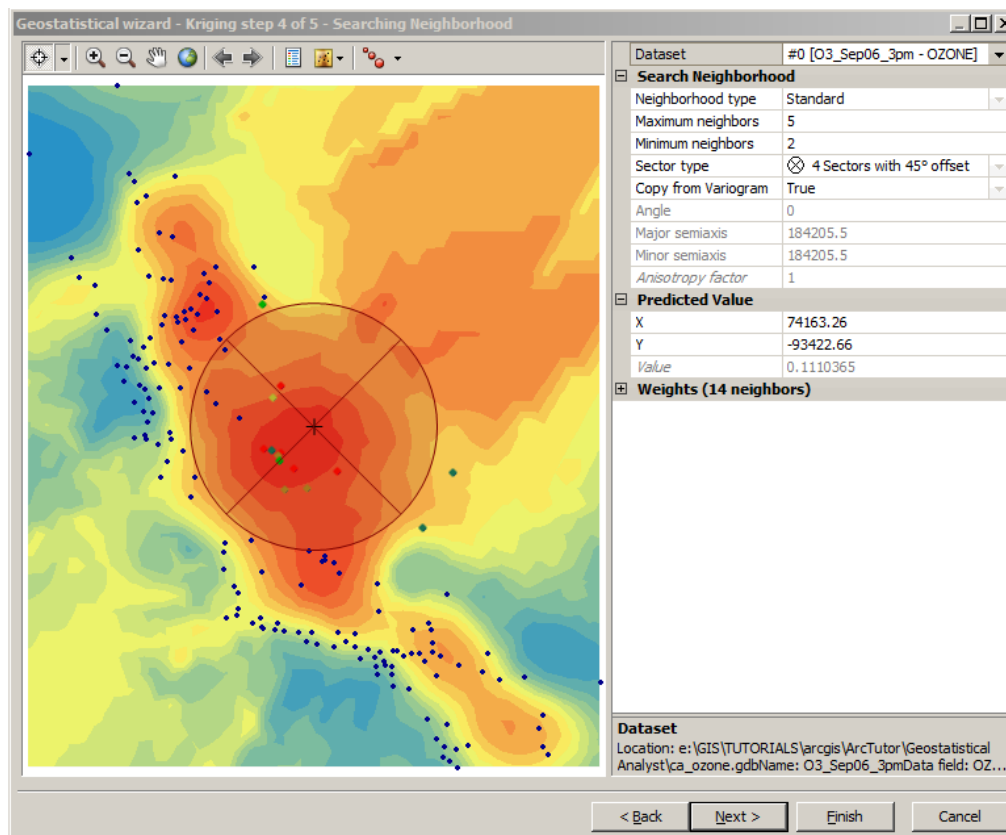


Fig. 11. Okno dialogowe poszukiwania punktów sąsiednich

7.10. Kliknij *Dalej (Next)*.

Diagram weryfikacji krzyżowej (tzw. cross-walidacja – ang. *cross-validation*) pozwala zweryfikować poprawność doboru modelu estymującego wartości badanego parametru w nieznanach lokalizacjach (Fig. 12).

