



**Uniwersytet im. Adama Mickiewicza w Poznaniu**

Wydział Nauk Geograficznych i Geologicznych

Instytut Geografii Fizycznej i Kształtowania Środowiska Przyrodniczego

**mgr Lech Kaczmarek**

**Potencjał informacyjny  
krajowych baz danych przestrzennych  
w kartograficznych badaniach  
środowiska przyrodniczego**

**Rozprawa doktorska**

napisana w Zakładzie Kartografii i Geomatyki UAM w Poznaniu

pod kierunkiem **prof. UAM dr hab. Beaty Medyńskiej-Gulij**

**Poznań 2011**

Szczególne podziękowania składam Pani Profesor Beacie Medyńskiej-Gulij  
za wnikliwe wsparcie merytoryczne podczas przygotowywania rozprawy

Ponadto wszystkim moim Najbliższym i Znajomym dziękuję za niezwykle  
wsparcie duchowe, które towarzyszyło powstaniu tej rozprawy

W szczególności dziękuję mojej Żonie Ewelinie  
za umacnianie mojej woli do pracy badawczej  
i Dzieciom Basi, Stasiowi i Tosi  
za twórcze wykorzystanie próbnych wydruków

*Rozprawę doktorską dedykuję Mojej Mamie  
w podziękowaniu za podjęty trud wychowania*

## Spis treści

<b>1. WSTĘP .....</b>	<b>5</b>
<b>1.1. Cel główny i cele cząstkowe .....</b>	<b>6</b>
<b>1.2. Przedmiot i obszar badań .....</b>	<b>7</b>
<b>1.3. Metody badawcze .....</b>	<b>8</b>
<b>1.4. Przebieg postępowania badawczego .....</b>	<b>10</b>
<b>2. PODSTAWY TEORETYCZNE MODELOWANIA     KARTOGRAFICZNEGO .....</b>	<b>12</b>
<b>2.1. Poglądy na metodykę kartograficznych badań środowiska         przyrodniczego .....</b>	<b>12</b>
<b>2.2. Modelowanie środowiska przyrodniczego w ujęciu cyfrowym .....</b>	<b>13</b>
2.2.1. Modelowanie kartograficzne a modelowanie danych geograficznych .....	13
2.2.2. Reprezentacja obiektów przestrzennych .....	19
2.2.3. Generalizacja kartograficzna i konceptualna .....	21
<b>2.3. Czynniki wpływające na jakość danych przestrzennych .....</b>	<b>23</b>
<b>2.4. Koncepcje skal pomiarowych w kartografii .....</b>	<b>27</b>
<b>2.5. Uwarunkowania tworzenia krajowych baz danych przestrzennych .....</b>	<b>31</b>
<b>2.6. Problemy badawcze do rozwiązania .....</b>	<b>33</b>
2.6.1. Dane cyfrowe jako źródło informacji o środowisku przyrodniczym .....	34
2.6.2. Jakość danych przestrzennych .....	35
2.6.3. Skale pomiarowe w ujęciu danych cyfrowych .....	35
2.6.4. Potencjał informacyjny baz danych .....	36
<b>3. ANALIZA ZAWARTOŚCI BAZ DANYCH     PRZESTRZENNYCH W KONTEKŚCIE MODELOWANIA     RZECZYWISTOŚCI .....</b>	<b>37</b>
<b>3.1. Założenia dotyczące analizy zawartości merytorycznej baz danych         przestrzennych .....</b>	<b>38</b>
<b>3.2. Zawartość merytoryczna baz danych przestrzennych w ujęciu modelu         danych .....</b>	<b>40</b>
<b>3.3. Analiza porównawcza danych o środowisku przyrodniczym w bazach ...</b>	<b>68</b>

<b>4. BADANIA JAKOŚCI DANYCH PRZESTRZENNYCH .....</b>	<b>74</b>
<b>4.1. Zakres badania jakości danych .....</b>	<b>74</b>
<b>4.2. Miara jakości danych .....</b>	<b>75</b>
<b>4.3. Wstępna ocena jakości baz danych .....</b>	<b>75</b>
<b>4.4. Badanie jakości wybranych danych przestrzennych .....</b>	<b>77</b>
4.4.1. Badanie jakości danych o wodach płynących .....	79
4.4.2. Badanie jakości danych o wodach stojących .....	89
<b>5. WYZNACZANIE I WYKORZYSTANIE POTENCJAŁU INFORMACYJNEGO DANYCH PRZESTRZENNYCH .....</b>	<b>98</b>
<b>5.1. Koncepcja potencjału informacyjnego danych przestrzennych .....</b>	<b>98</b>
<b>5.2. Poziomy ujęcia atrybutów .....</b>	<b>100</b>
<b>5.3. Pomiar potencjału informacyjnego danych przestrzennych – przykłady</b>	<b>104</b>
5.3.1. Pomiar potencjału informacyjnego dla wód płynących .....	106
5.3.2. Pomiar potencjału informacyjnego dla wód stojących .....	109
<b>5.4. Dynamiczny potencjał informacyjny .....</b>	<b>113</b>
<b>5.5. Dualizm pojęcia potencjału informacyjnego danych przestrzennych .....</b>	<b>123</b>
<b>6. PODSUMOWANIE .....</b>	<b>125</b>
Literatura .....	129
Spis rycin i tabel .....	136

## 1. WSTĘP

Mapa od zarania stanowi źródło informacji o przestrzeni geograficznej mimo ciągłych dyskusji związanych z jej istotą (Robinson i in. 1995, Makowski 2005a). Jej treść niesie nieocenioną wiedzę o rozmieszczeniu, charakterystyce i wzajemnych relacjach składowych środowiska przyrodniczego (Makowski 2005b). Saliszczew (1955) wprowadził pojęcie „kartograficznej metody badań”, zaliczając do niej wypracowywane już od końca XIX wieku metody analiz map. Do dziś jest to jedna z najważniejszych gałęzi kartografii. Metody analiz map były opisywane przez wielu autorów, m.in. Berlanta (1978), Rataskiego (1989) i Saliszczewa (1998).

