

Agnieszka Skowronek-Grądziel, Wiktor Kołwzan, Andrzej M. Dziubek

## Zastosowanie statystycznych metod wielowymiarowej analizy porównawczej w zarządzaniu rozwojem zaopatrzenia w wodę i sanitacji obszarów wiejskich w Polsce w latach 2000–2012

Wiedza w zakresie stanu i rozwoju infrastruktury technicznej na danym obszarze, odpowiedzialnej za zaopatrzenie w wodę oraz usuwanie i unieszkodliwianie ścieków i odpadów, jest niezwykle ważna na wszystkich etapach zarządzania daną aglomeracją, ponieważ daje podstawy do podejmowania optymalnych decyzji związanych z ogólnie pojętą inżynierią środowiska. Pozwala także na kompleksową ocenę bieżącej sytuacji na danym obszarze oraz na formułowanie przyszłych celów w tym zakresie. W obszarze inżynierii środowiska wiedza z zakresu stanu systemów wodociągowych i kanalizacyjnych jest niezbędna do formułowania celów i strategii ochrony środowiska w skali makro, organizowania podsystemu technicznego i podsystemu struktury, projektowania systemu informacyjno-decyzyjnego, kształtowania systemu kierowania jednostkami odpowiedzialnymi za ochronę środowiska w całym kraju, jak również do podejmowania decyzji dotyczących zmian na poziomie technologii stosowanych w zakresie zaopatrzenia w wodę i sanitacji obszarów wiejskich [1], czyli ogólnie – do usprawniania funkcjonowania organów państwa właściwych w zakresie środowiskowych działań ochronnych.

Jednym z głównych problemów związanych z jakością środowiska naturalnego jest ochrona wód przed zanieczyszczeniem związkami azotu, a zwłaszcza azotanami [2]. Problem ten jest szczególnie istotny na obszarach o rozwiniętej produkcji rolnej i hodowlanej, a także na terenach o słabej infrastrukturze w zakresie odprowadzania i oczyszczania ścieków oraz zagospodarowania odpadów. Źródłem nadmiernej ilości azotanów w wodach są spływy powierzchniowe z pól uprawnych oraz odcieki ze składowisk odpadów, a także ścieki nieobjęte systemem kanalizacji. Ponieważ usuwanie szkodliwych dla zdrowia azotanów z wody ujmowanej do celów wodociągowych jest bardzo kłopotliwe i kosztowne, dlatego duży nacisk kładzie się na zapobieganie ich przedostawaniu się do środowiska wodnego, przede wszystkim przez budowę systemów kanalizacji i oczyszczalni ścieków, a także przez działania wynikające z tzw. dyrektywy azotanowej (Dyrektywa Rady 91/676/EWG z 12 grudnia 1991 r.), dotyczącej ochrony wód przed zanieczyszczeniami powodowanymi przez azotany pochodzenia rolniczego. Przystępując do Unii Europejskiej

Polska zobowiązała się także do wdrożenia dyrektywy dotyczącej oczyszczania ścieków (Dyrektywa Rady 91/271/EWG z 21 maja 1991 r.), co w efekcie ma doprowadzić do uporządkowania stanu gospodarki ściekowej. W tym celu został utworzony krajowy program oczyszczania ścieków komunalnych, którego założeniem było uporządkowanie do końca 2015 r. infrastruktury technicznej odpowiedzialnej za stan gospodarki ściekowej w Polsce, tzn. objęcie systemami kanalizacji i oczyszczania ścieków praktycznie 100% mieszkańców miast i około 60% mieszkańców terenów wiejskich. Program ten przewidywał budowę i modernizację ponad 30 tys. km sieci kanalizacyjnej oraz budowę, rozbudowę i modernizację kilkuset oczyszczalni ścieków kosztem ponad 30 mld złotych.

Mając na uwadze skalę tego przedsięwzięcia, skuteczne wykorzystanie wyników uzyskanych w trakcie badań nad stanem infrastruktury technicznej służącej zaopatrzeniu w wodę i sanitacji obszarów wiejskich w dłuższym przedziale czasu powinno się przełożyć na ograniczenie zanieczyszczenia wód i powierzchni ziemi, a przez to do poprawy atrakcyjności inwestycyjnej danego obszaru.

W niniejszym artykule, stosując metody wielowymiarowej analizy statystycznej, starano się w wymiarze ogólnym całego kraju zdiagnozować poziom rozwoju tej infrastruktury oraz określić występujące w tym zakresie tendencje. W głównej hipotezie badawczej pracy założono, że od początku XXI wieku, w miarę upływu czasu, następuje systematyczny rozwój infrastruktury technicznej niezbędnej do poprawy stanu zaopatrzenia w wodę i sanitacji obszarów wiejskich w Polsce. Słuszność tego założenia starano się wykazać w logiczny sposób poprzez właściwy dobór metod badawczych, adekwatnych do natury danych o charakterze statystycznym. Ponadto otrzymane wyniki takiej analizy powinny mieć znaczący wymiar użyteczny.

### Dobór cech diagnostycznych

W artykule poddano analizie rozwój obszarów wiejskich w Polsce w latach 2000, 2003, 2005 i 2012 ze względu na zmiany zachodzące w zakresie zaopatrzenia w wodę, kanalizacji i zbiorczych oczyszczalni ścieków, a także indywidualnych wiejskich oczyszczalni ścieków i zorganizowanych składowisk odpadów komunalnych. Rok 2000 wytypowano jako wyjściowy – pierwszy po wprowadzeniu reformy terytorialnej w Polsce. Rok 2003, bezpośrednio przed akcesją Polski do Unii Europejskiej, uznano jako istotny w tym badaniu, ponieważ państwo polskie, chcąc przystąpić do struktur UE, musiało spełnić określone

Dr A. Skowronek-Grądziel, dr W. Kołwzan: Wyższa Szkoła Oficerska Wojsk Lądowych im. Generała Tadeusza Kościuszki we Wrocławiu, Wydział Zarządzania, ul. Piotra Czajkowskiego 109, 51-150 Wrocław  
a.skowronek\_gradziel@wso.wroc.pl