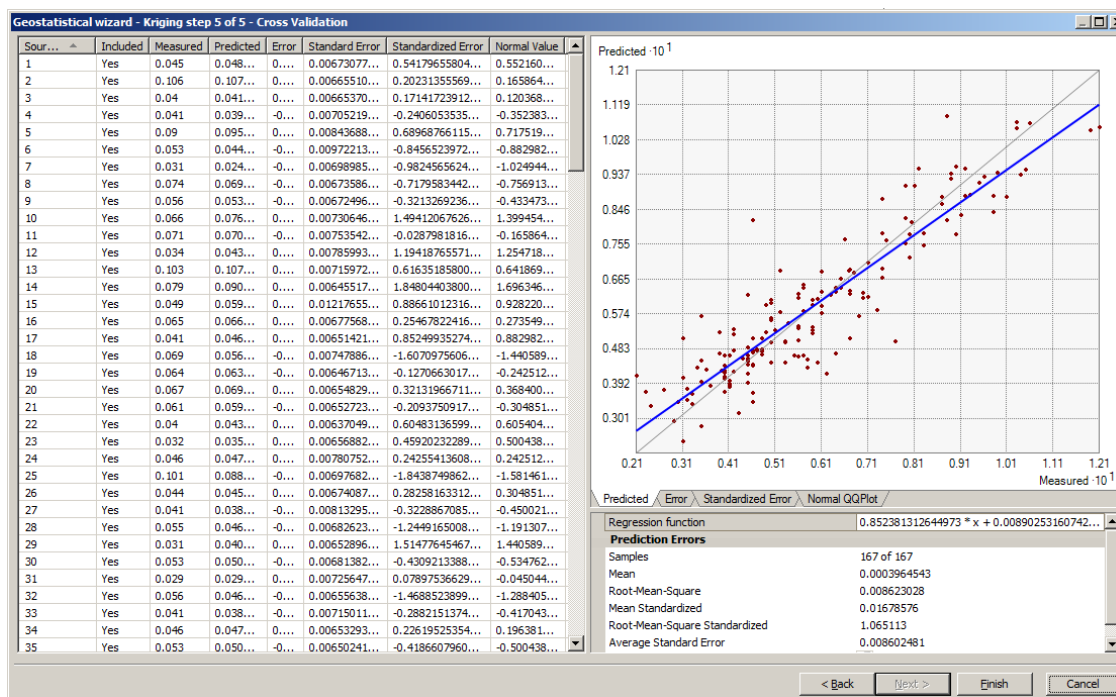


Fig. 12. Okno dialogowe cross-walidacji

Jak z niego korzystać i jak rozumieć wyświetlone statystyki dowiemy się w 4-tym etapie ćwiczenia.

7.11. Kliknij przycisk *Zakończ*.

W oknie dialogowym *Raport Metody (Method Report)* podsumowano informacje dotyczące metody (i powiązanych z nią parametrów), która będzie wykorzystana do utworzenia powierzchni interpolacyjnej (Fig. 13).

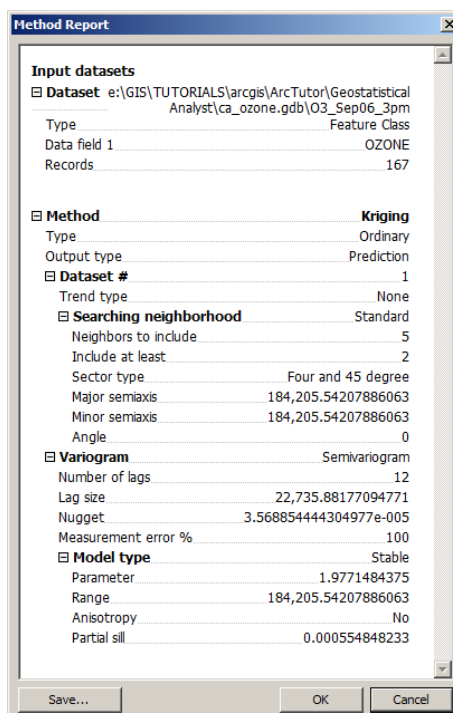


Fig. 13. Raport modelowania geostatystycznego

7.12. Kliknij przycisk *OK*.

Mapa zawartości ozonu zostaje dodawana jako górna warstwa do tabeli zawartości (Fig. 14).

