

Anna Janiga-Ćmiel

ANALIZA PRZESTRZENNA KOSZTU LIKWIDACJI SZKÓD W OCHRONIE ŚRODOWISKA

Wprowadzenie

Jednym z najważniejszych elementów zarządzania na różnych jego szczeblach jest podejmowanie decyzji. Najczęściej podjęcie decyzji ma miejsce w sytuacjach, gdy nieznane są jej skutki. Skutki te są zazwyczaj wymierne i dostrzegamy je głównie w długich okresach czasu. Podejmowanie decyzji w określonych warunkach wymaga od decydenta, aby mógł przewidzieć konsekwencje własnych decyzji. Wymierne skutki podejmowania decyzji można z dużą precyzją oszacować przed ich podjęciem. Służą do tego metody ilościowe, które odgrywają ważną rolę w wielu dziedzinach nauk, począwszy od techniki, ekonomii, a skończywszy na psychologii itp. Warunkiem prawidłowego określenia skutków podejmowania decyzji jest w miarę możliwości bezbłędnie wykonana analiza zjawiska. Wymaga się przy tym określenia stosownych metod oraz ustalenia ich kolejności na podstawie celów, którym są podporządkowane.

Każdą analizę rozpoczynamy od obserwacji statystycznej, w wyniku której ustala się możliwie obszerny zbiór danych statystycznych. Po dokonaniu oceny jakości pozyskanych danych przystępujemy do zbudowania modelu, a po jego gruntownej weryfikacji dokonujemy opisu stanu aktualnego zjawiska lub przystępujemy do oszacowania prognoz na przyszłość.

Celem opracowania jest ocena najistotniejszych czynników kształtujących koszt likwidacji szkód i ocena ich znaczenia dla problemu likwidacji szkód w przyszłości. W tym celu zostanie oszacowany odpowiedni model ekonometryczny, przy czym zmienne objaśniające zostaną wyznaczone metodą składowych głównych i dla sprawdzenia słuszności – metodą Hellwiga. Na-

stępnie zostanie wyznaczony syntetyczny miernik rozwoju pozwalający ocenić przydatność zmiennych objaśniających w przyszłości (syntetyczny miernik rozwoju określa poziom zmienności zjawiska, którego przyszły rozwój będzie dla użytkowników systemu zadowalający).

W każdym systemie zarządzania podejmuje się wiele różnych decyzji, z których każda ma za zadanie doprowadzić do realizacji postawionych celów. Syntetyczny wskaźnik rozwoju pozwala określić zakres zjawiska, który może nadal być kształtowany jak do tej pory, oraz zakres, w którym wymaga się zmian. Stąd jego znaczenie przy podejmowaniu decyzji zarządzania strategicznego jest ogromne.

1. Zakres analizy

W niniejszym opracowaniu przedstawiono plan czynności analizy; jego poszczególne etapy zrealizowano na podstawie przykładu analizy kosztów likwidacji szkód w ochronie środowiska.

Analizie poddano następujące zmienne:

Y – koszty likwidacji szkód (ochrona gruntów, rzek, powietrza) jako zmienną objaśnianą,

X₁ – grunty zdewastowane w tys. ha,

X₂ – ścieki przemysłowe nieoczyszczone w hm³,

X₃ – emisja zanieczyszczeń do atmosfery w tys. ton,

X₄ – odpady komunalne w ciągu roku w mln ton.

Tabela 1

Dane empiryczne

Lp.	Rok	X ₁	X ₂	X ₃	X ₄	Y
1	2	3	4	5	6	7
1	1987	123,4	1634,2	4352	145,5	690,5
2	1988	110,3	1601,20	4523	168,7	688,2
3	1989	100,2	1567,30	4023	156,7	645,3
4	1990	95,3	1421,40	3754	150,2	533,2
5	1991	93,7	1342,60	3210	143,9	532,4
6	1992	90,3	1432,00	3045	135,6	512,2
7	1993	84,3	1123,40	2867	133,4	496,5
8	1994	80,4	750,40	2576	126,8	495,2
9	1995	76,4	800,20	2412	125,4	486,4

cd. tabeli 1

1	2	3	4	5	6	7
10	1996	73,2	700,20	2376	123,3	485,1
11	1997	72,2	610,80	2368	122,7	455,2
12	1998	75,6	520,30	2181	124,5	452,1
13	1999	74,2	424,20	1897	133,1	444,4
14	2000	72,8	376,40	1719	126,3	432,1
15	2001	71,5	301,30	1511	125,5	412,5

Źródło: Rocznik Statystyczny (1986-2000).

1.1. Badanie zmienności przyjętych zmiennych

Przedstawione dane poddano wstępnej analizie opisowej. Wyniki tej analizy przedstawiono w tab. 2.