Dr inż. A.M. Dziubek: Politechnika Wroclawska, Wydział Inżynierii Środowiska, Katedra Technologii Oczyszczania Wody i Ścieków, wybrzeże Stanisława Wyspiańskiego 27, 50-370 Wrocław  
andrzej.dziubek@pwr.edu.pl

wymagania związane z funkcjonowaniem infrastruktury służącej ochronie środowiska. W tym też czasie odnotowano największe tempo wzrostu nakładów finansowych w analizowanym obszarze – 145% względem roku poprzedniego i ponad 172% w stosunku do roku 2000. Rok 2005 to czas bezpośrednio po uzyskaniu przez Polskę członkostwa w UE. Nakłady na infrastrukturę służącą ochronie środowiska na obszarach wiejskich zmniejszyły się w stosunku do roku poprzedniego o przeszło 15%, a względem roku wyjściowego (2000) wzrosły o niecałe 11%. Z kolei rok 2012 poddano analizie jako najbardziej aktualny do określenia zmian dokonanych w zakresie zaopatrzenia w wodę i sanitacji obszarów wiejskich. Tempo wzrostu nakładów inwestycyjnych w roku sprawozdawczym 2012 było – jak dotychczas – najniższe i w stosunku do roku poprzedniego wyniosło zaledwie 61,5%. Obserwowane spowolnienie mogło wynikać między innymi z faktu, iż rok 2012 był przedostatnim rokiem perspektywy finansowej 2007–2013 (przedział czasu, na który planowany jest budżet UE), zatem można przypuszczać, że kończy się realizacja wyznaczonych zadań, a tym samym i środki finansowe.

Określając zmienne do badania starano się, aby były poprawne merytorycznie oraz odpowiadały także statystycznym zasadom doboru [3]. W celu spełnienia założeń, do analiz zakwalifikowano dane uzyskane z Ministerstwa Rolnictwa i Rozwoju Wsi oraz z Banku Danych Lokalnych [4] mierzone względem najsilniejszej skali pomiarowej – ilorazowej, która daje możliwość stosowania wszelkich przekształceń. Wytypowane surowe dane przeliczono, dzięki czemu uzyskano wartości stosunkowe i możliwość porównywania obiektów. Dodatkowo zmienne musiały się charakteryzować stosownymi właściwościami dyskryminacyjnymi – odpowiednio wysoką zmiennością (w 2000 r. wartości współczynników zmienności wahały się od 25% do 263%, w 2003 r. od 36% do 360%, w 2005 r. od 30% do 122%, a w 2012 r. od 28% do 225%) oraz musiały być stymulantami (zmienna, której wzrost wartości świadczy o wzroście poziomu zjawiska złożonego). Kierując się tymi zasadami, do badania zakwalifikowano zbiór dwudziestu pięciu zmiennych diagnostycznych (tab. 1). Przeprowadzone obliczenia i analizy wykonano korzystając z pakietów komputerowych Excel, GRETL i Statistica.

Tabela 1. Zmienne diagnostyczne  
Table 1. Diagnostic variables

Symbol zmiennej	Opis zmiennej
X <sub>1</sub>	Wskaźnik zwodociągowania – liczba przyłączy wodociągowych przypadających na jedną miejscowość wiejską (szt./miejscowość)
X <sub>2</sub>	Liczba nowych przyłączy do budynków przypadających na jedną miejscowość wiejską (szt./miejscowość)
X <sub>3</sub>	Długość sieci wodociągowej przypadająca na powierzchnię obszarów wiejskich województwa (km/km <sup>2</sup> )
X <sub>4</sub>	Nakłady na wybudowanie sieci wodociągowej przypadające na jedną miejscowość wiejską (tys. zł/miejscowość)
X <sub>5</sub>	Długość sieci kanalizacji sanitarnej przypadającej na jedną miejscowość wiejską (km/miejscowość)
X <sub>6</sub>	Wskaźnik skanalizowania – liczba przyłączy kanalizacyjnych przypadających na jedną miejscowość wiejską (szt./miejscowość)
X <sub>7</sub>	Przeciętna liczba nowych przykanalików przypadających na jedną miejscowość wiejską województwa (szt./miejscowość)
X <sub>8</sub>	Nakłady na budowę sieci kanalizacyjnej w przeliczeniu na statystyczną miejscowość (tys. zł/miejscowość)
X <sub>9</sub>	Liczba nowych zbiorczych oczyszczalni ścieków przypadających na istniejące oczyszczalnie (%)
X <sub>10</sub>	Łączna przepustowość wszystkich zbiorczych oczyszczalni ścieków (średnia ważona na jedno województwo, m <sup>3</sup> /d)
X <sub>11</sub>	Przeciętne nakłady na nową oczyszczalnię zbiorczą w przeliczeniu na jedną miejscowość wiejską województwa (tys. zł/miejscowość)
X <sub>12</sub>	Liczba zmodernizowanych zbiorczych oczyszczalni ścieków przypadających na jedną miejscowość wiejską (szt./miejscowość)
X <sub>13</sub>	Wskaźnik zmodernizowanych oczyszczalni zbiorczych do istniejących oczyszczalni zbiorczych (%)
X <sub>14</sub>	Stosunek rocznych nakładów na modernizację zbiorczych oczyszczalni ścieków do nakładów na nowe zbiorcze oczyszczalnie ścieków (%)
X <sub>15</sub>	Oczyszczalnie przyzagrodowe przypadające na jedną miejscowość wiejską (szt./miejscowość)
X <sub>16</sub>	Przeciętne nakłady na indywidualną wiejską oczyszczalnię ścieków (tys. zł/szt.)
X <sub>17</sub>	Liczba zmodernizowanych stacji oczyszczania wody przypadających na jedną miejscowość wiejską (szt./miejscowość)
X <sub>18</sub>	Liczba oddanych do użytku stacji oczyszczania wody przypadających na istniejącą stację wodociągową na obszarach wiejskich województwa (%)
X <sub>19</sub>	Miejscowości wiejskie obsługiwane przez stacje oczyszczania wody (szt./miejscowość)
X <sub>20</sub>	Nakłady na modernizację stacji oczyszczania wody w przeliczeniu na jedną miejscowość wiejską (tys. zł/miejscowość)
X <sub>21</sub>	Nakłady na stacje oczyszczania wody w przeliczeniu na jedną miejscowość wiejską (tys. zł/miejscowość)
X <sub>22</sub>	Wskaźnik zmodernizowanych stacji oczyszczania wody do stacji istniejących (%)
X <sub>23</sub>	Stosunek rocznych nakładów na modernizację stacji oczyszczania wody do nakładów na nowe stacje wodociągowe (%)
X <sub>24</sub>	Przeciętna wielkość składowiska odpadów (ha)
X <sub>25</sub>	Roczne nakłady na nowe gminne składowiska odpadów (tys. zł/miejscowość)