Pojawienie się cyfrowego zapisu informacji kartograficznej zapoczątkowało analityczne podejście do badania zjawisk przestrzennych. Zainicjowana przez Toblera (1959, 1961) kartografia analityczna przeniosła warsztat badawczy kartografa na systemy komputerowe, a tradycyjne metody analizy map pochodzące z kartograficznej metody badań odniesiono do transformacji danych (Clarke 1995). Wyniki tak rozumianej transformacji źródeł danych są przedmiotem kompilowania kartograficznego (Dent 1990, 1999, Medyńska-Gulij 2007). Modelowanie rzeczywistości (Board 1967, Saliszczew 1967) ujęto w postaci krajobrazowego i kartograficznego modelu danych (Digital Landscape Model - DLM, Digital Cartographic Model - DCM; Kraak, Ormelling 1998). Ten pierwszy stanowi zbiór danych odnoszących się do wybranego obszaru, stąd często nazywany jest modelem topograficznym (Głazewski 2006). Cyfrowy model kartograficzny (DCM) powstaje w wyniku transformacji kartograficznej cyfrowego modelu krajobrazu.

Idące za rozwojem kartografii analitycznej nowe technologie zmieniły kartografię polską w ciągu ostatniego ćwierćwiecza w znaczący sposób (Czerny i in. 1999). Upowszechniająca się w naszym kraju już od lat 90. XX wieku komputerowa redakcja map stała się okazją do podjęcia decyzji o gromadzeniu zasobu mapowego w dwóch wersjach, cyfrowej i analogowej. Powstałe w tym okresie instrukcje i wytyczne techniczne zobligowały wykonawców map analogowych do budowania cyfrowych baz danych przestrzennych, które stawały się integralną częścią opracowań kartograficznych. W taki sposób powstawały mapy geologiczne (hydrogeologiczna, geologiczno-gospodarcza, geologiczna) oraz sozologiczne i hydrograficzne. To był początek okresu tworzenia infrastruktury informacji przestrzennej Polsce. Kolejny silny impuls w rozwoju danych przestrzennych przyniosła Dyrektywa INSPIRE (dyrektywa ustanawiająca infrastrukturę informacji przestrzennej we Wspólnocie Europejskiej nr 2007/2/WE z dnia 14 marca 2007 r.) oraz Ustawa o infrastrukturze informacji

przestrzennej (Dz. U. nr 76, poz. 489 z dnia 4 marca 2010 r.). Załączniki I, II i III do tych aktów prawnych zdefiniowały zakresy tematyczne niezbędnych danych, a same przepisy zobligowały do tworzenia uspołnionych infrastruktur informacji przestrzennej.

Istnienie baz danych przestrzennych pociągnęło za sobą wdrażanie nowych narzędzi badawczych w zakresie szeroko rozumianych badań środowiska przyrodniczego. Wykorzystanie technik geoinformacyjnych w kartograficznej metodzie badań otworzyło szereg nowych możliwości. Jak zauważył Zwoliński (2009) dynamiczny rozwój geoinformacji w drugiej połowie lat 90. XX wieku zmienił proporcje czasu poświęconego na pozyskiwanie danych i ich integrację a analizy przestrzenne na korzyść tych drugich. Przyczyniło się do tego między innymi masowe powstawanie baz danych przestrzennych. Nowe perspektywy badawcze w tym zakresie są jednak uwarunkowane faktycznym poziomem zasobności baz danych w informację przestrzenną oraz jakością tych danych. Konieczne staje się więc wdrożenie procedur oceny przydatności baz danych przestrzennych w konkretnym postępowaniu badawczym, które można określić jako potencjał informacyjny. **Potencjał informacyjny danych przestrzennych to miara zasobności modeli obiektów przestrzennych (encji) w informacje o cechach tych obiektów, istotnych z punktu widzenia użytkownika danych.**

### 1.1. Cel główny i cele cząstkowe

Głównym celem rozprawy jest **wskazanie kluczowej roli wyznaczania potencjału informacyjnego krajowych baz danych przestrzennych do badań środowiska przyrodniczego z wykorzystaniem kartograficznej metody badań.** Potencjał informacyjny odnosi się do najmniejszej jednostki generowania danych w bazach cyfrowych – do warstwy tematycznej.