Tabela 2

Zestawienie wyników analizy opisowej

	X_1	X_2	X_3	X_4	Y
Σ	1293,8	14605,9	42814	2041,6	7761,3
\bar{X}	86,25	973,73	2854,27	136,11	517,42
S_x^2	247,63	239587,40	897677,92	195,64	7905,16
S_x	15,74	489,48	947,46	13,99	88,91
V_x	18,24%	50,27%	33,19%	10,28%	17,18%

Pierwszy wiersz tab. 2 jest sumą wartości, drugi zawiera średnie wartości badanych zmiennych:

$$\bar{X} = \frac{1}{n} \sum X_i$$

trzeci przedstawia wariancję:

$$S_x^2 = \frac{1}{n} \sum (X_i - \bar{X})^2$$

natomiast czwarty ukazuje odchylenie standardowe, a wszystko to wyznaczono w celu oszacowania współczynnika zmienności V_x w piątym wierszu:

$$V_x = \frac{S_x}{\bar{X}} \cdot 100\%$$

Każdy spośród wyliczonych współczynników zmienności przekracza progową wartość poziomu zmienności wynoszącą 10%.

$$V_x > 10\% \quad \text{i} \quad V_y > 10\%$$

Rozpatrywane zmienne objaśniające wykazują wystarczającą zmienność, aby na ich podstawie można było przedstawić model ekonometryczny i wykorzystać je do wyjaśniania zmienności zmiennej objaśnianej przez model.

2. Wyznaczenie macierzy współczynników korelacji

Wyznaczamy macierz korelacji R między zmiennymi objaśniającymi oraz wektor współczynników korelacji R_0 między zmienną objaśnianą Y a zmiennymi objaśniającymi.

$$R = \begin{bmatrix} 1 & 0,90285 & 0,94123 & 0,83483 \\ 0,90285 & 1 & 0,95681 & 0,83084 \\ 0,94123 & 0,95681 & 1 & 0,88335 \\ 0,83483 & 0,83084 & 0,88335 & 1 \end{bmatrix}$$

$$R_0 = \begin{bmatrix} 0,94944 \\ 0,88215 \\ 0,95750 \\ 0,85799 \end{bmatrix}$$

3. Budowa macierzy regularnej Q w celu wykrycia efektu katalizy zmiennych

$$Q = [q_{ij}], \quad i, j \in 1, \dots, k$$

gdzie k to liczba wszystkich rozpatrywanych zmiennych objaśnianych i objaśniających.

$$q_{ij} = \min \left[\frac{r_i}{r_j}, \frac{r_j}{r_i} \right] \quad (1)$$

$$r_i = r_{y,x_i}, \quad i = 1, \dots, k$$

Macierz regularną Q rozpatrujemy jedynie wtedy, gdy współczynniki korelacji r_i są dodatnie.

1. Jeżeli $q_{ij} = r_{ij}$, to jedna zmienna objaśniająca X wystarcza za wszystkie (ta która jest najmocniej skorelowana z Y), a współczynnik determinacji jest równy kwadratowi maksymalnego współczynnika korelacji r_i .
2. Jeżeli $|q_{ij}| > |r_{ij}|$ dla każdego i, j , to wszystkie k -zmiennych stanowią podstawę budowy modeli i żadnej nie można odrzucić.
3. W pozostałych przypadkach należy posłużyć się metodą doboru zmiennych do modelu.
4. Jeżeli $|q_{ij}| < |r_{ij}|$, to ma miejsce katalizacja zmiennych, której metoda Hellwiga nie wykrywa. Katalizacja zmiennych jest wykrywana przy realizacji analizy czynnikowej.

$$Q = \begin{bmatrix} 1 & 0,92912 & 0,99159 & 0,90368 \\ 0,92912 & 1 & 0,92131 & 0,97261 \\ 0,99159 & 0,92131 & 1 & 0,89608 \\ 0,90368 & 0,97261 & 0,89608 & 1 \end{bmatrix}$$

Badając różnicę $r_{ij} - q_{ij}$ współczynników korelacji macierzy R i elementów macierzy regularnej Q , zauważamy, że są one ujemne dla każdej pary zmiennych z wyjątkiem pary (X_2, X_3) :

$$r_{23} - q_{23} = 0,03551$$

Zmienne X_2, X_3 podlegają katalizacji $r_2 < r_3$. Zmienna X_2 będąca słabiej skorelowana z Y, katalizuje zmienną X_3 mocniej skorelowaną z Y.

Tabela 3

Różnica odpowiednich elementów macierzy R i odpowiednich elementów macierzy regularnej Q

0,00000	-0,02627	-0,05036	-0,06884
-0,02627	0,00000	0,03551	-0,14178
-0,05036	0,03551	0,00000	-0,01273
-0,06884	-0,14178	-0,01273	0,00000

4. Analiza czynnikowa

Analiza czynnikowa to zespół metod i procedur statystycznych pozwalających na sprowadzenie dużej ilości badanych zmiennych do ilości znacznie mniejszej, czyli nowych czynników niewspółzależnych między sobą. Jest ona aktualnie najlepszą metodą doboru czynników kształtujących zmienność zjawiska. Za jej pomocą bez żadnych obciążeń można wyznaczyć zbiór zmiennych diagnostycznych. Ma ona przewagę nad innymi metodami w tym sensie, że przy ocenie poszczególnych zmiennych uwzględnia efekt katalizy zmiennych.