## Metody badawcze oparte na wielowymiarowej analizie statystycznej

Podjęcie decyzji w przypadku dużej liczby danych o charakterze losowym powinno być poprzedzone zastosowaniem analizy statystycznej (tzw. myślenie statystyczne) [5]. Takie podejście wymaga umiejętności zastosowania odpowiednich narzędzi statystycznych odpowiadających naturze zgromadzonych danych. Na ogół, kiedy badane obiekty charakteryzują się wieloma cechami o charakterze losowym (jak w omawianym przypadku), do ich interpretacji należy zastosować metody wielowymiarowej analizy statystycznej, które znajdują liczne zastosowania praktyczne ze względu na proste wymagania dotyczące danych statystycznych. Wystarczy mieć do dyspozycji kilka zmiennych, czyli cech o takiej samej liczebności, a więc pewną liczbę badanych obiektów, a można już z powodzeniem stosować techniki badawcze oparte na wielowymiarowej analizie statystycznej i uzyskać interesujące wyniki w postaci interpretacyjnej w nowym – mniejszym w stosunku do danych pierwotnych – wymiarze. Ponadto dane te mogą mieć zarówno charakter przekrojowy, czyli zostać pobrane w jednym czasie z różnych miejsc, jak i dynamiczny, a więc zarejestrowane w różnych przedziałach czasu. Dopuszcza się też do takiej analizy dane o charakterze jakościowym. Podstawa merytoryczna tych metod sprowadza się do kilku zasadniczych elementów [6–11]:

- taksonomicznego, polegającego na porządkowaniu i grupowaniu badanych obiektów,

- analizy czynnikowej, która polega na przedstawieniu danego zbioru cech (na ogół zmiennych losowych) badanych obiektów dowolnej natury (matematycznie badane obiekty są punktami w pewnej przestrzeni, natomiast wymiar tej przestrzeni stanowią badane cechy; zatem liczba cech stanowi wymiar przestrzeni matematycznej, a liczba obiektów – liczbę badanych punktów w tej przestrzeni), jako kombinacji liniowej nowych nieobserwowanych bezpośrednio charakterystyk, nazywanych czynnikami (składowymi); czynniki te mają ważną właściwość – są ortogonalne względem siebie (czyli niezależne) i jest ich mniej niż cech pierwotnie wziętych do analizy zmiennych, zatem tworzą nowy układ przestrzenny (kartyzjański); pozwala to na wyjaśnienie i interpretowanie struktury powiązań między cechami badanych obiektów,

- wieloczynnikowej lub wielowymiarowej analizy wariancji (ANOVA, MANOVA), która w obrębie ustalonej liczby czynników pozwala zbadać, czy wywierają one wpływ (i jak wielki) na kształtowanie się średnich wartości cech mierzalnych badanych obiektów.

Przedstawione metody badawcze w statystyce nazywane są wielowymiarową analizą porównawczą (WAP). Korzyści wynikające ze stosowania WAP w badaniu w zasadzie dowolnego procesu o układzie obiekty–cechy są następujące:

- zmniejszenie często ogromnej liczby informacji o badanych obiektach do zaledwie kilku bazowych kategorii, które mogą być dopiero poddane dokładniejszej analizie statystycznej i treściowej,

- otrzymanie jednorodnych grup obiektów i cech pod względem ich właściwości, co w konsekwencji prowadzi do znalezienia (nazwania) ich podstawowych właściwości, czyli wyrażenia ich w nowym wymiarze,

- skupienie uwagi na najbardziej typowych w przypadku rozważanego problemu badawczego zjawiskach, kategoriach, cechach itp.,

- określenie i wyjaśnienie – co jest najważniejsze – struktury powiązań zarówno między cechami badanych obiektów, jak i między samymi obiektami ze względu na cechy reprezentowane przez te obiekty.

W analizie problemu badanego w niniejszej pracy wykorzystano w szczególności:

- metodę głównych składowych, która umożliwia wyznaczenie czynników warunkujących badane zjawisko,

- analizę skupień, która stanowi metodę klasyfikacji obiektów (bądź cech) i jest w pewnym stopniu kontynuacją analizy składowych głównych, gdyż wydzielona liczba składowych (czynników) determinuje do pewnego stopnia liczbę klas wydzielanych przyjętą metodą klasyfikacji w obrębie analizy skupień, mimo iż formalnie w obrębie wybranej metody można tworzyć dowolną liczbę klas obiektów, lecz ograniczoną liczebnością badanych obiektów (bądź zmiennych),

- wieloczynnikową analizę wariancji, która pozwala zbadać interaktywne oddziaływania pomiędzy wyodrębnionymi grupami,

- analizę korelacji, pozwalającą z kolei badać bezpośrednie związki (liniowe) pomiędzy parami zmiennych, a nawet pomiędzy zbiorami zmiennych,

- analizę podstawowych statystyk z próby (takich jak średnia, odchylenie standardowe, współczynnik zmienności), która pozwala, bez testowania ich statystycznej istotności i sprawdzenia odpowiednich założeń statystycznych, orzec wstępnie, jak się one zachowują (np. w czasie, bądź w miarę upływu czasu); taka wstępna informacja pozwala na zastosowanie bardziej skomplikowanych metod analizy statystycznej,

- metodę Hellwiga, za pomocą której można na podstawie macierzy korelacji wydzielić zmienne wiodące (centralne) i tzw. zmienne satelitarne; prowadzi to do zmniejszenia liczby zmiennych do takich, które w danym analizowanym procesie mają charakter zmiennych wiodących (macierz korelacji bierze też pośrednio udział w klasyfikacji obiektów oraz w analizie czynnikowej).

Jako przykład formalnych procedur obliczeniowych związanych z wymienionymi wcześniej metodami można potraktować metodę Hellwiga, która jest pod względem złożoności logicznej bardzo prosta i przedstawia się następująco: w odniesieniu do każdej zmiennej liczy się wartość jej współczynnika zmienności; z kolei przy ustalonym arbitralnie progu (10% lub 15%) z listy zmiennych eliminuje się te, których zmienność jest mniejsza od przyjętej wartości progowej. Następnie w nowym układzie zmiennych wyznacza się macierz zachodzących między nimi korelacji i oblicza się przy danej liczbie stopni swobody wartość krytyczną współczynnika korelacji, która może być przyjęta arbitralnie (z reguły na poziomie 0,5) lub też wyznaczona według dwóch możliwych metod:

- w pierwszej wartości progową współczynnika korelacji wyznacza się na podstawie formuły:

$$r^* = \min_i \max_j |r_{ij}|, i, j = 1, 2, \dots, k \quad (1)$$

w której  $k$  oznacza liczbę zmiennych

- w drugiej formuła opisująca wartość progową współczynnika korelacji ma postać:

$$r^* = \sqrt{\frac{t_\alpha^2}{t_\alpha^2 + n - 2}} \quad (2)$$

w której  $t_{\alpha}^2$  jest wartością statystyki z tablicy testu t-Studenta przy  $n-2$  stopni swobody oraz przyjętym poziomie istotności  $\alpha$ , zaś  $n$  liczbą obserwacji w przypadku każdej zmiennej.

