INFRASTRUKTURA I EKOLOGIA TERENÓW WIEJSKICH
INFRASTRUCTURE AND ECOLOGY OF RURAL AREAS

Nr II/2/2014, POLSKA AKADEMIA NAUK, Oddział w Krakowie, s. 377–388
Komisja Technicznej Infrastruktury Wsi

DOI: http://dx.medra.org/10.14597/infraeco.2014.2.2.027

WYKORZYSTANIE ANALIZY SEMIWARIANCJI W OSZACOWANIU STĘŻENIA SO_2 W POWIETRZU ATMOSFERYCZNYM

Jolanta Kwiatkowska-Malina, Andrzej Szymon Borkowski
Politechnika Warszawska

APPLICATION OF SEMIVARIANCE ANALYSIS FOR ESTIMATING SO_2 CONCENTRATION IN ATMOSPHERIC AIR

Streszczenie

Cechą charakterystyczną zjawisk zachodzących w powietrzu atmosferycznym jest istnienie autokorelacji. Dzięki narzędziom geostatystyki możliwa jest identyfikacja, modelowanie i estymacja stanu skażenia powietrza atmosferycznego. Zjawisko autokorelacji wykazuje, że wyniki pomiarów, które sąsiadują ze sobą w czasie i przestrzeni są do siebie bardziej podobne niż wyniki pomiarów bardziej odległych. Istnienie autokorelacji czasowych może być podstawą do szacowania wartości pomiarów sąsiednich. Do badania autokorelacji czasowej stężenia gazów i aerozoli wykorzystano analizę semiwariancji. Znajomość parametrów semiwariancji pozwala na użycie najlepszej nieobciążonej liniowo metody estymacji jaką jest kriging.

W pracy przedstawiono przykład zastosowania analizy semiwariancji stanu jakościowego powietrza atmosferycznego dla obszaru województwa. Dokonano oceny autokorelacji przestrzennej i czasowej w województwie śląskim w oparciu o podstawowe parametry tj. zakres oddziaływania, próg oraz efekt samorodków. Wyniki analizy semiwariancji przedstawiają w postaci funkcji liniowej zależność między oddalaniem się punktów a stopniem ich podobieństwa. Istnienie autokorelacji przestrzennej daje podstawę do modelowania i estymacji przy użyciu metod stochastycznych.

Słowa kluczowe: semiwariancja, autokorealacja, interpolacja.

Summary

A characteristic feature of the phenomena occurring in the atmospheric air is the existence of autocorrelation. With tools geostatistics is possible to identify modeling and estimation of the state of contamination of the air. The phenomenon of autocorrelation shows that the results of measurements that are adjacent to each other in time and space are more alike than more distant measurements. The existence of autocorrelation time can be the basis for estimating the value of the measurement location. To study the autocorrelation time gas concentrations and aerosols used semivariance analysis. Knowledge of semivariance parameters allows you to use the best uncompressed linear estimation method which is kriging.

The paper presents an example of application semivariance analysis of ambient air quality status for the area of the province. An evaluation of spatial and temporal autocorrelation in the Silesia province based on the basic parameters such as range of impact, and the effect of the threshold nuggets. The analysis results show the semivariance as a function of the linear relationship between distant the points and the degree of their similarity. The existence of spatial autocorrelation provides a basis for modeling and estimation using stochastic methods.

Key words: *semivariance, autocorrelation, interpolation.*

WPROWADZENIE

Geostatystyka to zestaw modeli i narzędzi opracowanych dla analizy statystycznej danych ciągłych. Dane te mogą być mierzone w dowolnym czasie i przestrzeni, ale są one dostępne jedynie w ograniczonej liczbie punktów pomiaru. Ponieważ dane wejściowe są obarczone błędami, modele geostatystyczne są jedynie przybliżeniem rzeczywistości (Namysłowska-Wilczyńska, 2006).