**Cele cząstkowe** rozprawy, składające się na realizację celu głównego:

1. Analiza zawartości merytorycznej krajowych baz danych przestrzennych w zakresie danych o środowisku przyrodniczym.
2. Ocena jakości wybranych warstw baz danych przestrzennych w zakresie dokładności geometrycznej i lokalizacyjnej oraz kompletności.
3. Klasyfikacja funkcjonalna poziomów ujęcia atrybutów na potrzeby wyznaczania kompleksowego wskaźnika dokładności tematycznej danych przestrzennych.
4. Określenie teoretycznych podstaw wyznaczania potencjału informacyjnego.
5. Wyznaczanie wartości potencjału informacyjnego wybranych danych przestrzennych w oparciu o zaproponowany schemat postępowania.
6. Zdefiniowanie pojęć statycznego i dynamicznego potencjału informacyjnego.

**Celem użytkowym** niniejszej rozprawy stało się zwrócenie uwagi na przydatność krajowych baz danych przestrzennych w badaniach środowiska przyrodniczego oraz wskazanie metod kwantyfikacji miar potencjału informacyjnego danych.

## 1.2. Przedmiot i obszar badań

Przedmiotem zainteresowań badawczych w niniejszej rozprawie jest zasób cyfrowych baz danych przestrzennych dostępnych w Polsce (tab. 1.1.).

**Tabela 1.1. Zestawienie wybranych baz danych przestrzennych**

*Źródło: opracowanie własne*

Nazwa	Skrót stosowany w opracowaniu	Skala nominalna Układ współrzędnych
Baza Danych Ogólnogeograficznych	BDO	1:250 000 PUWG 1992
Mapa Wektorowa Poziomu Drugiego	VML2	1:50 000 WGS84
Baza Danych Topograficznych (Baza Danych Obiektów Topograficznych)	TBD (BDOT)	1:10 000 PUWG 1992
Mapa Hydrograficzna Polski	MHP	1:50 000 PUWG 1992
Mapa Sozologiczna Polski	MSP	1:50 000 PUWG 1992
Szczegółowa Mapa Geologiczna Polski	SMGP	1:50 000 PUW 1942
Mapa Hydrogeologiczna Polski (główny poziom wodonośny)	MHgP	1:50 000 PUW 1942
Mapa Geośrodowiskowa Polski	MGP	1:50 000 PUW 1942
Mapa Podziału Hydrograficznego Polski	MPHP	1:50 000 PUWG 1992
Mapa Glebowo-Rolnicza	MGR	1:25 000 (1:5 000) PUWG1992
Ewidencja Gruntów i Budynków	EGiB	1:5 000 PUWG 2000
Leśna Mapa Numeryczna i System Informatyczny Lasów Państwowych	LMN, SILP	1:5 000 PUWG1992
Numeryczny model terenu z Mapy Wektorowej Poziomu Drugiego	DTED2	1:50 000 WGS84
Baza Danych Pokrycia Terenu Corine Land Cover 2006	CLC2006	1:100 000 PUWG1992

Dysponentami tych baz danych są instytucje rządowe (Główny Urząd Geodezji i Kartografii, Ministerstwo Środowiska), samorządowe (Wojewódzkie Ośrodki Dokumentacji Geodezyjnej i Kartograficznej) i branżowe (np. Państwowy Instytut Geologiczny, Instytut Upraw i Nawożenia w Puławach).

Obszar badań ograniczono do województwa wielkopolskiego, głównie ze względu na dostęp do danych źródłowych. Ponadto wydaje się, że w sytuacji istnienia ogólnokrajowych instrukcji i wytycznych technicznych opracowywania większości z rozpatrywanych baz danych, obszar województwa wielkopolskiego stanowić będzie próbę reprezentatywną dla kraju. Przy rozpatrywaniu szczegółowych miar jakości danych badania przeprowadzono na wybranych polach testowych, np. wybranych arkuszach mapy topograficznej 1:10 000 w układzie współrzędnych 1992.

### **1.3. Metody badawcze**

Przeprowadzenie prac badawczych poprzedziła analiza stanu wiedzy z zakresu modelowania danych przestrzennych, badania jakości danych, koncepcji poziomów pomiarowych w kartografii oraz ewolucji polskich baz danych przestrzennych. Pozwoliła ona na zestawienie poglądów naukowych i wskazanie problemów do rozwiązania.

Głównym założeniem niniejszej rozprawy stało się wykorzystanie zasobu cyfrowych baz danych przestrzennych w kartograficznej metodzie badań środowiska przyrodniczego (Saliszczew 1955, 1998; Ratajski 1973, 1989; Berlant 1978; Żyszkowska 2003; Paślawski, Siwek 2010). Ponieważ metoda kartograficzna odnosi się w niniejszej rozprawie do materiału cyfrowego, więc towarzyszą jej metody przetwarzania geoinformacyjnego (Tobler 1959, Clarke 1995, O`Sullivan, Unwin 2003, Chang 2004, Longley i in 2006, Urbański 2008, Goodchild 2008, Gomarasca 2009). Badanie jakości danych przestrzennych odnoszono do stanu wiedzy (Zhang, Goodchild 2002, Bernhardsen 2002, Bielecka 2006), jak i do założeń w normach jakości (US Federal Geographic Data Committee, Spatial Data Transfer Standard, International Standards Organization). W szczególności rozprawa odnosi się do metodycznych rozwiązań badania jakości danych opisanych w normach ISO (PN-EN ISO 19113:2005, PN-EN ISO 19114:2005 i ISO / AVI 19138:2002).