Algorytm metody Hellwiga, jako szczególny przypadek tzw. metody parametrycznej (lub pojemnościowej), przebiega w następujących po sobie pięciu etapach (krokach):

- wyznaczenie sumy bezwzględnych wartości współczynników korelacji w każdej kolumnie (wierszu) macierzy korelacji,

- wybór kolumny (wiersza), której suma współczynników korelacji jest największa,

- w kolumnie (wierszu) o maksymalnej sumie dokonuje się wyboru elementów o wartościach bezwzględnych większych od wartości progowej  $r^*$  oraz odpowiadających tym elementom wierszy (kolumn); potencjalną zmienną odpowiadającą wyłonionej kolumnie (wierszowi) nazywa się zmienną centralną, a zmienne odpowiadające wyróżnionym wierszom (kolumnom) jej zmiennymi satelitarnymi (razem tworzą one tzw. skupienie); zmienne tworzące dane skupienie są ze sobą istotnie skorelowane, co oznacza, że każda z nich w znaczącym stopniu powiela te same informacje – stąd mogą być reprezentowane przez jedną tylko zmienną centralną,

- redukcja macierzy korelacji  $R$  przez wykreślenie wyróżnionej kolumny (wiersza) oraz wyróżnionych wierszy (kolumn),

- powtórzenie wcześniejszych kroków z użyciem zredukowanych macierzy korelacji; w ten sposób otrzymuje się kolejne skupienia, aż do wyczerpania zbioru zmiennych (skupienia mogą tworzyć pojedyncze zmienne nazywane izolowanymi).

Do ostatecznego zbioru zmiennych (diagnostycznych), będącego podstawą analiz porównawczych, wchodzi wszystkie zmienne centralne oraz zmienne izolowane [12]. Do przeprowadzenia dwuczynnikowej analizy wariancji w celu stwierdzenia, czy w badanym przedziale czasu zaszły statystycznie istotne zmiany w ramach analizowanych zmiennych i czy miała miejsce interakcja pomiędzy zmiennymi oraz parametrem czasu, z obliczonych statystyk opisowych w szczególności wykorzystano współczynnik zmienności, ponieważ jest on wyrażony w procentach i dlatego pozwala na porównywanie ze sobą wartości zmiennych wyrażonych w różnych jednostkach i miarach.

Ocenę rozwoju infrastruktury technicznej służącej zaopatrzeniu w wodę i sanitacji obszarów wiejskich w Polsce przeprowadzano na przykładzie szesnastu obiektów – województw, które zostały scharakteryzowane za pomocą dwudziestu pięciu cech o charakterze losowym (tab. 1).

## Analiza wyników

### Statystyki opisowe

W pierwszym kroku, mając wytypowane do badania 25 zmiennych z lat 2000, 2003, 2005 i 2012, oszacowano ich średnie arytmetyczne, odchylenia standardowe oraz wartości współczynników zmienności. Uzyskane wyniki przedstawiono w tabeli 2. Zamysłem autorów była analiza zachodzących w czasie zmian, tymczasem otrzymane w obrębie bardzo licznej i zróżnicowanej jakościowo liczby zmiennych wartości podstawowych statystyk z próby pozwalały na sformułowanie bardzo ogólnego poglądu na charakter zmian ilościowo-jakościowych. Analizowane zmienne miały różny charakter (różne miana – jednostki) i często zestawianie np. ich średnich wartości nie miało sensu. Porównanie staje się możliwe biorąc pod uwagę

współczynnik zmienności, który jest kombinacją (ilorazem) dwóch podstawowych parametrów wszelkich rozkładów (w przypadku których parametry te mają określone wartości liczbowe) – odchylenia standardowego i średniej. Wartości wyznaczonych współczynników zmienności wszystkich zmiennych okazały się spełniać postulowaną w statystyce wartość krytyczną (większą od 10% lub 15%).

### Analiza głównych składowych

Analiza głównych składowych była metodą, która posłużyła do określenia liczby postulowanych do badania czynników niezależnych. Podejmując decyzję na podstawie kryterium Kaisera (bo standaryzowano dane statystyczne) ustalono, że w latach 2000 i 2003 zmienne osiągnęły istotne ładunki w przypadku siedmiu pierwszych czynników. Pozostałe składowe, wraz z upływem czasu, wyjaśniały coraz mniej wariancji wszystkich zmiennych (w 2000 r. – 10,55%, a w 2003 r. – jedynie 8,9%). W 2005 r. liczba czynników została zredukowana do 6 głównych składowych, co skutkowało stratą 12,23% informacji zawartych w zmiennych wejściowych, a w 2012 r. – 12,17%. Podobne wyniki osiągnięto na podstawie kryterium osypiska.

W dalszej analizie, na podstawie macierzy współczynników korelacji pomiędzy zmiennymi charakteryzującymi poziom zaopatrzenia obszarów wiejskich województw w infrastrukturę służącą ochronie środowiska ustalono, które zmienne były istotnie skorelowane z wytypowanymi potencjalnymi składowymi głównymi. Przykładowo, główne składowe – szóste z lat 2000 i 2012 oraz siódma z 2003 r. cechowały wartości współczynników bliskie lub równe zero, dlatego składowe te uznano za wskazujące na niską świadomość środowiskową. W tabeli 3 przedstawiono nazwy głównych składowych i reprezentujące je zmienne.

### Metoda Hellwiga

Kolejnym zastosowanym narzędziem klasyfikacji była opisana wcześniej technika optymalnego wyboru predyktorów – metoda Hellwiga [12]. Pozwala ona wydzielić zmienne o znaczeniu centralnym (wiodącym), podporządkowane im zmienne satelitarne oraz zmienne izolowane (pojedyncze). W rezultacie następuje zmniejszenie analizowanego zbioru zmiennych i można przeprowadzić dwuczynnikową analizę wariancji z interakcją.