Jednym z ważniejszych czynników determinującym warunki klimatyczne jest stan powietrza atmosferycznego. Regularne pomiary dotyczące stanu jakościowego powietrza atmosferycznego w Polsce prowadzone są między innymi przez Wojewódzkie Inspektoraty Ochrony Środowiska (WIOŚ). W ramach systemu oceny jakości powietrza WIOŚ gromadzą, weryfikują, archiwizują i udostępniają bazy danych z wynikami pomiarów niektórych substancji w powietrzu. W województwie śląskim WIOŚ obsługuje 16 automatycznych stacji monitoringu powietrza. Opracowywane przez WIOŚ raporty zawierają wyniki pomiarów

stanu jakościowego powietrza, które są przedstawiane w cyklach dobowych, miesięcznych i rocznych.

Zawartość w powietrzu zanieczyszczeń tj. jednego z badanych parametrów np. dwutlenku azotu zmierzona o godzinie 15 jest bardziej zbliżona do wyniku pomiaru, który wykonano godzinę później niż do tego, wyniku pomiaru, który wykonano 5 godzin wcześniej. Dlatego wyników pomiarów nie można traktować jako zdarzeń losowych. Zależność przestrzenna (autokorelacja) istnieje wtedy, gdy badane zjawiska w jednej przestrzeni zwiększają lub zmniejszają prawdopodobieństwo wystąpienia tego zjawiska w przestrzeni sąsiadującej (Zawadzki, 2011). Celem badań było oszacowanie stężenia SO_2 w powietrzu atmosferycznym na obszarze województwa śląskiego na podstawie ograniczonych danych udostępnianych przez WIOŚ pozyskiwanych w stacjach automatycznego pomiaru jakości powietrza. W badaniach wykorzystano analizę semiwariancji, która pozwala na ustalenie zależności czasowej i/lub przestrzennej oraz interakcji pomiędzy wartościami badanych parametrów.

MATERIAŁY I METODY BADAŃ

W niniejszej pracy wykorzystano wyniki pomiarów (dane) wykonane i zgromadzone w ramach Państwowego Monitoringu Środowiska przez Wojewódzki Inspektorat Ochrony Środowiska w Katowicach. Przeanalizowano zawartość dwutlenku siarki w powietrzu atmosferycznym. W ostatnich latach zanotowano w niektórych stacjach pomiarowych województwa śląskiego przekroczenia dopuszczalnych poziomów stężeń tego związku (Rozporządzenie Ministra Środowiska z dnia 24 sierpnia 2012 r. w sprawie poziomów niektórych substancji w powietrzu (Dz. U. z 2012r. nr 0 poz. 1031). Do analizy semiwariancji użyto oprogramowanie typu GIS z modułem geostatystycznym.

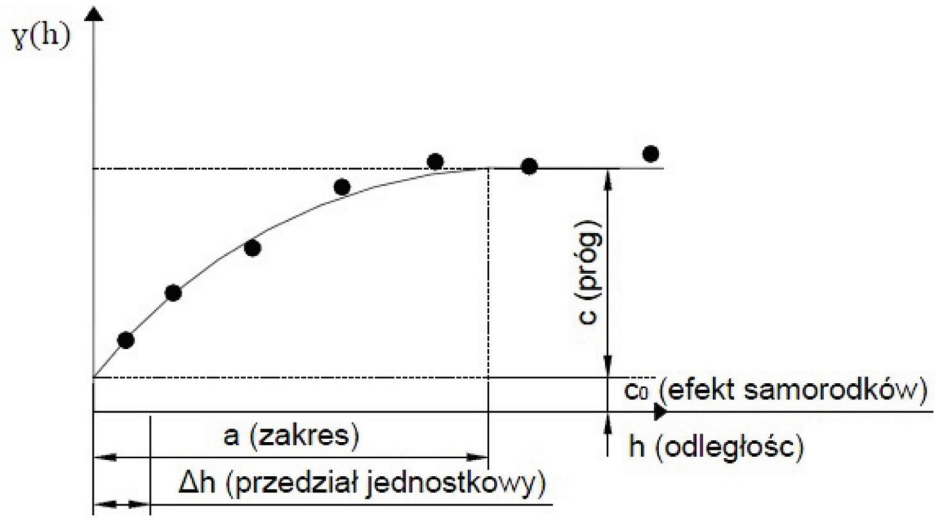
Analiza semiwariancji (Goovaerts, 1997) bada związki czasowe oraz przestrzenne i jest wyrażana jako połowa średniego kwadratu odchyleń między wartościami parametru (1):

$$\gamma = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} (z(u_\alpha - z(u_\alpha + h))^2 \quad (1)$$

gdzie:
$N(h)$ – liczba par danych dla danego odstępu czasu h między nimi,
$z(u\alpha)$ dla $\alpha = 1, 2, ..., n$ oznaczają zbiór n pomiarów danego parametru,
$u\alpha$ – wektor terminów wykonywanych pomiarów.