Wszystkie metody badawcze, które zastosowano w niniejszej rozprawie można umieścić w pięciu grupach:

1. **Analiza poglądów badawczych** – polegała na zestawieniu koncepcji naukowych w podziale na merytoryczne zagadnienia odnoszące się do problematyki pracy. Analiza została uzupełniona zestawieniami syntetyzującymi poszczególne zagadnienia takie, jak modelowanie danych przestrzennych (Board 1967, Saliszczew 1967, Ratajski 1973, 1978, Peuquet 1984, Kraak, Ormelling 1998, Zeiler 1999, Żyszkowska 2000, Brodlić i in. 2001, Hake 2002, Bernhardsen 2002, Longley i in. 2006, Ostrowski 2008), koncepcje poziomów pomiarowych w kartografii (Bertin 1983, Frączek 1981,

Żyszkowska 2000, Bernhardsen 2002, Korycka-Skorupa 2002a, 2002b, Chang 2004, Bielecka 2006, Ostrowski 2008) czy poglądy na czynniki wpływające na jakość danych przestrzennych (Zhang, Goodchild 2002, Bielecka 2006, Longley i in. 2006, Gomasasca 2009).

2. **Metody geoinformacyjne** – wykorzystano szereg operacji geoinformacyjnych pozwalających na analizę jakości danych przestrzennych dla dużego obszaru. Zastosowano m.in. nakładanie na wybrane warstwy tematyczne siatki regularnych kwadratów 10 x 10 km, generowanie wybranych parametrów z geometrii obiektów wektorowych w obrębie kwadratów regularnej siatki oraz dla całych warstw wektorowych, generowanie wybranych wartości atrybutów obiektów wektorowych w siatce kwadratów, obliczanie wskaźników geometrycznych obiektów, generowanie wskaźników błędów położenia obiektów i szereg innych.
3. **Metody statystyczne** - zastosowano proste miary statystyczne wartości wygenerowanych z wykorzystaniem metod geoinformacyjnych: sumy, średnie arytmetyczne, liczebność obiektów itp.
4. **Metody wizualizacji kartograficznej** – w celu wizualizacji przestrzennego rozkładu wskaźników jakości danych zastosowano podstawowe formy prezentacji kartograficznej – kartogramy wartości. W niektórych przypadkach zastosowano dane będące przedmiotem opracowania na tle siatki kwadratów z zastosowaną metodą kartogramu. Dla zilustrowania zagadnień jakości danych wykorzystano również prezentację warstw wektorowych na tle rastrowej mapy topograficznej lub zestawienia graficzne obrazu obiektów na różnych warstwach wektorowych
5. **Metody opisowe** – w celach analitycznych w pracy często stosowano metodę zestandaryzowanego opisu, szczególnie przy słownej charakterystyce baz danych przestrzennych (rozdział 3), gdzie istotne elementy opisu zgromadzono dla każdej bazy w dwunastopunktowym opisie. Pozwoliło to na porównywalność opisów oraz na zebranie precyzyjnej informacji o jakości danych, wynikającej z założeń tworzenia bazy. Uzupełnieniem charakterystyk oraz opisów badań nad jakością baz były zestawienia tabelaryczne. Zaprezentowanie w ten sposób metody potencjału informacyjnego baz danych uczyniło ocenę przejrzystą i w pełni udokumentowaną. Zastosowano również metodę uproszczonej oceny jakości baz danych (Kaczmarek, Medyńska-Gulij 2008) w celu wstępnego zwaloryzowania przedmiotowych baz.

#### **1.4. Przebieg postępowania badawczego**

Dla zrealizowania założonych celów rozprawy przyjęto procedurę badawczą, którą prezentuje ryc. 1.1. Przebieg postępowania badawczego zawiera się w kolejnych rozdziałach pracy, co zaznaczono na ryc. 1.1. i opisano poniżej.