Metodę analizy czynnikowej (Grabiński, Wydymus, Zeliaś, 1982) rozpoczynamy od wyznaczenia wielomianu charakterystycznego macierzy korelacji i jego pierwiastków (wartości własnych macierzy korelacji).

Wartości własne wyznaczono metodą Kryłowa i dla każdej wartości własnej λ_i wyznaczono składową główną F_i , której współrzędne przedstawiają i -ty wiersz poniższej macierzy.

Za podstawę wyznaczenia wartości własnych wzięto połączoną macierz korelacji zmiennej Y i zmiennych objaśniających, w odróżnieniu od wyżej rozpatrywanej macierzy regularnej, która ma wymiar równy ilości zmiennych objaśniających.

Tabela 4

Wartości własne i wektory własne

4,6015	0,45201	-0,1492	-0,53036	-0,53494	0,45386
0,19804	0,4502	-0,31392	-0,37308	0,73891	0,1166
0,1309	0,44467	-0,2676	0,74751	0,0576	-0,41057
0,05385	0,46086	-0,11578	0,14227	-0,38125	0,78014
0,01572	0,42766	0,89117	0,02274	0,13849	-0,05684

Suma wartości własnych wynosi 5

Wartości własne kumulujemy wyznaczając udział każdej z nich w kształtowaniu zmienności w badanym zjawisku. Ostatnia kolumna przedstawia skumulowane udziały kształtowania zjawiska przez poszczególne składowe. Pierwsza składowa kształtuje zjawisko w 92,03%. Pierwsza i druga łącznie 95,99% itd. Największy udział w kształtowaniu zjawiska ma składowa pierwsza.

Tabela 5

Udział czynnika głównego

Wartości własne	w_i	w_i cum
4,60150	0,92030	0,92030
0,19804	0,03961	0,95991
0,13090	0,02618	0,98609
0,05385	0,01077	0,99686
0,01572	0,00314	1,00000

Gdzie:

$$w_i = \frac{\lambda_i}{\sum \lambda_i}$$

to wskaźnik struktury udziału poszczególnych składowych głównych w kształtowaniu współzależności zmiennych.

Współrzędne wektorów własnych dzielimy przez współrzędne pierwsze (odpowiadające zmiennej objaśnianej Y). Taką normalizację wektorów można wykonywać wtedy, gdy współrzędne odpowiadające zmiennej objaśnianej Y są dodatnie.

Tabela 6

Udział składowej we współrzędnych składowej o najwyższej wartości λ

1	-0,33008	-1,17334	-1,18347	1,004093
1	-0,69729	-0,8287	1,641293	0,258996
1	-0,60179	1,681044	0,129534	-0,92331
1	-0,25123	0,308705	-0,82726	1,692792
1	2,083828	0,053173	0,323832	-0,13291

Z każdego wektora składowej głównej wyznaczamy maksimum wartości bezwzględnej współrzędnej i sumujemy je.

Tabela 7

Analiza składowych głównych

1	0,330081	1,173337	1,183469	1,004093	1,183469
1	0,69729	0,828698	1,641293	0,258996	1,641293
1	0,601795	1,681044	0,129534	0,923314	1,681044
1	0,251226	0,308705	0,827258	1,692792	1,692792
1	2,083828	0,053173	0,323832	0,132909	2,083828

max

2,083828	1,681044	2,824762	1,692792
----------	----------	----------	----------

Po uporządkowaniu stwierdzamy, że odpowiadają one zmiennym X_1, X_3 . Najwięcej najmniejszych wartości współrzędnych składowych głównych jest przy zmiennej X_2 , co oznacza, że X_2 nie wywiera istotnego wpływu na Y. Nieoczyszczone ścieki przemysłowe nie wywierają więc istotnego wpływu na kształtowanie kosztów. Według analizy czynnikowej jest to najlepsza kombinacja zmiennych.

Dla potwierdzenia słuszności wyników badanie powtórzono, posługując się metodą nośników informacji Hellwiga. Analiza nośników informacji potwierdziła, że kombinacja X_1, X_3 jest najlepszą kombinacją zmiennych objaśniających. Integralny nośnik informacji tej kombinacji jest maksymalny, wynosi 0,9366 i jest nieznacznie niższy od R^2 . Efekt katalizacji zmiennych jest następujący:

$$U = R^2 - H = 0,9414 - 0,9366 = 0,0048$$

i stanowi 0,48% całkowitej informacji o badanym zjawisku.

Wyniki obu metod są identyczne, ponieważ między zmiennymi X_1 i X_3 nie występuje kataliza.

Opierając się na wyżej wymienionych zmiennych dokonano oszacowania modelu w postaci:

$$Y_t = \alpha_1 X_{1t} + \alpha_3 X_{3t} + \alpha_0 + \xi_t \quad (2)$$

Estymacji dokonano za pomocą KMNK. Po oszacowaniu model ma postać:

$$Y_t = 2,388X_{1t} + 0,0525X_{3t} + 161,5341 + \xi_t$$

Jest to model kosztów ponoszonych w latach 1986-2000 w zakresie ochrony gruntów, wód i atmosfery. Następnie dokonano oceny błędu, jakim obarczone jest oszacowane równanie i oceny poziomu dopasowania modelu do danych empirycznych (Barczak, Biolik, 2002).