Model wariogramu stanowi charakterystykę ciągłości przestrzennej występującym w badanym zbiorze danych (Zawadzki, 2011). Modele te charakteryzuje się na podstawie trzech parametrów (Fortin, Dale, 2005):

- zakresu oddziaływania (range), który określa odległość, powyżej której nie istnieją korelacje przestrzenne; dystans, powyżej którego wartość semiwariancji jest stała,
- efektu samorodków (nugget effect), na który składa się wartość błędu pomiarowego,
- progu (sill), to wartość progowa funkcji odpowiadająca zasięgowi oddziaływania wyznaczonego przez range; określa wartość nasycenia semiwariogramu.



Źródło: opracowanie własne
Source: own study

Rysunek 1. Wykres zależności semiwariancji od odległości między punktami pomiarowymi h
Figure 1. Semivariance graph of the distance between the measuring points h

Wykres jest dopasowany do danych w taki sposób, aby przybliżyć w postaci liniowej zależność pomiędzy oddalaniem się punktów (h), a stopniem ich podobieństwa (γ) (rys. 1.). Pierwszym krokiem w analizie danych statystycznych powinno być badanie dwóch cech danych: zależności i stacjonarności. Stacjonarność oznacza, że właściwości statystyczne nie zależą od dokładnej lokalizacji (Brenning, 2012). Dlatego średnia (wartość oczekiwana) zmiennej w jednym miejscu jest równa średniej z dowolnej, innej lokalizacji, a korelacji pomiędzy dowolnymi dwoma miejscami zależy tylko od wektora, który je oddziela a nie od położenia.

CHARAKTERYSTKA OBSZARU BADAŃ

W ramach Państwowego Monitoringu Środowiska w województwie śląskim znajduje się 16 stacji automatycznego pomiaru jakości powietrza. W stacjach rejestrowane są w sposób ciągły wartości pomiarów stężeń substancji w powietrzu atmosferycznym. Stacje pomiarowe monitoringu środowiska w województwie śląskim są rozmieszczone nierównomiernie (rys. 2). Tworzą one sieć prostą losową (simple random). Właściwa lokalizacja stacji powinna dostarczyć jak największej liczby informacji dotyczącej rozkładu przestrzennego przy jak najmniejszych kosztach i innych nakładach. Pojawia się zatem pytanie – czy liczba 16 stacji automatycznego pomiaru na obszarze województwa śląskiego jest wystarczająca?