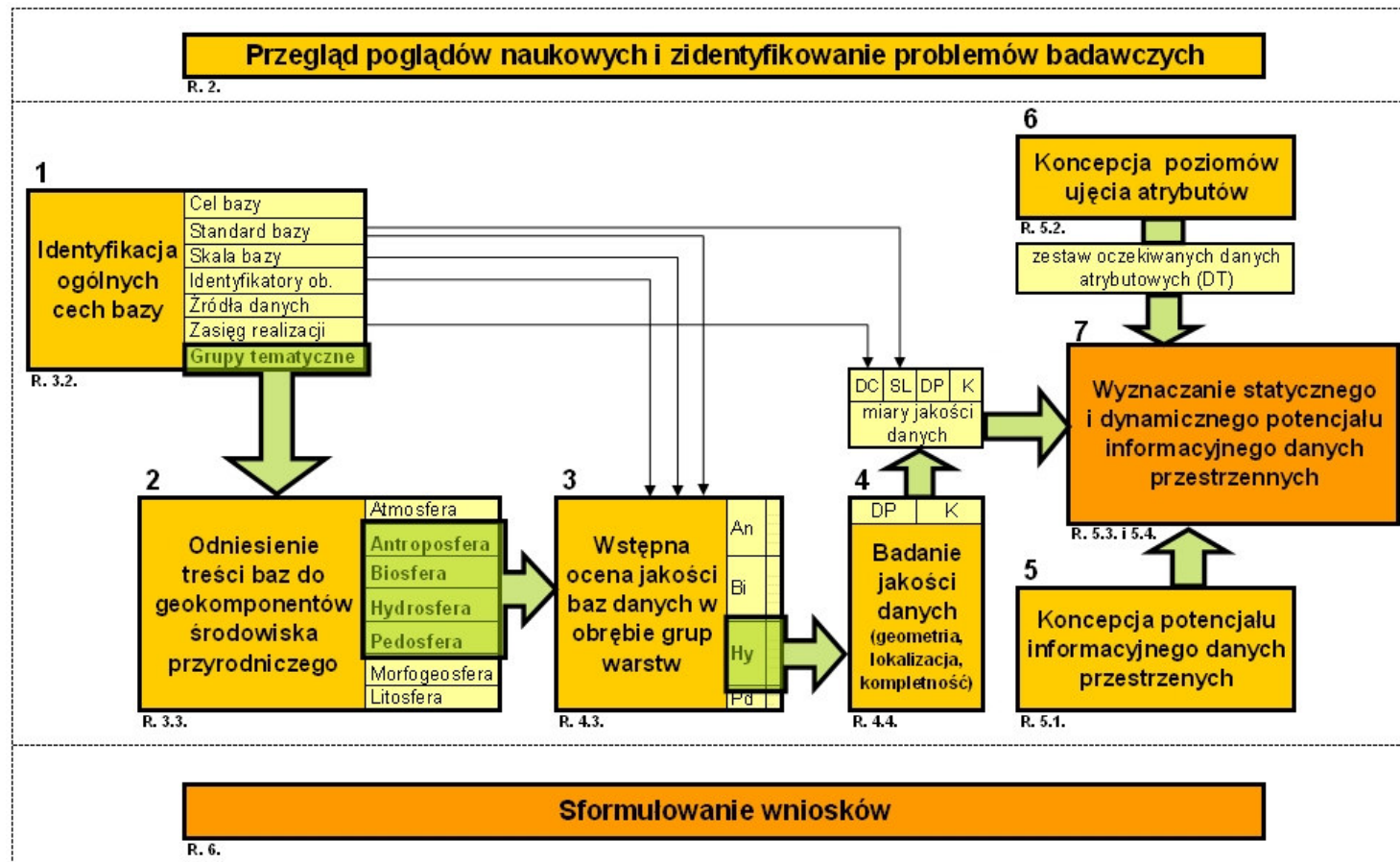
Rozdział 2 obejmuje przegląd poglądów naukowych w zakresie kartograficznych metod badań środowiska przyrodniczego, geoinformacyjnego modelowania środowiska, jakości danych przestrzennych, skal pomiarowych w odniesieniu do danych cyfrowych oraz uwarunkowań tworzenia krajowych baz danych przestrzennych. W podsumowaniu rozdziału zidentyfikowano problemy badawcze wymagające podjęcia w niniejszej pracy.

Rozdział 3 zawiera zestandaryzowany opis przedmiotu badań – najważniejszych krajowych baz danych przestrzennych. Opis służy porównaniu baz w zakresie zawartości merytorycznej, standardów wykonania, skali nominalnej i zastosowanych danych źródłowych. Wyniki analizy zestawiono w syntetycznych tabelach. Dla uwypuklenia istotności danych o komponentach środowiskowych wykonano analizę porównawczą zawartości tych danych w poszczególnych bazach. Na schemacie procedury badawczej (ryc. 1.1.) są to etapy 1 (identyfikacja ogólnych cech bazy) i 2 (odniesienie treści baz do geokomponentów środowiska przyrodniczego).

Rozdział 4 odnosi się do jakości danych przestrzennych zawartych w przedmiotowych bazach. Zastosowano metodę wstępnej oceny jakości baz danych, która pozwoliła na wytypowanie baz danych o cechach rokujących najwyższą jakość. Dalsze prace polegały na badaniu jakości wybranych danych, głównie w zakresie dokładności geometrycznej oraz kompletności danych z wykorzystaniem metod geoinformacyjnych. Na schemacie procedury badawczej (ryc. 1.1.) ten rozdział opisuje etapy 3 (wstępna ocena jakości baz danych w obrębie grup warstw) i 4 (badanie jakości danych).

W rozdziale 5 zaprezentowano koncepcję potencjału informacyjnego danych przestrzennych. Zaprezentowano propozycję klasyfikacji poziomów ujęcia atrybutów. Przedstawiono i empirycznie wyznaczono potencjał informacyjny wybranych danych. W podsumowaniu rozdziału zaproponowano rozróżnienie statycznego i dynamicznego potencjału informacyjnego danych przestrzennych. Rozdział 5 opisuje etap 5 (koncepcja poziomów ujęcia atrybutów), 6 (koncepcja potencjału informacyjnego danych przestrzennych) i 7 (wyznaczanie statycznego i dynamicznego potencjału informacyjnego danych przestrzennych) procedury badawczej (ryc. 1.1.).

Rozdział 6 stanowi zebranie wyników badań w postaci wniosków odniesionych do poszczególnych celów pracy.