Wyznaczając jako bazowy rok badania najbardziej aktualny przedział czasu, klasyfikację metodą Hellwiga przeprowadzono na zmiennych z 2012 r. W efekcie uzyskano sześć zmiennych centralnych i siedemnaście opisujących je zmiennych satelitarnych oraz dwie zmienne izolowane. Ostatecznie do dalszych analiz zakwalifikowano następujące zmienne:  $X_4$ ,  $X_9$ ,  $X_{10}$ ,  $X_{17}$ ,  $X_{19}$ ,  $X_{20}$ ,  $X_{23}$  i  $X_{25}$ . Informacje zawarte w wytypowanych zmiennych zinterpretowano w następujący sposób:

- zmienna centralna  $X_4$  oraz zmienne satelitarne  $X_1$ ,  $X_2$ ,  $X_5$ ,  $X_6$ ,  $X_7$ ,  $X_8$ ,  $X_{11}$  i  $X_{16}$  tworzą grupę „zwodociągowanie i sanitacja obszarów wiejskich”,

- zmienna centralna  $X_9$  oraz satelitarna do niej zmienna  $X_{15}$  wygenerowały grupę „oczyszczalnie ścieków na obszarach wiejskich”,

- zmienną centralną  $X_{10}$  i reprezentującą ją zmienną satelitarną  $X_{22}$  nazwano „jakością infrastruktury służącej oczyszczaniu wody”,

- zmienna centralna  $X_{17}$  i zmienne satelitarne  $X_{12}$  i  $X_{13}$  stworzyły grupę „modernizacja urządzeń służących poprawie jakości wody”,

Tabela 2. Wybrane statystyki opisowe badanych zmiennych w latach 2000, 2003, 2005 i 2012  
 Table 2. The selected summary statistics for investigated variables in the years 2000, 2003, 2005 and 2012

Symbol zmiennnej	Nazwa statystyki												
	średnia				odchylenie standardowe								współczynnik zmienności
	2000 r.	2003 r.	2005 r.	2012 r.	2000 r.	2003 r.	2005 r.	2012 r.	2000 r.	2003 r.	2005 r.	2012 r.	
X <sub>1</sub>	65,59	74,79	79,19	93,6	42,28	42,77	41,73	47,18	0,64	0,57	0,53	0,50	
X <sub>2</sub>	2,08	2,66	1,58	2,0	1,09	2,41	0,68	1,25	0,53	0,91	0,43	0,62	
X <sub>3</sub>	0,54	0,57	0,64	0,7	0,20	0,21	0,21	0,23	0,37	0,36	0,34	0,32	
X <sub>4</sub>	11,40	16,61	12,07	19,5	4,36	12,50	4,83	9,30	0,38	0,75	0,40	0,48	
X <sub>5</sub>	0,51	0,94	1,23	2,4	0,45	0,82	1,03	1,82	0,88	0,87	0,84	0,76	
X <sub>6</sub>	8,76	15,51	20,49	37,6	7,41	12,72	15,94	27,99	0,85	0,82	0,78	0,75	
X <sub>7</sub>	1,60	2,87	2,24	3,3	1,68	2,63	1,92	3,01	1,05	0,92	0,86	0,91	
X <sub>8</sub>	22,72	40,01	35,67	77,2	16,58	32,81	27,37	93,07	0,73	0,82	0,77	1,21	
X <sub>9</sub>	9,49	3,37	6,68	2,5	3,86	2,69	7,49	1,77	0,41	0,80	1,12	0,70	
X <sub>10</sub>	49219,88	63180,75	70473,46	118708,00	30210,94	30154,23	32824,63	87316,50	0,61	0,48	0,47	0,74	
X <sub>11</sub>	7,49	14,38	9,53	12,4	4,70	29,24	7,54	11,81	0,63	2,03	0,79	0,95	
X <sub>12</sub>	2,53	3,63	2,92	4,8	2,14	3,04	2,37	2,88	0,84	0,84	0,81	0,60	
X <sub>13</sub>	2,53	3,63	2,92	4,8	2,14	3,04	2,37	2,88	0,84	0,84	0,81	0,60	
X <sub>14</sub>	16,75	46,68	43,46	52,2	16,73	28,78	23,58	30,20	1,00	0,62	0,54	0,58	
X <sub>15</sub>	1,03	0,74	0,89	2,3	2,70	1,04	1,07	1,45	2,63	1,40	1,20	0,62	
X <sub>16</sub>	5,24	28,79	4,42	9,0	6,64	97,82	3,55	3,04	1,27	3,40	0,80	0,34	
X <sub>17</sub>	2,01	2,43	2,50	4,2	1,17	1,47	1,22	2,22	0,58	0,61	0,49	0,53	
X <sub>18</sub>	1,16	1,82	1,06	1,2	1,31	1,34	1,21	1,24	1,12	0,74	1,14	1,00	
X <sub>19</sub>	0,17	0,20	0,20	0,2	0,11	0,15	0,16	0,13	0,64	0,76	0,79	0,69	
X <sub>20</sub>	55,75	60,50	66,75	70,5	29,41	26,77	20,35	19,52	0,53	0,44	0,30	0,28	
X <sub>21</sub>	1,54	1,81	1,89	5,2	0,90	0,93	1,14	2,89	0,58	0,51	0,60	0,56	
X <sub>22</sub>	2,01	2,43	2,50	4,2	1,17	1,47	1,22	2,22	0,58	0,61	0,49	0,53	
X <sub>23</sub>	55,75	60,50	66,75	70,5	29,41	26,77	20,35	19,52	0,53	0,44	0,30	0,28	
X <sub>24</sub>	2,02	2,30	2,64	3,6	0,52	0,86	1,10	1,49	0,25	0,37	0,42	0,42	
X <sub>25</sub>	1,82	5,70	0,60	2,8	2,00	20,52	0,73	6,21	1,10	3,60	1,22	2,25	

Tabela 3. Główne składowe i reprezentujące je zmienne w latach 2000, 2003, 2005 i 2012  
 Table 3. Principal components and their variables in the years 2000, 2003, 2005 and 2012