Źródło: opracowanie własne
Source: own study

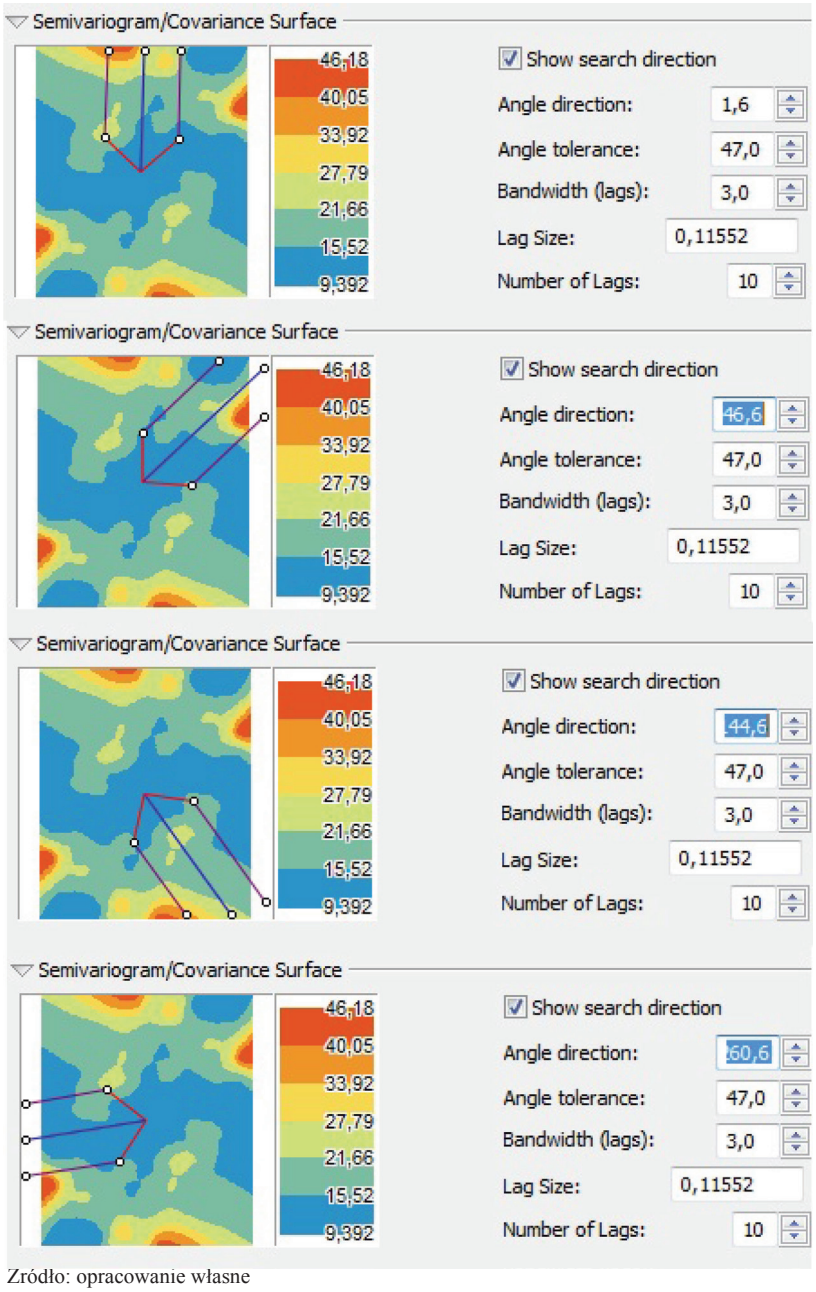
Rysunek 2. Stacje pomiaru automatycznego
Figure 2. Automatic measurement stations

W 2012 roku w trzech stacjach pomiarowych na terenie województwa śląskiego zanotowano przekroczenie dopuszczalnego poziomu stężeń SO_2 w powietrzu (20 µg/m^3) (Dz. U. z 2012r. nr 0 poz. 1031). Możliwe jest zatem, że na obszarach województwa śląskiego których nie obejmuje zasięg pomiaru punktowego (stacje pomiarowe) przekraczane są dopuszczalne godzinowe lub dobowe poziomy stężeń SO_2 w powietrzu . Do badań korelacji przestrzennej (semiwariancji) oraz do interpolacji przy użyciu metody krigingu zwykłego wykorzystano uśrednione roczne wyniki pomiarów (tab. 1).

Tabela 1. Wartości zmiennej zregionalizowanej stężeń SO_2 2012 w roku oraz ich położenie w województwie śląskim
Table 1. Regionalised variable values for concentrations of SO_2 in 2012, and their position in the voivodship

Nazwa stacji	Z (dla SO_2)	X	Y
Bielsko Biala	17,0	19,03690	49,80880
Cieszyn	14,0	18,65100	49,75580
Częsochowa ul. Armii Karajowej	17,0	19,12100	50,83830
Częstochowa ul. Baczyńskiego	12,0	19,10290	50,84780
Dąbrowa Górnicza	17,0	19,28280	50,35290
Gliwice	16,0	18,65370	50,29480
Katowice A4	15,0	19,02100	50,23010
Katowice ul. Kossutha	17,0	18,97350	50,25040
Rybnik	**23,0**	18,54160	50,11110
Sosnowiec	19,0	19,19170	50,27670
Tychy	19,0	19,00700	50,11960
Ustron	9,0	18,82630	49,71390
Wodzislaw	20,0	18,45870	50,00380
Zabrze	**21,0**	18,78890	50,30980
Złoty Potok	10,0	19,42590	50,70220
Żory	16,0	18,69480	50,03470
Żywiec	**24,0**	19,21350	49,69260