Model charakteryzuje się tym, że wartości teoretyczne w stosunku do rzeczywistych wartości badanego wskaźnika różnią się średnio o $\pm 24,07$. W celu zbadania oceny poziomu dopasowania modelu do danych empirycznych obliczono współczynnik determinacji $R^2 = 0,9414$ i współczynnik zbieżności $\varphi^2 = 0,0586$. Oznacza to, że oszacowany model jest dopasowany do danych empirycznych w 94,14% zakresu jego zmienności i w takim zakresie model wyjaśnia zmienność kosztu. Model jest niedopasowany do danych empirycznych w 5,86% i w takim zakresie nie jest przez model wyjaśniona zmienność rozpatrywanego kosztu. W przypadku oceny istotności wpływu zmiennych objaśniających na zmienną objaśnianą hipotezy testowe są następujące:

$$H_0 : \alpha_i = 0; H_1 : \alpha_i \neq 0$$

Wartość sprawdzianu hipotezy H_0 dla zmiennej X_1 wynosi $t = 1,973$. Podobnie dla zmiennej X_3 , wartość statystyki t wynosi $t = 2,6127$. Wartość krytyczna odczytana z tablic rozkładu t-Studenta dla poziomu istotności $\alpha = 0,1$ wynosi $t_\alpha = 1,782$.

Zatem na poziomie istotności $\alpha = 0,1$ stwierdzamy, że zmienne X_1, X_3 , czyli grunty zdewastowane i emisja zanieczyszczeń do atmosfery, mają istotny wpływ na kształtowanie kosztów likwidacji szkód.

5. Analiza taksonomiczna obiektów

W celu szczegółowej analizy kosztu wykonano grupowanie obserwacji. Analizę taksonomiczną wykonano na podstawie przedstawionego wcześniej materiału empirycznego. W celu jego ujednoczenia przeprowadzono standaryzację wszystkich wartości, dla których wyznaczono macierz odległości taksonomicznych (Chomątowski, Sokołowski, 1978). Odległość taksonomiczna została zastosowana do kolejnych okresów. Okresy są numerowane za pomocą wskaźników $i, j = 1, 2, \dots, 15$. Natomiast k to kolejne wartości cech $k = 1, 2, \dots, 5$. Za funkcję odległości taksonomicznej przyjmujemy:

$$d_{ij} = \sqrt{\sum_{k=1}^5 (X_{ik} - X_{jk})^2} \quad (3)$$

Tabela 8

Macierz odległości taksonomicznych

Dist	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0,00	1,87	1,79	2,65	2,92	3,32	3,90	4,57	4,83	5,10	5,39	5,39	5,57	5,92	6,25
2	1,87	0,00	1,29	2,55	3,09	3,70	4,21	4,95	5,20	5,46	5,73	5,73	5,72	6,18	6,51
3	1,79	1,29	0,00	1,44	1,89	2,46	2,98	3,76	3,98	4,25	4,53	4,56	4,60	5,03	5,37
4	2,65	2,55	1,44	0,00	0,75	1,34	1,83	2,70	2,91	3,20	3,42	3,45	3,50	3,93	4,25
5	2,92	3,09	1,89	0,75	0,00	0,72	1,19	2,07	2,27	2,58	2,82	2,84	2,92	3,32	3,64
6	3,32	3,70	2,46	1,34	0,72	0,00	0,80	1,74	1,87	2,19	2,43	2,50	2,71	3,02	3,34
7	3,90	4,21	2,98	1,83	1,19	0,80	0,00	0,98	1,12	1,43	1,66	1,73	1,96	2,26	2,60
8	4,57	4,95	3,76	2,70	2,07	1,74	0,98	0,00	0,35	0,58	0,83	0,87	1,28	1,46	1,82
9	4,83	5,20	3,98	2,91	2,27	1,87	1,12	0,35	0,00	0,33	0,62	0,74	1,20	1,31	1,65
10	5,10	5,46	4,25	3,20	2,58	2,19	1,43	0,58	0,33	0,00	0,39	0,59	1,13	1,15	1,48
11	5,39	5,73	4,53	3,42	2,82	2,43	1,66	0,83	0,62	0,39	0,00	0,37	0,99	0,91	1,22
12	5,39	5,73	4,56	3,45	2,84	2,50	1,73	0,87	0,74	0,59	0,37	0,00	0,72	0,65	0,99
13	5,57	5,72	4,60	3,50	2,92	2,71	1,96	1,28	1,20	1,13	0,99	0,72	0,00	0,56	0,83
14	5,92	6,18	5,03	3,93	3,32	3,02	2,26	1,46	1,31	1,15	0,91	0,65	0,56	0,00	0,36
15	6,25	6,51	5,37	4,25	3,64	3,34	2,60	1,82	1,65	1,48	1,22	0,99	0,83	0,36	0,00
MAX	59,47	62,19	47,94	37,94	33,02	32,14	28,66	27,96	28,38	29,86	31,31	31,14	33,70	36,05	40,32

Podstawą decyzji grupowania jest poziom miary podobieństwa obiektów:

$$P_{ij}^* = \sum_{s=1}^k \min(p_{is}, p_{js}) \quad (4)$$

gdzie:

i, j – numery porównywanych obiektów,

$k = 5$ – liczba zmiennych Y, X_1, X_2, X_3, X_4 .