Główna składowa	Zmienne reprezentujące główną składową							
	2000 r.		2003 r.		2005 r.		2012 r.	
	numer czynnika	symbol zmiennej	numer czynnika	symbol zmiennej	numer czynnika	symbol zmiennej	numer czynnika	symbol zmiennej
Dostępność infrastruktury	1	X <sub>1</sub> , X <sub>2</sub> X <sub>4</sub> , X <sub>5</sub> X <sub>6</sub> , X <sub>7</sub> X <sub>8</sub> , X <sub>10</sub> X <sub>11</sub> , X <sub>25</sub> X <sub>17</sub> , X <sub>18</sub> , X <sub>25</sub>	2	X <sub>1</sub> , X <sub>4</sub> X <sub>5</sub> , X <sub>6</sub> X <sub>7</sub> , X <sub>8</sub> X <sub>15</sub> , X <sub>25</sub>	1	X <sub>1</sub> , X <sub>2</sub> X <sub>5</sub> , X <sub>6</sub> X <sub>7</sub> , X <sub>8</sub> X <sub>10</sub> , X <sub>11</sub> X <sub>15</sub> , X <sub>25</sub>	1	X <sub>1</sub> , X <sub>2</sub> X <sub>4</sub> , X <sub>5</sub> X <sub>6</sub> , X <sub>7</sub> X <sub>8</sub> , X <sub>11</sub> X <sub>17</sub> , X <sub>18</sub>
Wysoka świadomość znaczenia jakości stosowanych rozwiązań	3	X <sub>19</sub> , X <sub>20</sub> , X <sub>23</sub>	1	X <sub>2</sub> , X <sub>10</sub> X <sub>11</sub> , X <sub>14</sub> X <sub>16</sub> , X <sub>18</sub> X <sub>20</sub> , X <sub>23</sub>	6	X <sub>4</sub>	2	X <sub>14</sub> , X <sub>16</sub> X <sub>19</sub> , X <sub>21</sub> X <sub>24</sub>
Innowacje	2	X <sub>12</sub> , X <sub>13</sub> X <sub>22</sub> , X <sub>24</sub>	5	X <sub>12</sub> , X <sub>13</sub>	2	X <sub>3</sub> , X <sub>9</sub> X <sub>12</sub> , X <sub>13</sub> X <sub>14</sub> , X <sub>16</sub> X <sub>19</sub> , X <sub>21</sub> X <sub>22</sub> , X <sub>24</sub> *	3	X <sub>9</sub> , X <sub>12</sub> X <sub>13</sub> , X <sub>22</sub>
Nakłady na inwestycje	7	X <sub>14</sub>	6	X <sub>21</sub>	3	X <sub>20</sub> , X <sub>23</sub>	4	X <sub>23</sub> , X <sub>25</sub>
Istniejąca infrastruktura	5	X <sub>3</sub>	3	X <sub>3</sub> , X <sub>9</sub> X <sub>19</sub> , X <sub>24</sub>	5	X <sub>18</sub>	5	X <sub>3</sub> , X <sub>10</sub> X <sub>15</sub>
Niska świadomość środowiskowa (wartości współczynników bliskie lub równe zero)	6	X <sub>1</sub> , X <sub>5</sub> X <sub>7</sub> , X <sub>8</sub> X <sub>9</sub> , X <sub>11</sub> X <sub>25</sub>	7	X <sub>5</sub> , X <sub>7</sub> X <sub>8</sub> , X <sub>11</sub> X <sub>12</sub> , X <sub>13</sub> X <sub>15</sub> , X <sub>16</sub> X <sub>21</sub> , X <sub>25</sub>	–	–	6	X <sub>1</sub> , X <sub>3</sub> X <sub>4</sub> , X <sub>12</sub> X <sub>13</sub> , X <sub>17</sub> X <sub>18</sub> , X <sub>20</sub> X <sub>25</sub>
Wysoka świadomość środowiskowa	4	X <sub>9</sub> , X <sub>21</sub>	4	X <sub>17</sub> , X <sub>22</sub>	4	X <sub>17</sub>	–	–

\*w miarę upływu czasu, ze względu na możliwość stosowania nowoczesnych rozwiązań technologicznych, maleje liczba małych składowisk odpadów na rzecz składowisk o dużej powierzchni

– zmienną X<sub>19</sub> reprezentującą zmienne X<sub>3</sub>, X<sub>14</sub>, X<sub>21</sub> i X<sub>24</sub> nazwano łącznie „uzbrojenie obszarów wiejskich w infrastrukturę służącą ochronie środowiska”,

– zmienna X<sub>20</sub> i zależna od niej zmienna X<sub>18</sub> łącznie tworzą grupę „stacje oczyszczania wody”,

– zmienna izolowana X<sub>23</sub> określa „stosunek nakładów na modernizację do nakładów na nowe inwestycje w zakresie sanitacji i oczyszczania wody”,

– zmienna izolowana X<sub>25</sub> określa „nakłady na nowe inwestycje w zakresie gminnych składowisk odpadów”.

### Dwuczynnikowa analiza wariancji z interakcją zmiennych centralnych w typowanych metodą Hellwiga

W celu zbadania, czy istnieje statystycznie istotny związek pomiędzy czasem (czynnik A) a zmiennymi (czynnik B), na podstawie wyznaczonych metodą Hellwiga zmiennych wiodących oraz otrzymanych wartości współczynnika zmienności z próby przeprowadzono dwuczynnikową analizę wariancji z interakcją. Zarówno z naukowego, jak i praktycznego punktu widzenia interesujące było porównanie danych ze skrajnych przedziałów czasu badania. W rezultacie biorąc pod uwagę czynnik czasu wygenerowano dwa poziomy uwzględniające lata 2000 i 2005 oraz 2010 (dotychczas w analizie danych rok 2010 nie był brany pod

uwagę. W analizie wariancji zostały wykorzystane lata 2010 i 2012 (gdyż wedle założeń model analizy wariancji z interakcją wymaga w jednej celi minimum dwóch obserwacji), którym przyporządkowano indeksy 1 i 2. W przypadku czynnika związanego ze zmiennymi, wykorzystanie zmiennych wiodących (wyznaczonych dzięki zastosowaniu metody Hellwiga) pozwoliło wydzielić osiem poziomów, które w kolejności występowania otrzymały indeksy od 1 do 8.

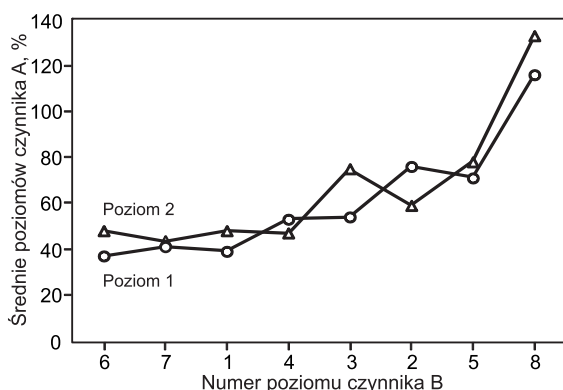
Centralnym punktem dwuczynnikowej analizy wariancji z interakcją było udzielenie odpowiedzi na pytanie, czy parametr czasu miał istotny, w sensie statystycznym, wpływ na zachowanie się (zmianę wartości zmienności w czasie) całego badanego zbioru zmiennych? Otrzymana w modelu interakcji wartość statystyki F-Snedecora z próby wyniosła 7,81. Z kolei odczytana z tablic rozkładu F-Snedecora wartość graniczna statystyki F\* przy przyjętym poziomie istotności  $\alpha=0,05$  oraz 16 i 7 stopniach swobody przyjęła wartość 3,49. Zatem pomiędzy analizowanymi czynnikami występuje wyraźna interakcja (3,49 jest wyraźnie mniejsze od wartości 7,81). Jeśli zachodzi interakcja pomiędzy badanymi czynnikami, wówczas nie przeprowadza się osobnej analizy statystycznej odnośnie równości średnich w poszczególnych czynnikach, gdyż wystąpienie interakcji świadczy o tym, że te średnie w przypadku obu czynników są różne [13].