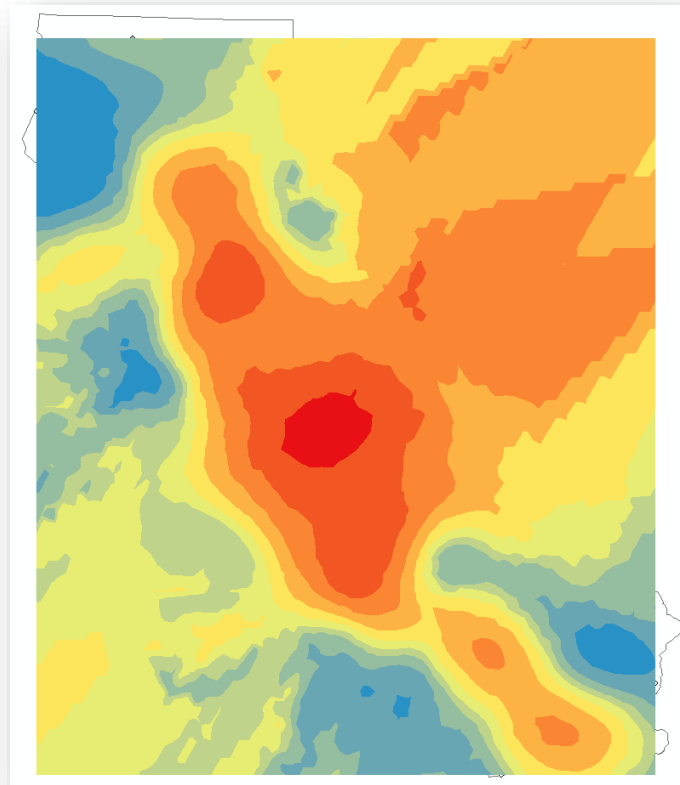


Fig. 14. Mapa zmienności zawartości ozonu w powietrzu dodana do sceny ArcMap

- 7.13. Aby otworzyć okno dialogowe *Właściwości Warstwy*, kliknij dwukrotnie warstwę w tabeli zawartości.
- 7.14. Kliknij zakładkę *Ogólne (General)* i zmień nazwę warstwy na *Default Kriging* (Fig. 15).

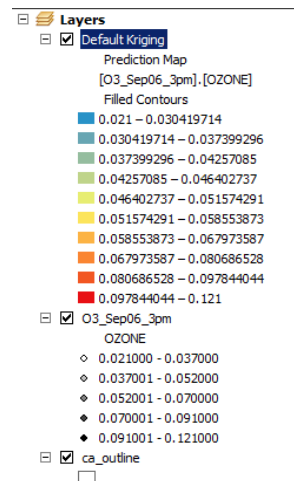


Fig. 15. Tabela zawartości z interpolowaną mapą zmienności ozonu (Default Kriging)

7.15. Kliknij przycisk *OK*.

Zmiana nazwy warstwy pomoże nam odróżnić tę warstwę od tej, którą utworzymy w 4-tym etapie ćwiczenia.

7.16. Aby zapisać plik projektu, kliknij ikonę *Zapisz* na standardowym pasku narzędzi.

Zauważ, że interpolowana powierzchnia jest kontynuowana także w oceanie. Dzieje się tak dlatego, że zasięg warstwy jest taki sam jak zakres danych wejściowych (O3_Sep06_3pm).

7.17. Aby ograniczyć powierzchnię predykcji do obszaru Kalifornii, kliknij ppm warstwę *Default Kriging* i kliknij przycisk *Właściwości Warstwy (Layer Properties)*.

7.18. Kliknij zakładkę *Zakres (Extent)*.

7.19. Kliknij strzałkę *Ustaw zakres na (Set the extent to:)*, wybierz opcję *prostokątny rozmiar ca_outline (the rectangular extent of ca_outline)*, a następnie kliknij przycisk *OK*.

Obszar interpolowany rozciąga się tak, aby obejmował całą Kalifornię.

7.20. W tabeli zawartości kliknij ppm ramkę danych *Layers*, kliknij polecenie *Właściwości (Properties)*, a następnie wybierz zakładkę *Ramka Danych (Data Frame)*.

7.21. W polu *Opcje Wycinania (Clip Options)*, wybierz opcję *Wytnij do kształtu (Clip to shape)*, a następnie kliknij przycisk *Określ Kształt... (Specify Shape...)*.

- 7.22. W oknie dialogowym *Przycinanie Ramki Danych (Data Frame Clipping)*, wybierz opcję *Zarys Obiektów (Outline of Features)*, z menu rozwijanego *Warstwa (Layer)* wybierz warstwę `ca_outline`.
- 7.23. Kliknij przycisk *OK*, a następnie ponownie kliknij *OK*.

Interpolowana powierzchnia jest przycięta i nie wyświetla danych spoza granic stanu (Fig. 16).