Ryc. 1.1. Procedura badawcza

Objaśnienie skrótów: An – antroposfera, Bi – biosfera, Hy – hydrosfera, Pd – pedosfera, DC – dokładność czasowa, SL – spójność logiczna, DP – dokładność przestrzenna, K – kompletność, DT – dokładność tematyczna;

R. 2. – odnośnik do fragmentu rozprawy (numeru rozdziału) opisującego dany etap procedury badawczej

Źródło: opracowanie własne

## **2. PODSTAWY TEORETYCZNE MODELOWANIA KARTOGRAFICZNEGO**

### **2.1. Poglądy na metodykę kartograficznych badań środowiska przyrodniczego**

Konieczność kompleksowego badania środowiska przyrodniczego była szczególnie silnie podkreślana przez badaczy w II połowie XX wieku (Bartkowski 1977, Kondracki 1965, Przewoźniak 1979, Richling 1972, Żynda 1978). Wykreowane w tym okresie koncepcje badań objęły różne podejścia do przedmiotu badań, obejmujące strukturę krajobrazu, funkcjonowanie środowiska czy też ocenę jego walorów. Pojęcie kartograficznej metody badań Saliszczew (1955) uznał za podstawę kartograficznej metody poznania. O metodach analizy map pisali m.in. Berlant (1978), Ratajski (1989). Ten ostatni określił cztery grupy metod analizy kartograficznej: analizy rozmieszczenia, metody korelacji, metody tendencji i metody transformacji. W kartografii amerykańskiej wprowadzono pojęcie kartografii analitycznej (Tobler 1959) z elementami metod ilościowych. Było to wynikiem prób wykorzystywania technik komputerowych. Clarke (1995) wskazuje na rolę kartografii analitycznej w ilościowej transformacji danych w zakresie: wymiarów obiektów, odwzorowań kartograficznych, transformacji międzyobektowych i innych. Różnica między kartograficzną metodą badań a kartografią analityczną wynika z zakresu problematyki badawczej (Żyszkowska 2003). O ile pierwsza wiąże się z formalną stroną metod kartograficznych, o tyle druga jest związana z cyfrowym przetwarzaniem danych kartograficznych. wielu autorów wiąże kartografię analityczną z analizami przestrzennymi przeprowadzanymi w Systemach Informacji Geograficznej. Autorka porządkuje metody analiz na (Żyszkowska 2003):

1. Podstawowe: integracja danych i selekcja informacji
2. Topologiczne: odnoszące się do obiektów powierzchniowych (wyznaczanie regionów – agregacja w oparciu o atrybut, buforowanie; podział obszaru – poligony Voronoi'a, overlay ), obiektów liniowych (lokalizacja; wyznaczenie najkrótszej trasy; wyznaczenie odległości z uwzględnieniem kosztów)
3. Kartometryczne
4. Na tabelach atrybutowych: grupowanie i reklasyfikacja, metody statystyczne
5. Rozkładu przestrzennego
6. Powierzchni 3D: aproksymacja, morfometryczne parametry powierzchni
7. Generalizacja

Pojęcie „kartograficznej metody badań” można odnieść również do użytkowania map (Muehrcke i in. 2001). Wieloetapowy proces kartograficznej metody badań polega na pomiarze, przekształceniu i interpretacji informacji zawartej na mapie (Paślawski, Siwek 2010). Celem jest uzyskanie danych, które są poszerzeniem informacji zawartej na mapie lub stanowią jakościowo nową informację.

Zastosowanie różnych rodzajów analiz i operacji w GIS jest określane jako modelowanie kartograficzne (Tomlin 1990, Delaney 1999). Longley i in. (2006) zwrócili uwagę, że delimitacja homogenicznych jednostek przestrzennych i zakresu ich oddziaływania stanowi tradycyjną metodę stosowaną w naukach geograficznych, jak również w nowoczesnych systemach geoinformacyjnych.

Ważnym elementem nowoczesnego ujęcia kartograficznej metody badań jest wizualizacja danych. Kartograficzna wizualizacja danych, która obejmuje wykorzystanie wcześniejszych opracowań (warstw tematycznych) jest określana jako kompilacja kartograficzna (Dent 1990, Medyńska-Gulij 2007). W procesie analizy i kompilacji danych przestrzennych dużą rolę mogą pełnić istniejące dane cyfrowe w postaci warstw wektorowych urzędowych lub branżowych baz danych przestrzennych (Kaczmarek, Medyńska-Gulij 2007, 2008).

## **2.2. Modelowanie środowiska przyrodniczego w ujęciu cyfrowym**

Zagadnienie modelowania rzeczywistości zostanie przedstawione w trzech aspektach: schematu procesu modelowania w ujęciu kartograficznym i geoinformacyjnym (rozdział 2.2.1.), wyboru reprezentacji obiektów w modelu środowiska przyrodniczego (rozdział 2.2.2.) oraz znaczenia generalizacji w modelowaniu rzeczywistości (rozdział 2.2.3.).

### 2.2.1. Modelowanie kartograficzne a modelowanie danych geograficznych

Modelowanie rzeczywistości wywodzi się z kartografii. Ideę traktowania mapy jako modelu opisującego i wyjaśniającego zjawiska przestrzenne wprowadzili w 1967 r. równocześnie Saliszczew i Board. W literaturze kartograficznej zarysowały się dwie koncepcje spojrzenia na modelowanie: subiektywistyczna (Board 1967) oraz materialistyczno-obiektywistyczna (Saliszczew 1967, Ratajski 1978, Czerny 1994a). Pierwsze podejście zakłada, że mapa jest subiektywnym modelem, będącym obrazem świata, konstruowanym przez autora na zasadzie analogii (Żyszkowska 2000). Natomiast koncepcja materialistyczno-obiektywistyczna zakłada, że mapa jest obiektywnym obrazem