Tabela 4. Wartości średnie współczynników zmienności dwuczynnikowej analizy wariancji  
Table 4. Average values of coefficients of variation of two-way analysis of variance (ANOVA)

A/B	1	2	3	4	5	6	7	8	$\bar{x}_i$
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)	(1,7)	(1,8)	–
$\bar{x}_{1j}$	39	76	54	53	71	37	41	116	61
2	(2,1)	(2,2)	(2,23)	(2,4)	(2,5)	(2,6)	(2,7)	(2,8)	–
$\bar{x}_{2j}$	48	59	75	47	78	48	43	133	67
$\bar{x}_j$	48	68	64	50	74	42	42	124	–

Chcąc określić, jak zmieniały się wartości średnich współczynników zmienności, pogłębiając zakres prowadzonych analiz wyznaczono średnie w grupach (na poziomach czynników, reprezentują je symbole  $\bar{x}_i$  oraz  $\bar{x}_j$ ) i podgrupach (w miejscu krzyżowania się poziomów obu czynników –  $\bar{x}_{ij}$ ). Wyniki obliczeń przedstawiono w tabeli 4. W przypadku drugiego i czwartego poziomu czynnika B wartości średnich uległy zmniejszeniu. W pozostałych przypadkach obserwowano wzrost poziomu średnich.

Aby lepiej zobrazować wykazaną analitycznie interakcję, na podstawie danych zawartych w tabeli 4 wykonano rysunek 1. W tym celu średnie poziomów w przypadku czynnika B uporządkowano według wartości niemalejących (od najmniejszej do największej) i w rezultacie otrzymano następującą kolejność poszczególnych poziomów: 6, 7, 1, 4, 3, 2, 5, 8. Na osi odciętych odłożono w równych odległościach punkty, które oznaczono numerami poziomów w takiej kolejności, w jakiej uporządkowały się wartości średnie. Następnie na osi rzędnych naniesiono wartości średnie poziomów 1 i 2 czynnika A względem czynnika B (są to średnie w tzw. celach, np. w przypadku (1,1)  $\bar{x}_{11}=39$ ). W ten sposób otrzymano dwie linie łamane, których przecięcie się świadczy o istnieniu interakcji (gdyby linie łamane były chociaż w przybliżeniu równoległe, oznaczałoby to brak interakcji).



Rys. 1. Ilustracja interakcji między czynnikami A i B

Fig. 1. The graphical illustration of interaction between A and B factors

## Dyskusja wyników

Cel poznawczy pracy – polegający na tym, aby zdiagnozować w wymiarze ogólnokrajowym poziom rozwoju systemów zaopatrzenia w wodę i sanitacji obszarów wiejskich w Polsce oraz określić występujące w czasie tendencje

– osiągnięto, stosując wielowymiarowe metody analizy danych. Logiczną implikacją zastosowania w tym badaniu metod statystycznych było:

- zmniejszenie znacznej ilości informacji o badanych obiektach do zaledwie kilku bazowych kategorii, które poddano dokładniejszej analizie statystycznej i treściowej,
- otrzymanie jednorodnych grup cech pod względem ich właściwości, co w konsekwencji doprowadziło do znalezienia (nazwania) ich podstawowych właściwości, czyli wyrażenia ich w nowym i liczbowo mniejszym wymiarze,
- skupienie uwagi na najbardziej typowych w rozważanym problemie badawczym zjawiskach, kategoriach, cechach itp.,
- oraz najważniejsze, określenie i wyjaśnienie struktury powiązań między cechami badanych obiektów (obiektami badanych cech były poszczególne województwa).

Centrum dociekań artykułu w zasadzie stanowiła analiza mająca doprowadzić do przydzielenia zmiennych do stosunkowo jednorodnych wewnętrznie grup. Można odnieść wrażenie, że zastosowane w pracy narzędzia statystyczne nie wykazują jednoznaczności w podziale. Tymczasem klasyfikacja nie musi być jednoznaczna, gdyż statystyka – jako nauka – nie rozwiązuje analizowanych problemów w sposób deterministyczny. Istotą tej dziedziny jest niejednoznaczność, przy czym wskazana niejednoznaczność musi spełniać podstawowy warunek statystyki – powinna być statystycznie istotna lub nieistotna. Dlatego statystyczne metody analizy obiektów mają jednak siłę dyskryminacyjną (statystycznie istotną) wydzielenia obiektów zachowujących się w sposób podobny, a ponadto zmieniających swoją przynależność do grup i klas w miarę upływu czasu, który to czynnik był jednym z podstawowych w wykonanej analizie statystyczno-merytorycznej.

Przeprowadzona analiza statystyczna objęła zatem pierwszy rok po wprowadzeniu reformy terytorialnej kraju i jej ostatni dostępny statystycznie okres – rok 2012. Wyniki otrzymane z analizy wariancji (model z interakcją) dają podstawy do stwierdzenia, że istnieje wyraźny związek (interakcja) pomiędzy naturą istotnych statystycznie zmiennych (centralnych) a parametrem czasu. Na podstawie analizy głównych składowych nasuwa się wniosek, że na przestrzeni pierwszej dekady XXI wieku zmienne ujednoliciły się – dążyły do jednolitej struktury – co oznacza, że infrastruktura techniczna służąca ochronie środowiska na obszarach wiejskich w Polsce, w miarę upływu czasu, wykazuje coraz mniejsze zróżnicowanie przestrzenne pomiędzy województwami, czyli w rezultacie zjawisko ulega uproszczeniu. Niezmiennie najważniejszymi składowymi

głównymi od 2000 r. pozostają zarówno dostępność infrastruktury technicznej w zakresie zaopatrzenia w wodę oraz sanitacji, jak i wysoka świadomość znaczenia jakości stosowanych rozwiązań (innowacje).

