

ZASTOSOWANIE PAKIETU STATYSTYCZNEGO BOOT W OBLICZENIACH ROZPRZESTRZENIANIA SIĘ ZANIECZYSZCZEŃ W POWIETRZU

Przemysław SZCZYGŁOWSKI, Marian MAZUR
Akademia Górniczo-Hutnicza, Wydział Geodezji Górniczej i Inżynierii Środowiska,
Katedra Kształtowania i Ochrony Środowiska, Al. Mickiewicza 30, 30-059 Kraków
szczyklo@agh.edu.pl

STRESZCZENIE

Zagadnienie modelowania rozprzestrzeniania się zanieczyszczeń wymaga podjęcia szeregu decyzji mających na celu właściwy dobór opcji modelu związanych z meteorologią czy współczynnikami dyspersji. Każda ich zmiana powoduje różnicę w wynikach obliczeń stanu zanieczyszczeń. Zastosowane w pracy nowe wskaźniki statystyczne pozwalają na dokładniejszą ocenę wpływu zmiany poszczególnych parametrów na wyniki modelowania. Szczególnie przydatnymi, stosunkowo nowymi wskaźnikami statystycznymi, mogą okazać się ujemne i dodatnie komponenty błędu systematycznego (FB) bądź geometrycznego błędu średniego (MG), które pozwalają na ocenę przeszacowania lub niedoszacowania wyników obliczeń w przypadku obliczeń rozkładu zanieczyszczeń na określonym obszarze.

1. Podstawowe wskaźniki statystyczne

Metodyka porównywania poszczególnych opcji modeli oparta została na wskaźnikach statystycznych opracowanych dla modeli rozprzestrzeniania się zanieczyszczeń. Wskaźniki te zebrano w pakiecie statystycznym BOOT Statistical Model Evaluation Software Package, Version 2.0 [1]. Podstawowe wskaźniki statystyczne wykorzystywane w porównaniach zostały zarekomendowane przez EPA (Environmental Protection Agency), a następnie zebrane w aplikacji BOOT [2, 3]. Podstawowe wskaźniki statystyczne to: FB – błąd systematyczny (Fractional Bias), MG – geometryczny błąd średni (Geometric Mean Bias), NMSE – względne odchylenie standardowe (Normalized Mean Square Error), VG – geometryczna średnia wariancja (Geometric Variance), R – współczynnik korelacji (Correlation Coefficient), FAC2 – wskaźnik charakteryzujący jakość modelu. Poza podstawowymi sześcioma podstawowymi wskaźnikami zdefiniowanymi powyżej wyznaczono również dodatkowe wskaźniki takie jak FB_{FN} – ujemny komponent FB (false-negative), FB_{FP} – dodatni komponent FB (false-positive), ujemny (MG_{FN}) i dodatni (MG_{FP}) komponent geometrycznego błędu średniego MG.

FB_{FN} – ujemny komponent FB (false-negative), rozpatruje tylko te pary, (C_o , C_p), w których $C_p < C_o$:

$$FB_{FN} = \frac{0.5 \cdot \sum_i [|C_{O_i} - C_{P_i}| + (C_{O_i} - C_{P_i})]}{0.5 \cdot \sum_i (C_{O_i} - C_{P_i})} \quad (1)$$

FB_{FP} – dodatni komponent FB (false-positive), rozpatruje tylko te pary (C_o , C_p), w których $C_p > C_o$:

$$FB_{FP} = \frac{0.5 \cdot \sum_i [|C_{O_i} - C_{P_i}| + (C_{P_i} - C_{O_i})]}{0.5 \cdot \sum_i (C_{O_i} - C_{P_i})} \quad (2)$$

gdzie: C_p – stężenie wyznaczone przez model,
 C_o – stężenie rzeczywiste.

W podobny sposób można wyznaczyć ujemny (MG_{FN}) i dodatni (MG_{FP}) komponent geometrycznego błędu średniego MG.

$$MG_{FN} = \exp \left[\frac{1}{2N} \sum_i \left[|\ln C_{O_i} - \ln C_{P_i}| + (\ln C_{O_i} - \ln C_{P_i}) \right] \right] \quad (3)$$

$$MG_{FP} = \exp \left[\frac{1}{2N} \sum_i \left[|\ln C_{O_i} - \ln C_{P_i}| + (\ln C_{P_i} - \ln C_{O_i}) \right] \right] \quad (4)$$

2. Opcje obliczeń

Informację o rzeźbie i użytkowaniu terenu pochodzącą z U.S. Geological Survey [4]. Obliczenia wykonano w siatce 40x40 km z rozdzielczością 1x1 km. Do obliczeń wykorzystano dane meteorologiczne dwojakiego rodzaju: pochodzące z naziemnych stacji meteorologicznych: Katowice (12560), Kraków (12566), Bielsko-Biała (11111) oraz 4 stacji aerologicznych: Legionów (12374), Wrocław (12425), Praga (11520) i Poprad (11952).