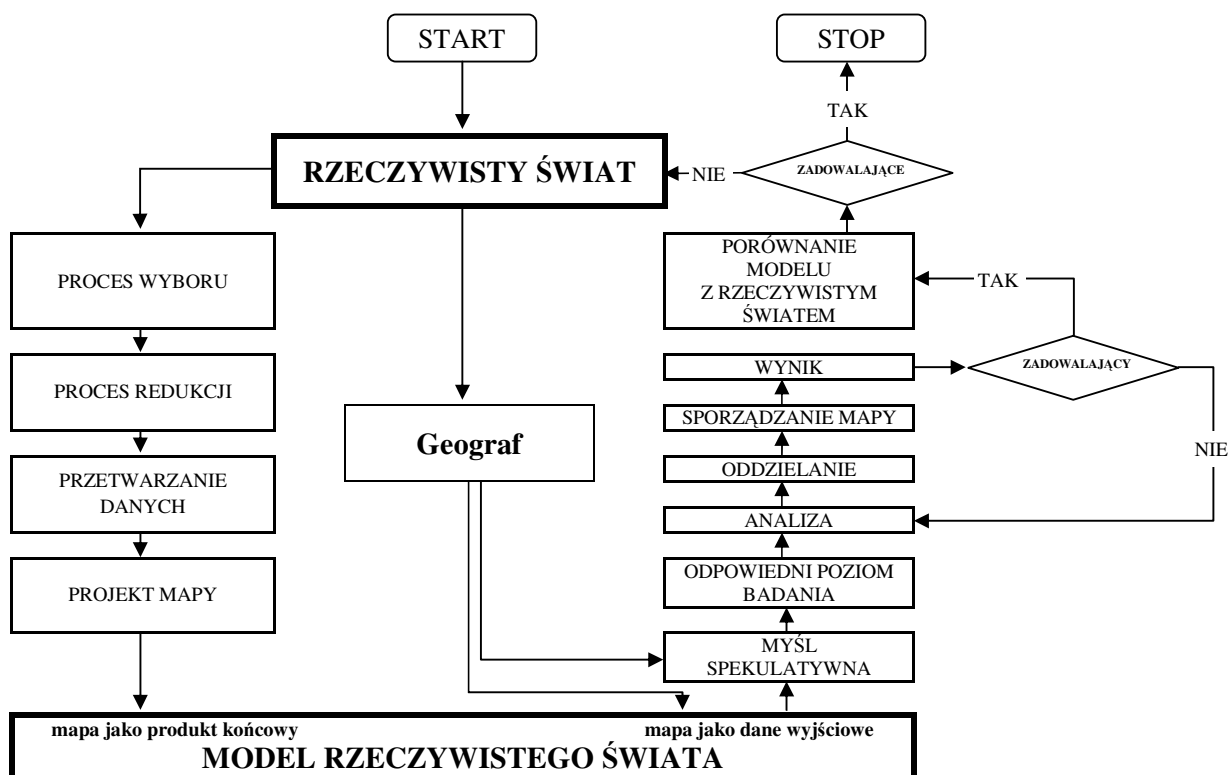
rzeczywistości, a modelowanie kartograficzne jest odwzorowaniem tworzącym jednoznaczne relacje między rzeczywistością a mapą (Czerny 1994).

Powyższe uwarunkowania rozwoju koncepcji modelowania rzeczywistości ugruntowały pojęcie **modelowania kartograficznego**. Należy je rozumieć jako wizualizację kartograficzną, która obejmuje analizę zjawisk przestrzennych i ich przedstawianie (Żyszkowska 2000: 20). Modelowe podejście wzmocniło się w kartografii w ostatnich dwóch dekadach, przede wszystkim w odniesieniu do systemów informacji geograficznej (Kraak, Ormeling 1998, Ostrowski 2008). Stąd zauważalne są koncepcje opierające się na **modelowaniu danych** (Bernhardsen 2002), odnoszące modelowanie rzeczywistości do projektowania cyfrowych baz danych (Peuquet 1984, Zeiler 1999, Longley i in. 2006). Model danych Peuquet określa jako abstrakcję świata rzeczywistego zawierającą tylko te właściwości, które są istotne w konkretnym zastosowaniu lub zastosowaniach, zazwyczaj ludzkiego rozumienia rzeczywistości (Peuquet 1984). Longley i in (2006) określają model danych jako uporządkowany cyfrowy opis służący do reprezentacji wybranych cech świata rzeczywistego. Przedmiotem modelowania są zarówno dane o charakterze przestrzennym, jak i atrybuty obiektów zwane danymi nieprzestrzennymi (Bielecka 2006). Geometria obiektów, czyli sposób odniesienia przestrzennego danych lub reprezentacja w modelu danych, zależy od ich charakteru – wyraźnie zdelimitowanych przestrzennie jednostek (obiektów dyskretnych) lub ciągłej reprezentacji cech środowiska (pól; Longley i in. 2006). Dane atrybutowe, będące drugim elementem modelowanym, zostaną omówione w rozdziale 2.3.

Koncepcje **kartograficznego modelowania** rzeczywistości były realizowane w szerokim lub wąskim zakresie znaczeniowym, odpowiednio jako etapowe przejście do mapy albo wyłącznie jako proces wizualizacji.