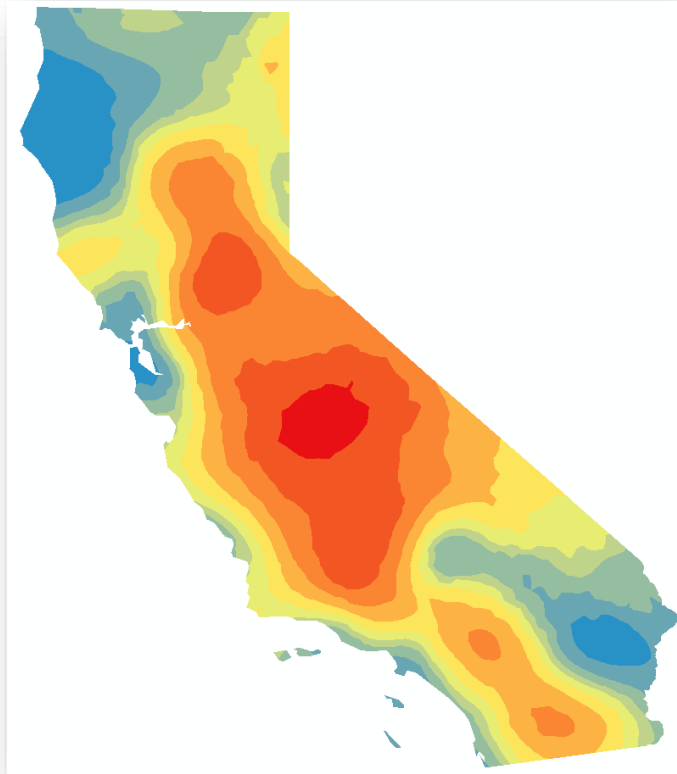


Fig. 16. Interpolowana powierzchnia przycięta do granic stanu Kalifornia

- 7.24. Przeciągnij warstwę `03_Sep06_3pm` na początek tabeli zawartości.

Wizualnie oceń, jak dobrze warstwa `Default Kriging` oddaje zmienność mierzonych wartości ozonu. Czy wysokie przewidywania dotyczące ozonu występują w tych samych obszarach, w których zmierzono wysokie stężenia ozonu? W etapach 3 i 4 nauczymy się w jaki sposób ilościowo oceniać, jak dobrze wykonuje się model i jak porównać dwa modele.

- 7.25. W tabeli zawartości, kliknij ppm warstwę `Default Kriging` i z menu kontekstowego wybierz *Validation/Prediction*. Spowoduje to otwarcie narzędzia geoprocesingu *GA Layer To Points* z wybranym źródłem geostatystycznych danych wejściowych `Default Kriging` (Fig. 17).

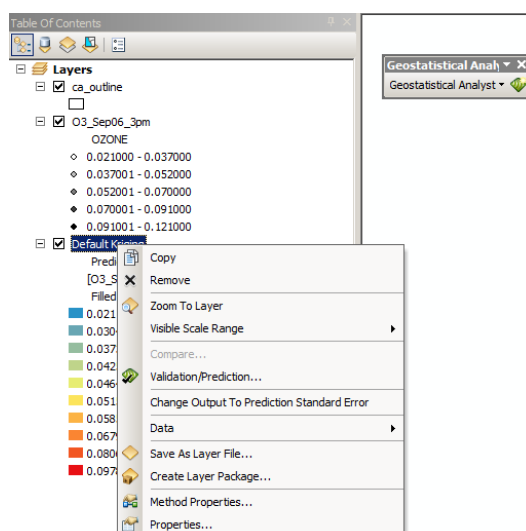


Fig. 17. Tabela zawartości z menu kontekstowym dla interpolowanej powierzchni

- 7.26. W oknie dialogowym *GA Layer To Points*, *Wejściowa warstwa geostatystyczna* (*Input geostatistical layer*) powinna być automatycznie ustawiona na *Default Kriging* (Fig. 18). W polu *Lokalizacje punktów obserwacji* (*Point observation locations*) wybierz geobazę zawierającą dane dla tego ćwiczenia i następnie wskaż zbiór danych *ca_cities*. Pole formularza *Pole do sprawdzenia poprawności (opcjonalne)* (*Field to validate on (optional)*) pozostaw puste, ponieważ chcemy wygenerować prognozy zawartości ozonu w głównych miastach, a nie chcemy sprawdzać przewidywanych wartości względem wartości zmierzonych. W polu *Statystyki wyjściowe w punktach* (*Output statistics at point locations*) przejdź do folderu ćwiczeniowego (tam gdzie jest zapisany plik projektowy *Ozone Prediction Map.mxd*) i nadaj mu nazwę *CA_cities_ozone.shp*.