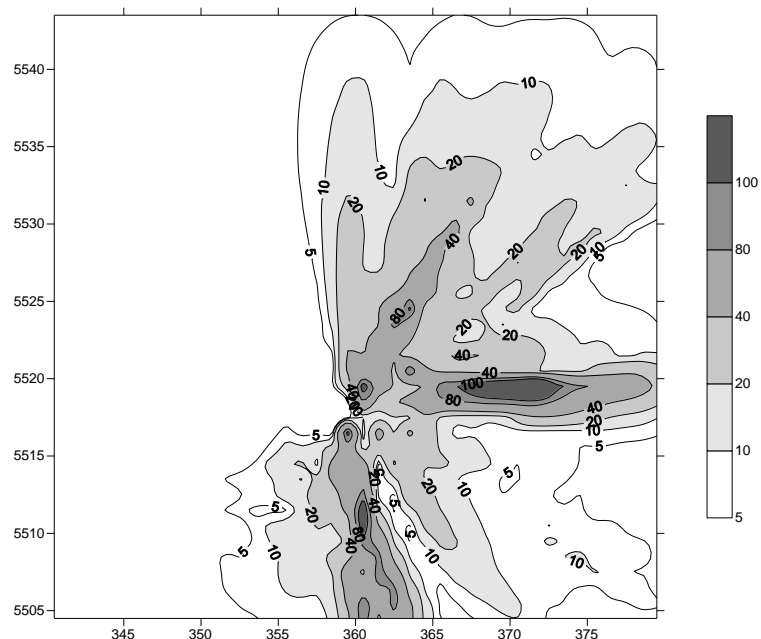
Zródłem emisji był komin elektrociepłowni zlokalizowany w Bielsku-Białej, w punkcie o współrzędnych (360.0, 5518.0), wysokość podstawy 420 m n.p.m. Emisja SO_2 wynosiła 100 g/s. Wysokość komina wynosiła 100 m (wysokość podstawy 420 m n.p.m), a jego średnica na wylocie 5,0 m. Prędkość wylotowa spalin wynosiła 15,0 m/s przy temperaturze gazów 400 K [5].

Przyjęto następujące opcje odpowiadające ilości danych meteorologicznych użytych do obliczeń rozkładu stężeń zanieczyszczeń:

- 1) 3 stacje naziemne (Bielsko-Biała, Kraków, Katowice), 4 aerologiczne (Wrocław, Legionów, Praga, Poprad) – 4up3surf,
- 2) 2 stacje naziemne (Kraków, Katowice), 4 aerologiczne (Wrocław, Legionów, Praga, Poprad) – 4up2surf,
- 3) 1 stacja naziemna (Bielsko-Biała), 4 aerologiczne (Wrocław, Legionów, Praga, Poprad) – 4up1surf,
- 4) 3 stacje naziemne (Bielsko-Biała, Kraków, Katowice), 2 aerologiczne (Wrocław, Poprad) – 2pow3surf,
- 5) 3 stacje naziemne (Bielsko-Biała, Kraków, Katowice), 2 aerologiczne (Legionów, Praga) – 2prl3surf,
- 6) 3 stacje naziemne (Bielsko-Biała, Kraków, Katowice), 1 aerologiczna (Legionów) – leg3surf,
- 7) 3 stacje naziemne (Bielsko-Biała, Kraków, Katowice), 1 aerologiczna (Wrocław) – wroc3surf,
- 8) 3 stacje naziemne (Bielsko-Biała, Kraków, Katowice), 1 aerologiczna (Praga) – pra3surf,
- 9) 3 stacje naziemne (Bielsko-Biała, Kraków, Katowice), 1 aerologiczna (Poprad) – pop3surf.

3. Wyniki obliczeń

Dla każdej z opisanych opcji, wykonano mapy rozkładów 1-godzinnych stężeń maksymalnych. Przykładową mapę rozkładu zanieczyszczeń przedstawiono na rys. 1.



Rys. 1. Rozkład 1-godzinnych stężeń maksymalnych (dane meteorologiczne - 4up 3surf, emitor punktowy h = 100 m)

W tabeli 1 przedstawiono zestawienie wskaźników statystycznych wyznaczonych z rozkładu 1-godzinnych stężeń zanieczyszczeń SO₂ w funkcji ilości dostępnych danych meteorologicznych.

Tabela 1. Zestawienie wskaźników statystycznych wyznaczonych z rozkładu 1-godzinnych stężeń zanieczyszczeń SO₂ w zależności od ilości danych meteorologicznych (emitor punktowy h = 100 m)