Źródło: opracowanie własne na podstawie danych Wojewódzkiego Inspektoratu Ochrony Środowiska
Source: own study on the basis of Voivodeship Inspection for Environmental Protection data



Zródło: opracowanie własne
Source: own study

Rysunek 3. Badanie izotropowości zjawiska
Figure 3. Examination of the isotropic phenomenon

WYNIKI

Charakterystykę statystyczną analizowanych pomiarów SO_2 w roku 2012 przedstawiono w tabeli 2. Wartości odchyleń standardowych i wariancja wskazują na duże rozproszenie danych statystycznych od średniej arytmetycznej (Cieciura, Zacharski, 2011).

Tabela 2. Statystyki opisowe analizowanych pomiarów SO_2
Table 2. Descriptive statistics of the analyzed measurements of SO_2

Substancja	Min.	Max.	X	Me	SD	Var	Sk	Ku
SO_2	9,0	24,0	17,0	17,0	3,99	16,9	– 0,04	-0,12

Źródło: opracowanie własne
Source: own study

Oznaczenia:
Min. – minimum, Max. – maksimum, X – średnia arytmetyczna, Me – mediana, SD – odchylenie standardowe, Var – wariancja, Sk – skośność, Ku – kurtoza.

Biorąc pod uwagę wszystkie dotychczas wymienione aspekty: rodzaj sieci oraz wartości charakterystyk podstawowych, wykonano analizę semiwariancji dla wartości stężeń SO_2 dla potrzeb interpolacji metodą krigingu. Uważa się, że kriging jest najlepszym liniowym nieobciążonym estymatorem. Nieznane wartości są szacowane przez kriging ważonymi kombinacjami dostępnych danych. Do wyliczenia wag wykorzystuje się semiwariogram. W krigingu zakłada się istnienie autokorelacji przestrzennej. Zgodnie z jego założeniami, zróżnicowanie przestrzenne badanej zmiennej dzieli się na trzy główne składowe (Burrough, McDonnell, 1998):

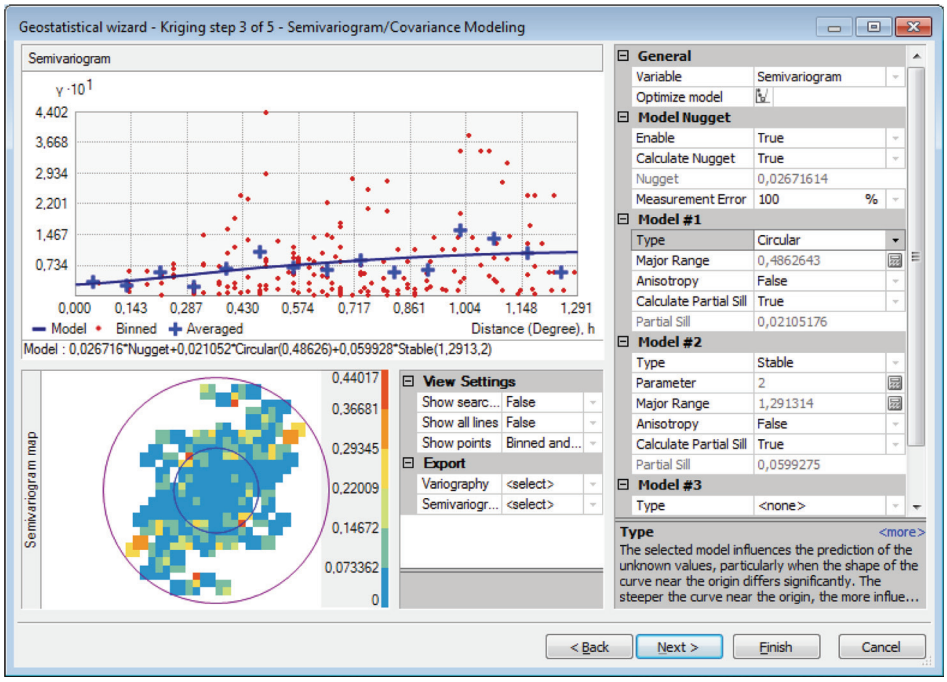
- ogólny trend zmienności danych, o stałej średniej,
- lokalną zmienność losową skorelowaną przestrzennie,
- nieskorelowany przestrzennie „szum" (wynikający np. z błędów pomiarowych).