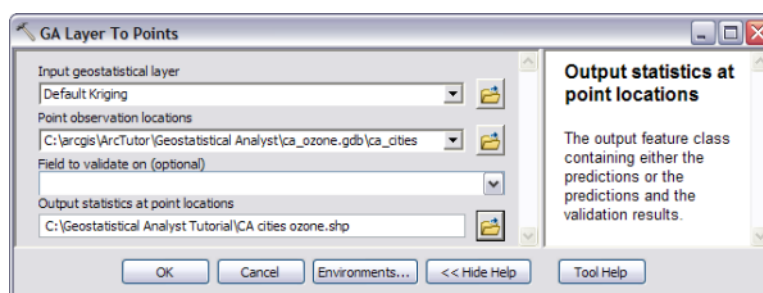


Fig. 18. Okno dialogowe *GA Layer To Points*

- 7.27. Aby uruchomić narzędzie, kliknij przycisk *OK*.

Narzędzie geoprocesingu będzie wykonywało operacje w tle. Postęp pracy będzie można obserwować na pasku stanu ArcMap (u dołu ekranu ArcMap).

- 7.28. Po zakończeniu pracy narzędzia, wybierz ikonę *Dodaj dane* znajdującą się na standardowym pasku narzędzi.

7.29. Przejdź do folderu z danymi i wybierz plik `CA_ozone_cities.shp`, a następnie kliknij przycisk *Dodaj (Add)*

Warstwa punktów obserwacji została dodana do mapy (Fig. 19).

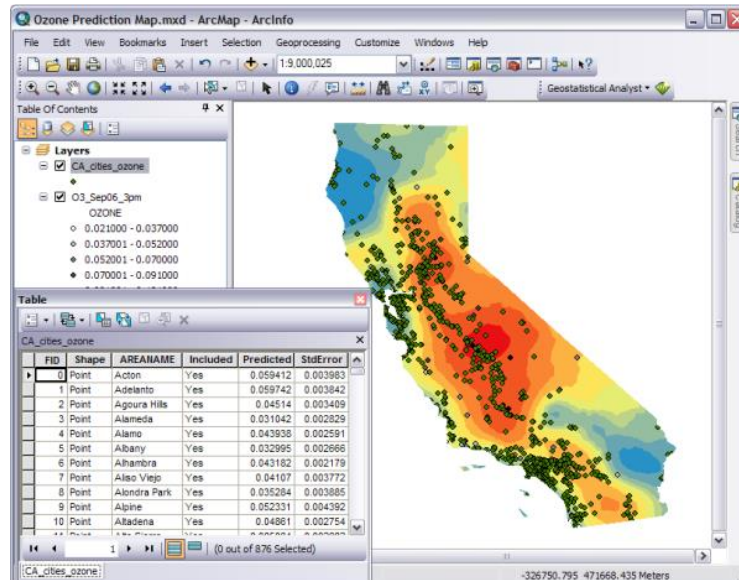


Fig. 19. Scena pliku projektowego z dodaną warstwą punktów obserwacji `CA_ozone_cities.shp`

7.30. W tabeli zawartości kliknij ppm warstwę `CA_cities_ozone` i z menu kontekstowego wybierz polecenie *Otwórz Tabelę Atrybutową (Open Attribute Table)*.

Zauważ, że każde miasto ma przewidywaną zawartość ozonu (atrybut `Prediction`), a także wartość błęd standardowego (atrybut `StdError`), która wskazuje na poziom niepewności związany z estymowaną zawartością ozonu dla każdego miasta).

7.31. Zamknij okno tabeli atrybutowej.

7.32. Aby usunąć warstwę `CA_cities_ozone` z projektu, w tabeli zawartości kliknij na niej ppm i z menu kontekstowego wybierz polecenie *Usuń (Remove)*.