Wskaźnik	Opcje obliczeniowe przyjęte ze względu na rodzaj danych meteorologicznych wykorzystanych w obliczeniach stanu zanieczyszczenia powietrza								
	4up 3surf	4up 2surf	4up 1surf	2pow 3surf	2prl 3surf	Leg 3surf	wroc 3surf	pra 3surf	pop 3surf
FB	0	0.24	-0.01	-0.46	-0.18	0.11	-0.69	0.09	-0.55
MG	1	1.22	1.01	3.87	0.51	0.62	0.35	0.43	3.50
NMSE	0	1.66	0.02	3.38	5.94	4.85	12.35	3.70	6.25
VG	1	1.34	1.01	2.43E+04	4.24E+05	2.93E+04	106.00	2.70E+04	3.36E+04
R	1	0.73	1.00	0.61	0.18	0.20	0.28	0.16	0.42
FAC2	1	0.85	1.00	0.43	0.27	0.33	0.27	0.33	0.36
Średnia	11.5	9.0	11.5	18.2	13.7	10.3	23.6	10.4	20.2
Odch. s.	18.7	14.5	19.3	32.3	27.7	19.1	58.8	13.0	40.8
Max	178	147	187	382	300	295	784	145	528
FB _{FN}	0	0.32	0.01	0.14	0.47	0.58	0.23	0.58	0.16
FB _{FP}	0	0.08	0.02	0.60	0.65	0.47	0.92	0.49	0.71
MOE _{FN}	1	0.72	0.99	0.82	0.49	0.45	0.66	0.45	0.78
MOE _{FP}	1	0.91	0.98	0.52	0.41	0.50	0.32	0.49	0.44
MG _{FN}	1	1.29	1.02	5.55	2.61	2.40	1.40	2.11	5.58
MG _{FP}	1	1.05	1.01	1.43	5.14	3.87	4.01	4.93	1.60

Największą zgodność z przyjętym wzorcem dla błędu statystycznego odnotowano dla opcji 4up1surf (FB = -0,01), została ona dodatkowo potwierdzona niskimi wartościami jej ujemnego i dodatniego komponentu ($FB_{FN} = 0,01$, $FB_{FP} = 0,02$). Dobry wskaźnik FB charakteryzował również opcję 4up2surf (FB = 0,24, $FB_{FN} = 0,32$, $FB_{FP} = 0,08$). W przypadku pozostałych opcji stosunkowo niski współczynnik FB nie został potwierdzony ujemnym i dodatnim komponentem (np. dla opcji leg3surf FB = 0,11 ale $FB_{FN} = 0,58$, $FB_{FP} = 0,47$). Dla współczynnika korelacji R wartości powyżej 0,7 osiągnęły jedynie opcje: 4up1surf (R = 1,00), 2pow3surf (R = 0,73). Również tylko te opcje charakteryzowały się wskaźnikiem FAC2 powyżej 0,5 i tak, dla 4up1surf FAC2 wyniósł 1,00, a 4up2surf FAC2 = 0,85.

Już pobieżna analiza rozkładów stężeń zanieczyszczeń pokazała, że bardzo istotnym elementem właściwego określenia poziomu emisji na danych obszarze jest wykorzystanie jak największej ilości danych meteorologicznych. Przy czym brak danych meteorologicznych ze stacji naziemnych, choć powoduje zmianę rozkładu stężeń zanieczyszczeń jest jeszcze akceptowalny, natomiast brak danych o pionowym profilu atmosfery może znacznie wpłynąć na wyniki obliczeń. Miało to miejsce w przypadku opcji od 6 do 9, w których wykorzystywano informacje o pionowym profilu stanu atmosfery pochodzącym tylko z jednej stacji aerologicznej. Zarówno kształty rozkładu zanieczyszczeń, jak i wskaźniki statystyczne były zupełnie różne od wzorca.

W wyniku przeprowadzonych obliczeń i analiz można stwierdzić, że zastosowanie jedynie podstawowych wskaźników statystycznych nie daje pełnego obrazu wyników modelowania. Przykładowa niska wartość błędy systematycznego FB może spowodować niesłuszną ocenę badacza, że wybrany model w dobry sposób przybliży rzeczywisty stan, jakości powietrza. Jednak dodatkowe wyznaczenie jego ujemnego i dodatniego komponentu może doprowadzić do całkiem odmiennych wniosków. Może okazać się, że w wybranym obszarze część wyników jest znacznie zawyżona (duża wartość ujemnego komponentu FB), bądź zaniżona (duża wartość dodatniego komponentu FB).

Praca została wykonana w ramach badań statutowych AGH nr 11.11.150.171.

Literatura

1. Joseph C. Chang, Steven R. Hanna: Technical Descriptions and User's Guide for the BOOT Statistical Model Evaluation Software Package, Version 2.0, 2005.
2. ASTM, 2000: Standard guide for statistical evaluation of atmospheric dispersion model performance. American Society for Testing and Materials, Designation D6589-00. ASTM, 100 Barr Harbor Drive, West Conshohocken, PA 19428-2959.
3. Cox, W.M., Tikvart J.A.: A statistical procedure for determining the best performing air quality simulation model. Atmospheric Environment, 1990, vol. 24A, No 9, 2387-2395.
4. <http://edc.usgs.gov/>
5. Szczygłowski P.: Ocena przydatności wybranych modeli gaussowskich w obliczeniach stanu zanieczyszczenia powietrza. Praca doktorska. Wydział Geodezji Górniczej i Inżynierii Środowiska AGH, Kraków 2007.