Analiza zachodzących w czasie zmian na poziomie stosowanej w zakresie wodociągów i sanitacji obszarów wiejskich wymiaru technologii wskazuje na zmianę nastawienia w skali kraju do znaczenia tego problemu. Na podstawie przeprowadzonej analizy można stwierdzić, że wraz z upływem czasu wskaźniki określające infrastrukturę techniczną służącą ochronie środowiska na obszarach wiejskich w Polsce ulegają stałemu rozwojowi, czyli w sposób wyraźny kształtuje się świadomość w zakresie ochrony środowiska. Świadczą o tym wartości średnich (trend liniowy) w metodzie Hellwiga w odniesieniu do zmiennych centralnych, których łączna wartość, w przypadku ośmiu wyznaczonych poziomów zmiennych, wzrosła z 61 w okresie początkowym do 67 w najbardziej aktualnym czasie badania. Wprawdzie w przypadku drugiego i czwartego poziomu czynnika B odnotowano spadek wartości średnich, jednak w tym wypadku należy przyjąć interpretację, iż następuje stopniowe nasylenie obszarów wiejskich nowymi i zmodernizowanymi urządzeniami służącymi poprawie jakości zasobów wodnych.

## Podsumowanie

Wyniki przeprowadzonej analizy statystycznej dotyczącej badania stopnia rozwoju infrastruktury technicznej związanej z zaopatrzeniem w wodę i sanitacją obszarów wiejskich w Polsce dostarczają decydom znaczącą wiedzę ilościową i merytoryczną, niezbędną do podejmowania decyzji o znaczeniu ogólnokrajowym. Umożliwiają one kompleksowe spojrzenie na państwo jako system i na bieżącą ocenę sytuacji w zakresie zaopatrzenia w wodę oraz odprowadzania ścieków i zagospodarowania odpadów na danym obszarze. Łączy się to ściśle także z podstawowymi funkcjami zarządzania [14, 15] w zakresie podejmowania decyzji na szczeblu centralnym (np. przez Ministerstwo Ochrony Środowiska czy Ministerstwo Rolnictwa i Rozwoju Wsi). Ponieważ z naukowego punktu widzenia interesujące jest również, jak w badanym czasie zmieniał się rozwój analizowanych cech w ramach poszczególnych województw, dlatego celowe jest przeprowadzenie analogicznej analizy statystycznej na płaszczyźnie wojewódzkiej.

**Skowronek-Grądział, A., Kołwzan, W., Dziubek, A.M. Multivariate Statistical Analysis of Development in Water Supply and Sanitation in Rural Areas in Poland in the Period 2000–2012. *Ochrona Środowiska* 2014, Vol. 36, No. 4, pp. 61–68.**

**Abstract:** A research hypothesis was formulated that since the beginning of the 21<sup>st</sup> century a systematic progress had been made in technical infrastructure serving water supply and sanitation of rural areas in Poland. Concurrently, it was assumed that multivariate methods of statistical analysis could help achieve the set goal. The analysis of 25 variables from the years 2000, 2003, 2005 and 2012 allowed for the estimation of their arithmetic means, variances, standard deviations and coefficients of variation. In terms of water supply and sewerage systems, the statistical analysis demonstrated a steady development of indicators for technical infrastructure in rural areas in Poland. Thus, awareness in respect of the need for water supply protection

## LITERATURA

1. I. BODIK, P. RIDDERSTOLPE [Eds.]: Sustainable Sanitation in Central and Eastern Europe – addressing the needs of small and medium-size settlements. Global Water Partnership Central and Eastern Europe, Nitra (Slovakia) 2007.
2. A.M. DZIUBEK: Some of the problems concomitant with the removal of nitrogen compounds in Poland. *Environment Protection Engineering* 2012, Vol. 38, No. 4, pp. 163–169.
3. E. NOWAK: Metody taksonomiczne w klasyfikacji obiektów społeczno-gospodarczych. Polskie Wydawnictwo Ekonomiczne, Warszawa 1990.
4. Z. JANKOWSKA [red.]: Sprawozdanie z realizacji inwestycji w zakresie wodociągów i sanitacji wsi w roku 2000, 2003, 2005, 2010, 2012. Ministerstwo Rolnictwa i Rozwoju Wsi, Warszawa 2000, 2003, 2005, 2010, 2012.
5. W. OSTASIEWICZ: Myślenie statystyczne. Wolters Kluwer, Warszawa 2012.
6. T. PANEK: Statystyczne metody wielowymiarowej analizy porównawczej. Szkoła Główna Handlowa, Warszawa 2009.
7. J.F. HAIR: Multivariate Data Analysis. Upper Saddle River. Pearson Prentice Hall, New York 2006.
8. M. SZTEMBERG-LEWANDOWSKA: Analiza czynnikowa w badaniach marketingowych. Uniwersytet Ekonomiczny, Wrocław 2008.
9. A. ZELIĄS: Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym. Akademia Ekonomiczna, Kraków 2000.
10. T. GALANC, W. KOŁWZAN, J. PIERONEK: A quantitative method supporting the management of the teaching process. *Operations Research and Decisions* 2011, Vol. 21, No. 1, pp. 33–51.
11. A. SKOWRONEK-GRĄDZIAL, W. KOŁWZAN: Zastosowanie metody głównych składowych do analizy obszarów wiejskich w zakresie infrastruktury służącej ochronie środowiska. Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2013.
12. Z. HELLWIG: Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę wykwalifikowanych kadr. *Przegląd Statystyczny* 1968, nr 4, ss. 307–326.
13. C. PLATT: Problemy rachunku prawdopodobieństwa i statystyki matematycznej. PWN, Warszawa 1981.
14. F.E. KAST, J.E. ROSENZWEIG: Organization and Management: A System and Contingency Approach. McGraw-Hill, New York 1979.
15. S. RAMZAN, M.I. KHAN, F.M. ZAHID, S. RASUL, S. RAFIG: Socio-economic differences among districts of the Punjab: A cluster analysis approach based on multiple indicator cluster survey. *World Applied Sciences Journal* 2013, Vol. 21, No. 2, pp. 293–300.

clearly increases. The above conclusion is supported by the means in the Hellwig's method (linear trend) in relation to the central variables. The combine value of the latter, calculated for the eight sets of variables, increased from 61 in the initial period (year 2000) to 67 in the most recent study period (2012). The results received provide the decision-makers with extensive quantitative and qualitative knowledge, essential for taking decisions at the national level. In addition, they allow for comprehensive approach and on-going evaluation of the situation with regard to water supply as well as sewage disposal and waste management in any given area. This is strictly connected with currently understood fundamental functions of management with regard to decision-making at the central level (e.g. by the Ministry of the Environment).

**Keywords:** Water supply system, sewerage system, landfill, multivariate analysis, ANOVA, Hellwig's method, correlation analysis.