7.33. Zapisz dokument *ArcMap*.

8. Metodologia dopasowania powierzchni

Uworzyliśmy mapę koncentracji ozonu i tym samym ukończyliśmy 1-szy etap ćwiczeń z zakresu modelowania geostatystycznego.

Pomimo tego, że zadanie utworzenia mapy zmienności analizowanego parametru z wykorzystaniem kreatora geostatystycznego wydaje się być proste, ważne jest aby wykonując analizy geostatystyczne przebiegały one w ściśle przestrzeganej kolejności (Fig. 20).

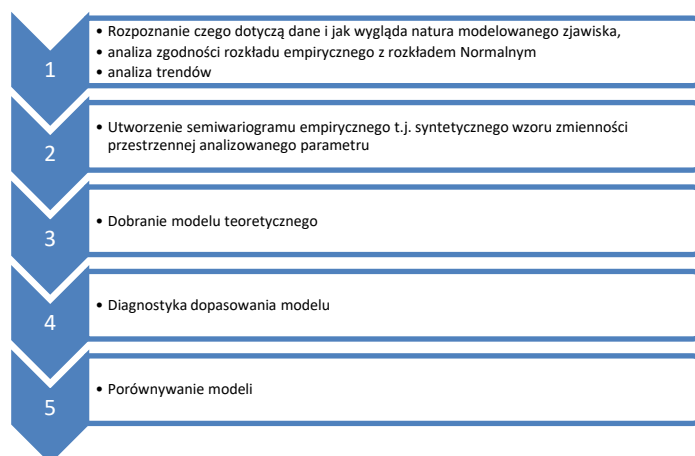


Fig. 20. Główne etapy procedury modelowania geostatystycznego

Podczas następnych etapów tego ćwiczenia będziemy ćwiczyli procedurę modelowania geostatystycznego. Ponadto, w etapie 5-tym, utworzymy powierzchnię pokazującą prawdopodobieństwo, że stężenia ozonu przekraczają określony próg. Zauważ, że właściwie to wykonaliśmy dopiero pierwszy krok w tym procesie – wyeksplorowaliśmy nasze dane. Wiemy teraz o nich całkiem sporo. Możemy więc przystąpić do kolejnych, bardziej szczegółowych etapów analizy.

9. Bibliografia

- Armstrong, M., 1998. *Basic linear geostatistics*. Springer – Verlag, Berlin, 156.
- Isaaks, E.H., Srivastava, R.M., 1988. Spatial continuity measures for probabilistic and deterministic geostatistics. *Mathematical Geology*, 20(4), 313–341.
- Isaaks, E.H., Srivastava, R.M., 1989. *An introduction to Applied Geostatistics*. Oxford University Press, New York, Oxford.
- Kokesz, Z., Mucha, J., 1984. Anizotropia parametrów złożowych – cele i metody badania. *Techn. Poszuk. Geol.*, 2, 40–49.
- Meisel, J.E. & Turner, M.G., 1998. Scale detection in real and artificial landscapes using semivariance analysis. *Landscape Ecology*, 13, 347–362.
- Mucha, J., 1994. *Metody geostatystyczne w dokumentowaniu złóż*. Skrypt, Katedra Geologii Kopalnianej, Wyd. AGH, Kraków, 155.
- Namysłowska-Wilczyńska, B., 2006. Geostatystyka. Teoria i zastosowania. Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław, 356.
- Orłóci, L., 1978. *Multivariate analysis in vegetation research*. Springer. 468.
- Parysek, J.J., 1982. *Modele klasyfikacji w geografii*. Wydawnictwo Naukowe Uniwersytetu Adama Mickiewicza, Seria Geografia, 31, Poznań, 223.

Surfer 8. User's Guide, 2002, Golden Software, Inc. Colorado, 640.

Tobler, W.R., 1970. A Continuous Transformation Useful for Districting. *Annals of the New York Academy of Sciences*, 215–220.

Urbański, J., 2011. *GIS w badaniach przyrodniczych*. Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk, 252.

Zawadzki, J., 2011. *Metody geostatystyczne dla kierunków przyrodniczych i technicznych*. Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa, 132.