Wyrażenie P_{ij}^* może przyjmować wartości z przedziału $\langle 0,1 \rangle$. $P_{ij}^* = 1$, gdy obiekty są identyczne, a $P_{ij}^* = 0$, gdy są całkowicie odmienne.

Tabela 9

Macierz podobieństwa dla początkowej obserwacji

Dist	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1,00	0,95	0,93	0,88	0,82	0,77	0,65	0,50	0,47	0,45	0,43	0,42	0,47	0,45	0,48
2	0,95	1,00	0,94	0,86	0,80	0,75	0,64	0,49	0,46	0,44	0,42	0,41	0,46	0,44	0,47
3	0,93	0,94	1,00	0,89	0,83	0,76	0,64	0,48	0,44	0,42	0,41	0,40	0,44	0,42	0,45
4	0,88	0,86	0,89	1,00	0,91	0,84	0,70	0,52	0,49	0,47	0,45	0,44	0,48	0,46	0,49
5	0,82	0,80	0,83	0,91	1,00	0,90	0,76	0,58	0,54	0,51	0,49	0,48	0,52	0,50	0,53
6	0,77	0,75	0,76	0,84	0,90	1,00	0,84	0,64	0,60	0,57	0,55	0,54	0,58	0,56	0,59
7	0,65	0,64	0,64	0,70	0,76	0,84	1,00	0,78	0,74	0,71	0,69	0,68	0,72	0,69	0,71
8	0,50	0,49	0,48	0,52	0,58	0,64	0,78	1,00	0,95	0,91	0,88	0,86	0,86	0,84	0,83
9	0,47	0,46	0,44	0,49	0,54	0,60	0,74	0,95	1,00	0,95	0,92	0,89	0,87	0,85	0,84
10	0,45	0,44	0,42	0,47	0,51	0,57	0,71	0,91	0,95	1,00	0,95	0,93	0,89	0,88	0,86
11	0,43	0,42	0,41	0,45	0,49	0,55	0,69	0,88	0,92	0,95	1,00	0,96	0,90	0,90	0,88
12	0,42	0,41	0,40	0,44	0,48	0,54	0,68	0,86	0,89	0,93	0,96	1,00	0,92	0,92	0,89
13	0,47	0,46	0,44	0,48	0,52	0,58	0,72	0,86	0,87	0,89	0,90	0,92	1,00	0,96	0,95
14	0,45	0,44	0,42	0,46	0,50	0,56	0,69	0,84	0,85	0,88	0,90	0,92	0,96	1,00	0,96
15	0,48	0,47	0,45	0,49	0,53	0,59	0,71	0,83	0,84	0,86	0,88	0,89	0,95	0,96	1,00

$P_{ij} = 1 - P_{ij}^*$ jest prawdopodobieństwem braku podobieństwa między strukturami. Macierz $P = [P_{ij}]$ jest macierzą symetryczną.

Tabela 10

Macierz braku podobieństwa

Dist	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0,00	0,05	0,07	0,12	0,18	0,23	0,35	0,50	0,53	0,55	0,57	0,58	0,53	0,55	0,52
2	0,05	0,00	0,06	0,14	0,20	0,25	0,36	0,51	0,54	0,56	0,58	0,59	0,54	0,56	0,53
3	0,07	0,06	0,00	0,11	0,17	0,24	0,36	0,52	0,56	0,58	0,59	0,60	0,56	0,58	0,55
4	0,12	0,14	0,11	0,00	0,09	0,16	0,30	0,48	0,51	0,53	0,55	0,56	0,52	0,54	0,51
5	0,18	0,20	0,17	0,09	0,00	0,10	0,24	0,42	0,46	0,49	0,51	0,52	0,48	0,50	0,47
6	0,23	0,25	0,24	0,16	0,10	0,00	0,16	0,36	0,40	0,43	0,45	0,46	0,42	0,44	0,41
7	0,35	0,36	0,36	0,30	0,24	0,16	0,00	0,22	0,26	0,29	0,31	0,32	0,28	0,31	0,29
8	0,50	0,51	0,52	0,48	0,42	0,36	0,22	0,00	0,05	0,09	0,12	0,14	0,14	0,16	0,17
9	0,53	0,54	0,56	0,51	0,46	0,40	0,26	0,05	0,00	0,05	0,08	0,11	0,13	0,15	0,16
10	0,55	0,56	0,58	0,53	0,49	0,43	0,29	0,09	0,05	0,00	0,05	0,07	0,11	0,12	0,14
11	0,57	0,58	0,59	0,55	0,51	0,45	0,31	0,12	0,08	0,05	0,00	0,04	0,10	0,10	0,12
12	0,58	0,59	0,60	0,56	0,52	0,46	0,32	0,14	0,11	0,07	0,04	0,00	0,08	0,08	0,11
13	0,53	0,54	0,56	0,52	0,48	0,42	0,28	0,14	0,13	0,11	0,10	0,08	0,00	0,04	0,05
14	0,55	0,56	0,58	0,54	0,50	0,44	0,31	0,16	0,15	0,12	0,10	0,08	0,04	0,00	0,04
15	0,52	0,53	0,55	0,51	0,47	0,41	0,29	0,17	0,16	0,14	0,12	0,11	0,05	0,04	0,00