Kriging zwykły (ordinary) zakłada, że badane zjawisko jest izotropowe. To znaczy, że funkcja autokorelacji przestrzennej jest taka sama we wszystkich kierunkach.

Badanie zmienności cechy w zależności od kierunku wykazało, że zjawisko zanieczyszczenia powietrza atmosferycznego SO_2 ma charakter izotropowy (rys.3). Na wykresach nie widać zwiększonej korelacji niezależnie od zmiany kierunku złączeń.

Następnie wygenerowano semiwariogram dla stężeń SO_2 w powietrzu atmosferycznym. Nieregularny kształt wariogramu (tzw. kształt „piły") można

tłumaczyć występowaniem na przemian niskich i wysokich wartości wariogramu (rys.4). Spowodowane jest to niewystarczającą liczbą punktów pomiarowych.

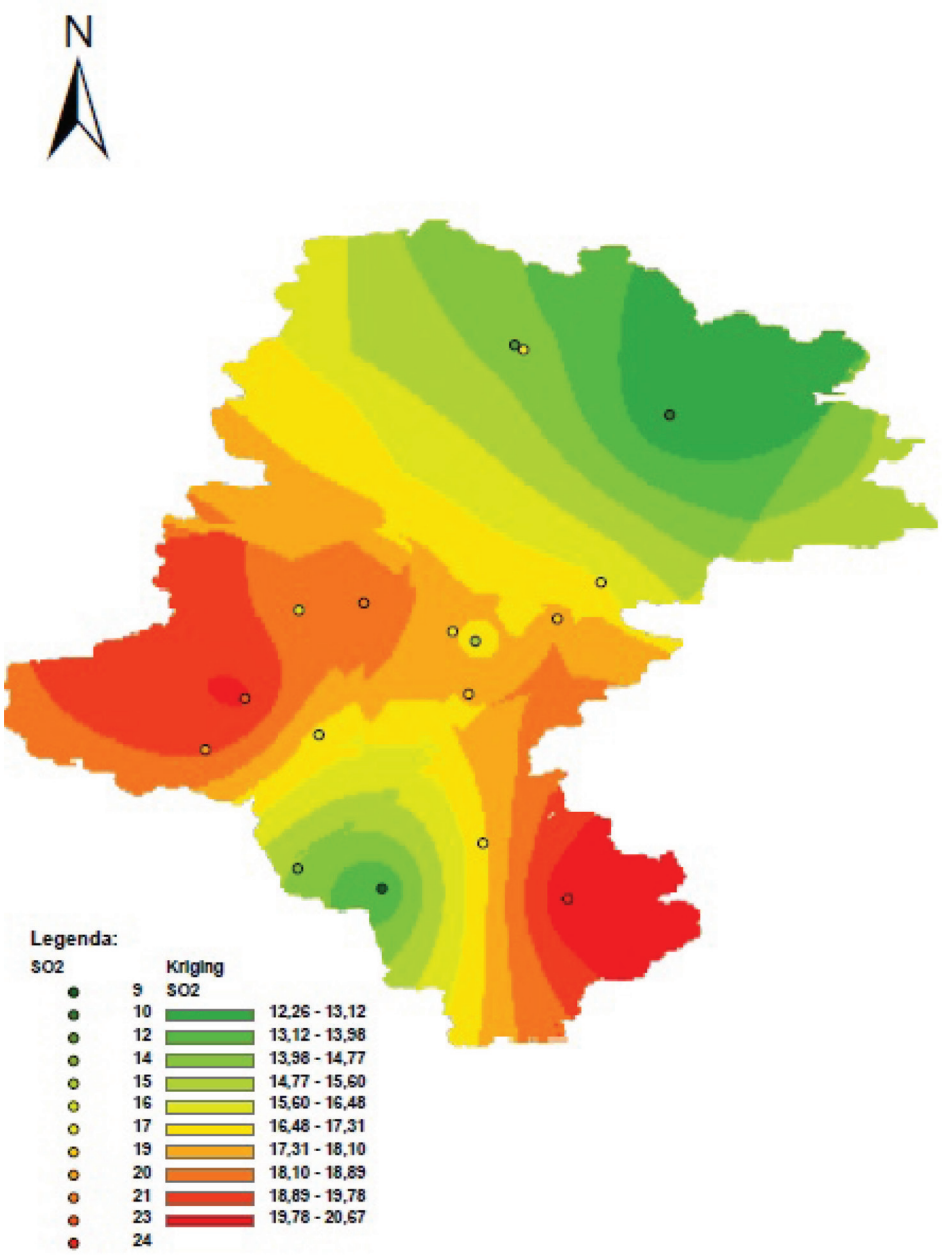


Źródło: opracowanie własne
Source: own study

Rysunek 4. Model semiwariogramu (typ kołowy)
Figure 4 Semivariogram model (type circular)

Otrzymana funkcja zbliżona jest do liniowej, prawie równoległej do osi liniowej wskazuje na wysoką losowość tj. niską korelację przestrzenną (dość wysoki efekt samorodków). Efekt samorodków spowodowany jest błędami pomiarowymi oraz skończoną liczbą obserwacji. Dlatego nie można jednoznacznie stwierdzić jak zachowuje się zjawisko zanieczyszczenia powietrza atmosferycznego SO_2 przy odległościach mniejszych niż minimalna odległość między punktami pomiarowymi na obszarze województwa śląskiego.

Na podstawie otrzymanego semiwariogramu wygenerowano mapę metodą krigingu zwykłego (rys. 5).



Źródło: badania własne
Source: own study

Rysunek 5. Stacje pomiaru automatycznego.
Figure 5. Automatic measurement stations.

PODSUMOWANIE

Wyniki analizy semiwariancji dla pomiarów otrzymanych ze stacji automatycznych pokazują, że jest to niewystarczający materiał do estymacji i generowania map rozkładu stężeń dla gazów i aerozoli w powietrzu atmosferycznym. Otrzymany efekt interpolacji metodą krigingu jest bardzo podobny do jednej z deterministycznych form estymacji tj. metody najbliższego sąsiada.

Niska autokorelacja przestrzenna oraz wysoki rozrzut danych (wariancja) nie pozwalają na prawidłowy dobór wag dla potrzeb interpolacji metodą krigingu. Pomocne w tym przypadku mogą być dodatkowe pomiary np. manualne (200 stanowisk pomiarowych w województwie śląskim) wykorzystane, jako zmienna pomocnicza w interpolacji metodą kokrigingu. Musi jednak być spełnione założenie, że istniejące dane są skorelowane z badaną zmienną i dzięki temu mogą dostarczyć dodatkowej informacji do prowadzonej estymacji w nieznanych miejscach.

LITERATURA

Burrough P., McDonnell R., (1998) Principles of Geographical Information Systems, New York: Oxford University Press Inc.

Brenning, A. (2012). Geostatistics without Stationarity Assumptions within Geographical Information Systems. Freiberg Online Geoscience.

Cieciura, M., Zacharski, J. (2011). Podstawy probalistyki z przykładami zastosowań w informatyce. Część II. Statystyka opisowa, Warszawa, s.17.

Fortin, M.-J., Dale, M.R.T. (2005). Spatial Analysis. A Guide for ecologists. Cambridge University Press.

Goovaerts, P. (1997). Geostatistics for Natural Resources Evaluation. Oxford University Press, 1-483.

Namysłowska-Wilczyńska, B. (2006). Geostatystyka. Teoria i zastosowania. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej, 1-356.

Rozporządzenie Ministra Środowiska z dnia 24 sierpnia 2012 r. w sprawie poziomów niektórych substancji w powietrzu (Dz. U. z 2012r. nr 0 poz. 1031)

www.katowice.pios.gov.pl/

Zawadzki, J. (2011). Metody geostatystyczne dla kierunków przyrodniczych i technicznych. Warszawa: Oficyna Wydawnicza Politechniki Warszawskiej, 25-28.

Dr hab. Jolanta Kwiatkowska – Malina Prof. PW
E-mail: j.kwiatkowska@gik.pw.edu.pl
Mgr inż. Andrzej Szymon Borkowski
E-mail: a.borkowski@gik.pw.edu.pl
Wydział Geodezji i Kartografii,
Politechnika Warszawska
Pl. Politechniki 1
00-661 Warszawa,