Przyjmijmy krytyczny poziom braku podobieństwa $\alpha = 0,1$. Budujemy macierz zero-jedynkową wstawiając zero, gdy $d < \alpha$ oraz jeden, gdy $d \geq \alpha$ i zliczamy wartości wierszami (wektor ostatni P_0). Wyznaczamy największą współrzędną wektora P_0 i eliminujemy obserwację, która jej odpowiada. Z macierzy P usuwamy odpowiedni jej wiersz zastępując kolumnę symbolem „*“.

Tabela 11

Macierz po usunięciu obserwacji odpowiadającej największej współrzędnej wektora P_0

Dist	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	$P_0 = \sum P_{ij}$
1	0,00	0,00	0,00	1,00	1,00	1,00	*	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	11,00
2	0,00	0,00	0,00	1,00	1,00	1,00	*	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	11,00
3	0,00	0,00	0,00	1,00	1,00	1,00	*	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	11,00
4	1,00	1,00	1,00	0,00	0,00	1,00	*	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	12,00
5	1,00	1,00	1,00	0,00	0,00	0,00	*	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	11,00
6	1,00	1,00	1,00	1,00	0,00	0,00	*	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	12,00
7	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0,00
8	1,00	1,00	1,00	1,00	1,00	1,00	*	0,00	0,00	0,00	1,00	1,00	1,00	1,00	1,00	11,00
9	1,00	1,00	1,00	1,00	1,00	1,00	*	0,00	0,00	0,00	0,00	1,00	1,00	1,00	1,00	10,00
10	1,00	1,00	1,00	1,00	1,00	1,00	*	0,00	0,00	0,00	0,00	0,00	1,00	1,00	1,00	9,00
11	1,00	1,00	1,00	1,00	1,00	1,00	*	1,00	0,00	0,00	0,00	0,00	0,00	0,00	1,00	8,00
12	1,00	1,00	1,00	1,00	1,00	1,00	*	1,00	1,00	0,00	0,00	0,00	0,00	0,00	1,00	9,00
13	1,00	1,00	1,00	1,00	1,00	1,00	*	1,00	1,00	1,00	0,00	0,00	0,00	0,00	0,00	9,00
14	1,00	1,00	1,00	1,00	1,00	1,00	*	1,00	1,00	1,00	0,00	0,00	0,00	0,00	0,00	9,00
15	1,00	1,00	1,00	1,00	1,00	1,00	*	1,00	1,00	1,00	1,00	1,00	0,00	0,00	0,00	11,00
															MAX	12,00

Postępujemy tak do czasu, gdy P_0 będzie wektorem zerowym, a nieusuńnięte elementy z macierzy dystansów wyznaczą grupę obiektów podobnych na poziomie prawdopodobieństwa $1 - \alpha$.

Tabela 12

Zmodyfikowana macierz dystansów dla ostatniego kroku I etapu

Dist	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	$P_0 = \sum P_{ij}$
1	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0,00
2	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0,00
3	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0,00
4	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0,00
5	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0,00
6	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0,00
7	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0,00
8	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0,00
9	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0,00
10	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0,00
11	*	*	*	*	*	*	*	*	*	*	0,00	0,00	0,00	0,00	*	0,00
12	*	*	*	*	*	*	*	*	*	*	0,00	0,00	0,00	0,00	*	0,00
13	*	*	*	*	*	*	*	*	*	*	0,00	0,00	0,00	0,00	*	0,00
14	*	*	*	*	*	*	*	*	*	*	0,00	0,00	0,00	0,00	*	0,00
15	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	0,00
															MAX	0,00

W przeprowadzonym pierwszym etapie została wyznaczona grupa $G_1 = (11, 12, 13, 14)$ obiektów podobnych.

Procedurę powtarzamy od nowa, biorąc macierz z usuniętymi wierszami i kolumnami odpowiadającymi wybranej grupie.

Etap drugi

Tabela 13

Zmodyfikowana macierz dystansów dla ostatniego kroku II etapu

Dist	1	2	3	4	5	6	7	8	9	10	15	$P_0 = \sum P_{ij}$
1	*	*	*	*	*	*	*	*	*	*	*	0,00
2	*	*	*	*	*	*	*	*	*	*	*	0,00
3	*	*	*	*	*	*	*	*	*	*	*	0,00
4	*	*	*	*	*	*	*	*	*	*	*	0,00
5	*	*	*	*	*	*	*	*	*	*	*	0,00
6	*	*	*	*	*	*	*	*	*	*	*	0,00
7	*	*	*	*	*	*	*	*	*	*	*	0,00
8	*	*	*	*	*	*	*	0,00	0,00	0,00	*	0,00
9	*	*	*	*	*	*	*	0,00	0,00	0,00	*	0,00
10	*	*	*	*	*	*	*	0,00	0,00	0,00	*	0,00
15	*	*	*	*	*	*	*	*	*	*	*	0,00
											MAX	0,00

$G_2 = (8, 9, 10)$ to druga grupa obiektów podobnych.

Etap trzeci

Tabela 14

Zmodyfikowana macierz dystansów dla ostatniego kroku III etapu

Dist	1	2	3	4	5	6	7	15	$P_0 = \sum P_{ij}$
1	0,00	0,00	0,00	*	*	*	*	*	0,00
2	0,00	0,00	0,00	*	*	*	*	*	0,00
3	0,00	0,00	0,00	*	*	*	*	*	0,00
4	*	*	*	*	*	*	*	*	0,00
5	*	*	*	*	*	*	*	*	0,00
6	*	*	*	*	*	*	*	*	0,00
7	*	*	*	*	*	*	*	*	0,00
15	*	*	*	*	*	*	*	*	0,00
								MAX	0,00

$G_3 = (1, 2, 3)$ to trzecia grupa obiektów podobnych.

Etap czwarty

Tabela 15

Zmodyfikowana macierz dystansów dla ostatniego kroku IV etapu

Dist	4	5	6	7	15	$P_0 = \sum P_{ij}$
4	*	*	*	*	*	0,00
5	*	0,00	0,00	*	*	0,00
6	*	0,00	0,00	*	*	0,00
7	*	*	*	*	*	0,00
15	*	*	*	*	*	0,00
					MAX	0,00

W wyniku powyższego etapu otrzymaliśmy następujące grupy:

$$G_1 = (11,12,13,14),$$

$$G_2 = (8,9,10),$$

$$G_3 = (1,2,3),$$

$$G_4 = (5,6),$$

$$G_5 = (4,7,15).$$

Sprawdzamy, czy wynik grupowania jest poprawny. W tym celu obliczamy wariancję wewnątrzgrupową:

$$S_w^2 = \frac{1}{n} \sum_{j=1}^{k+1} S_j^2 \quad (5)$$

oraz wariancję międzygrupową:

$$S_M^2 = \frac{1}{n} \sum_{j=1}^{k+1} (\bar{X}_j - \bar{X})^2 \quad (6)$$

W tab. 16 obliczono różnicę wariancji międzygrupowej i wariancji wewnątrzgrupowej. Wariancja międzygrupowa powinna być większa od wariancji wewnątrzgrupowej w porównywanych parach grup.

Tabela 16

Różnica wariancji międzygrupowej i wariancji wewnątrzgrupowej

	16447,99	679538,3	34856,51	2526,263
16447,99		819295	35561,84	18248,35
668997,2	819295		1091118	584281,1
34068,26	34773,59	1091118		78821,09
-301902	-302323	274251	-240962	

Grupa trzecia ma wiersz i kolumnę różnic dodatnią, jest więc na pewno prawidłowo dobrana, natomiast grupa piąta ma w wierszu trzy ujemne wartości różnic. Oznacza to, że grupy piątej nie można traktować jako kompletu. Porównanie wartości zmiennej X_3 dla obserwacji piętnastej sugeruje odrzucenie tej obserwacji z powodu wartości zmiennej X_3 odstającej od pozostałych. Zatem obserwację piętnastą eliminujemy jako nietypową. Grupa piąta zawiera obserwację 5 oraz 7.

Podsumowując, obserwację 15 odrzucamy jako nietypową, a grupy jednorodnie są następujące:

$$G_1 = (11, 12, 13, 14),$$

$$G_2 = (8, 9, 10),$$

$$G_3 = (1, 2, 3),$$

$$G_4 = (5, 6),$$

$$G_5 = (4, 7).$$

Ponownie wyznaczamy macierz różnic między wariancją międzygrupową i wewnątrzgrupową.

Tabela 17

Macierz różnic między wariancją międzygrupową i wewnątrzgrupową

	16447,991	679538,32	34856,51	130854,9
16447,991		819294,96	35561,844	198456,96
668997,18	819294,96		1091117,6	209063,98
34068,256	34773,589	1091117,6		327144,9
62047,946	113506,98	134655,14	242983,18	

Wszystkie elementy macierzy różnic między wariancją międzygrupową i wewnątrzgrupową są dodatnie, co świadczy o prawidłowo dobranych grupach. Lata 1997-2000 charakteryzują się podobnymi kosztami likwidacji szkód (ochrona gruntów, rzek, powietrza) oraz podobnym poziomem zdewastowanych gruntów, zanieczyszczonych ściekami i odpadami komunalnymi. Podobna sytuacja ma miejsce w pozostałych wyznaczonych grupach. Wartości odpowiednich zmiennych w okresach zaliczonych do tych samych nie wykazują statystycznych różnic, również koszty kształtują się na podobnym poziomie. Natomiast między okresami zaliczonymi do różnych grup, różnice między wartościami empirycznymi są istotne statystycznie.

6. Opracowanie wskaźnika syntetycznego rozwoju

Na podstawie zdefiniowanych danych statystycznych w wyodrębnionych grupach wyznaczono syntetyczne wskaźniki rozwoju (Strahl, 1990). Dla grupy G_1 syntetyczny miernik rozwoju wynosi $M_1 = 0,0553818$, dla grupy G_2 : $M_2 = 0,033127$, dla G_3 : $M_3 = 0,0729116$, dla G_4 : $M_4 = 0,148748$ oraz dla G_5 : $M_5 = 0,189565$. Zatem:

$$M = \sum_{i=1}^5 M_i = 0,49973$$

co oznacza, że w najbliższych latach badane zjawisko w 49,97% będzie kształtowało się w analogiczny sposób, jak w badanych latach. W pozostałych 50,03% należy podjąć decyzję dotyczącą nowych kierunków rozwoju zjawiska.

Wnioski

W niniejszym opracowaniu zbadano zależność kosztów likwidacji szkód od czterech potencjalnych zmiennych objaśniających: X_1 (grunty zdewastowane w tys. ha), X_2 (ścieki przemysłowe nieoczyszczone w hm^3), X_3 (emisja zanieczyszczeń do atmosfery w tys. ton), X_4 (odpady komunalne w ciągu roku w mln ton).

Przez zastosowanie analizy czynnikowej i metody Hellwiga spośród tych zmiennych wybrano zmienne X_1 i X_3 , mające największy wpływ na wielkość kosztów likwidacji szkód.

Analiza czynnikowa i metoda Hellwiga dały jednakowy rezultat, ponieważ wśród wyznaczonych zmiennych nie ma zmiennych katalitycznych. Do ostatecznego zbioru zmiennych objaśniających weszła tylko jedna z par zmiennych katalitycznych.

Dla wszystkich zmiennych zastosowano grupowanie obiektów w celu wyodrębnienia jednorodnych grup $G_1 = (11,12,13,14)$, $G_2 = (8,9,10)$, $G_3 = (1,2,3)$, $G_4 = (5,6)$, $G_5 = (4,7)$. W grupach tych wyznaczono syntetyczny wskaźnik rozwoju $M = \sum_{i=1}^5 M_i = 0,49973$. Oznacza to, że w 49,97% zmienność

kosztu będzie w przyszłości kształtowana przez czynniki, które zostały uwzględnione w przedstawionym modelu (inne czynniki również będą miały istotny wpływ na kształtowanie się kosztu w przyszłości). Zmienność kosztu będzie kształtowana przez te czynniki w 50,03% zakresu zmienności kosztu i w przyszłości trzeba je będzie wykryć oraz dołączyć do analizy.

Literatura

- Ostasiewicz W.: *Statystyczne metody analizy danych*. AE, Wrocław 1998.
- Strahl D.: *Metody programowania rozwoju społeczno-gospodarczego*. PWE, Warszawa 1990.
- Grafiński T., Wydymus S., Zeliaś A.: *Metody doboru zmiennych w modelach ekonometrycznych*. PWN, Warszawa 1982.
- Grafiński T., Wydymus S., Zeliaś A.: *Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych*. PWN, Warszawa 1989.
- Barczak A., Biolik J.: *Podstawy ekonometrii*. AE, Katowice 2002.
- Chomątowski S., Sokołowski A.: *Taksonomia struktur*. „Przegląd Statystyczny” 1978.
- Rocznik Statystyczny*. GUS, Warszawa 1986-2000.

**THE SPATIAL ANALYSIS
OF THE COST OF THE LIQUIDATION OF THE DAMAGES
IN THE PROTECTION OF THE ENVIRONMENT**

Summary

The paper investigates the dependence of the costs of damage liquidation according to four explanatory variables. To select the factors forming the variability of the phenomenon the best method at the present moment has been chosen, i.e. factor analysis by means of which it is possible to set a class of the diagnostic variables without any encumbrances. Grouping the objects has enabled to separate the homogenous groups $G_1 = (11,12,13,14)$, $G_2 = (8,9,10)$, $G_3 = (1,2,3)$, $G_4 = (5,6)$, $G_5 = (4,7)$. In the groups the synthetic determinant of the development has been formed up to the standard $M_1 = 0,0553818$, $M_2 = 0,033127$, $M_3 = 0,0729116$, $M_4 = 0,148748$, $M_5 = 0,189565$. In consequence of the groups being separable the synthetic measure of the whole collectivity is an additive total of fragmentary measures and it amounts to

$$M = \sum_{i=1}^5 M_i = 0,49973.$$