Board (1967) określa modelowanie kartograficzne jako iteracyjny cykl mapa-model (ryc. 2.1.). Obraz rzeczywistości w procesie tworzenia mapy podlega przekształceniu poprzez generalizację obejmującą wybór i redukcję informacji. W dalszych etapach przetwarzania danych oraz projektowania mapy następuje określenie sposobu ujęcia kartograficznego. Mapa jako produkt końcowy jest jednocześnie materiałem wyjściowym w procesie badawczym dla badacza środowiska przyrodniczego (geografa w ujęciu Boarda; ryc. 2.1., prawa strona schematu). Jego potrzeby badawcze definiują sposób filtrowania danych, analogicznie do procesu tworzenia mapy. Analiza jest przez tego autora rozumiana jako przetwarzanie danych, a proces oddzielania polega na zidentyfikowaniu źródła zmienności i odróżnieniu go

od zakłóceń rzeczywistości. Wynik analizy w postaci mapy staje się nowym spojrzeniem na rzeczywisty świat. Jeśli jest on akceptowany podlega weryfikacji względem rzeczywistości. Jeśli wynik nie jest zadowalający następuje powtórzenie procesu analizy (ryc. 2.1., prawa strona schematu). Ostateczna weryfikacja mapy polega na jej porównaniu ze światem rzeczywistym. Pozytywny wynik weryfikacji kończy cykl iteracyjny, negatywny – uruchamia modelowanie od nowa (ryc. 2.1., prawy górny fragment schematu)

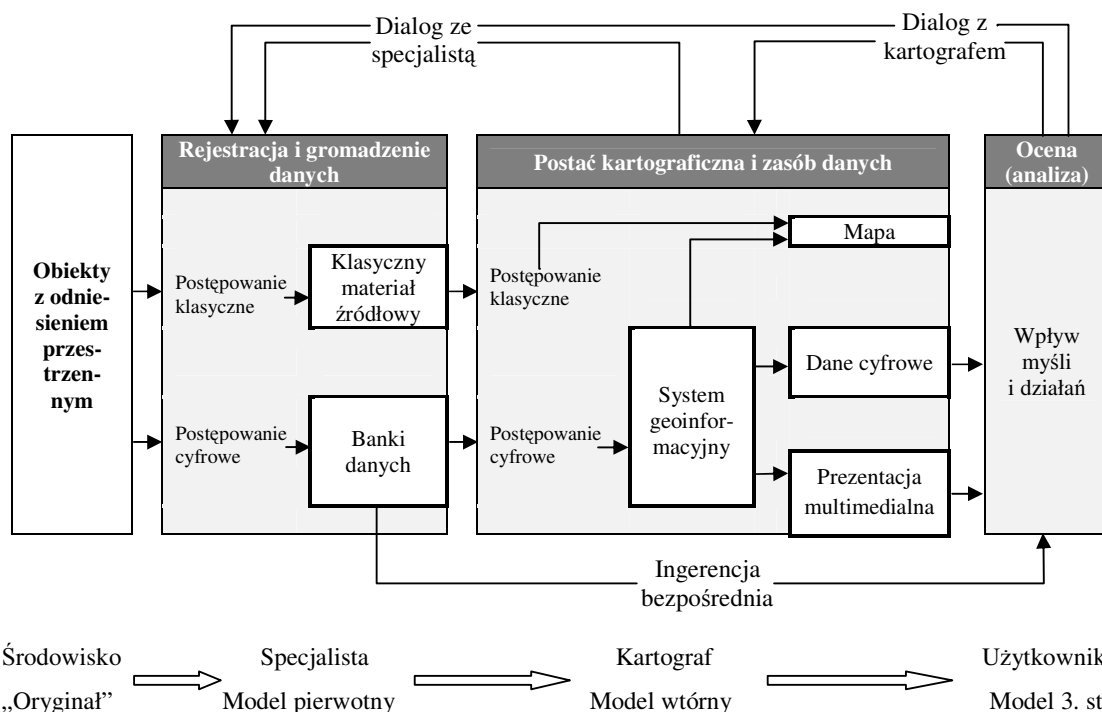


Ryc. 2.1. Cykl mapa-model

Źródło: opracowanie własne na podstawie Boarda (1967)

Modele kartograficzne zdaniem Hake i in. (2002) mają szerszy zakres przedmiotowy i występują trójstopniowo w postaci: modelu pierwotnego, będącego wynikiem bezpośredniej obserwacji rzeczywistości, modelu wtórnego – prezentacji kartograficznej oraz modelu trzeciego stopnia, stanowiącego wyobrażenie przedstawionych na mapie obiektów. Każdemu poziomowi modelowemu przypisany jest inny uczestnik procesu komunikacyjnego w kartografii, odpowiednio: w modelu pierwotnym specjalista (np. topograf, geolog, statystyk), we wtórnym kartograf, a w modelu trzeciego stopnia użytkownik. Taki trójstopniowy schemat ze zdefiniowanymi interakcjami między uczestnikami określony został przez autorów mianem kartograficznej sieci komunikacyjnej (ryc. 2.2.). Autorzy w obrębie sieci rozróżniają modele analogowe i cyfrowe.

Podobnie jak Hake i in. (2002) również Brodlie i in. (2001) wyznaczyli trzy obszary transformacji informacji (ryc. 2.3.). Transformacja 1 obejmuje proces rozpoznania informacji geograficznej na drodze opisu, pomiarów, rejestracji oraz kompilowania informacji. Transformacja 2 ma charakter transformacji kartograficznej, natomiast transformacja 3 prowadzi do stworzenia u użytkownika wyobrażenia o mapie.

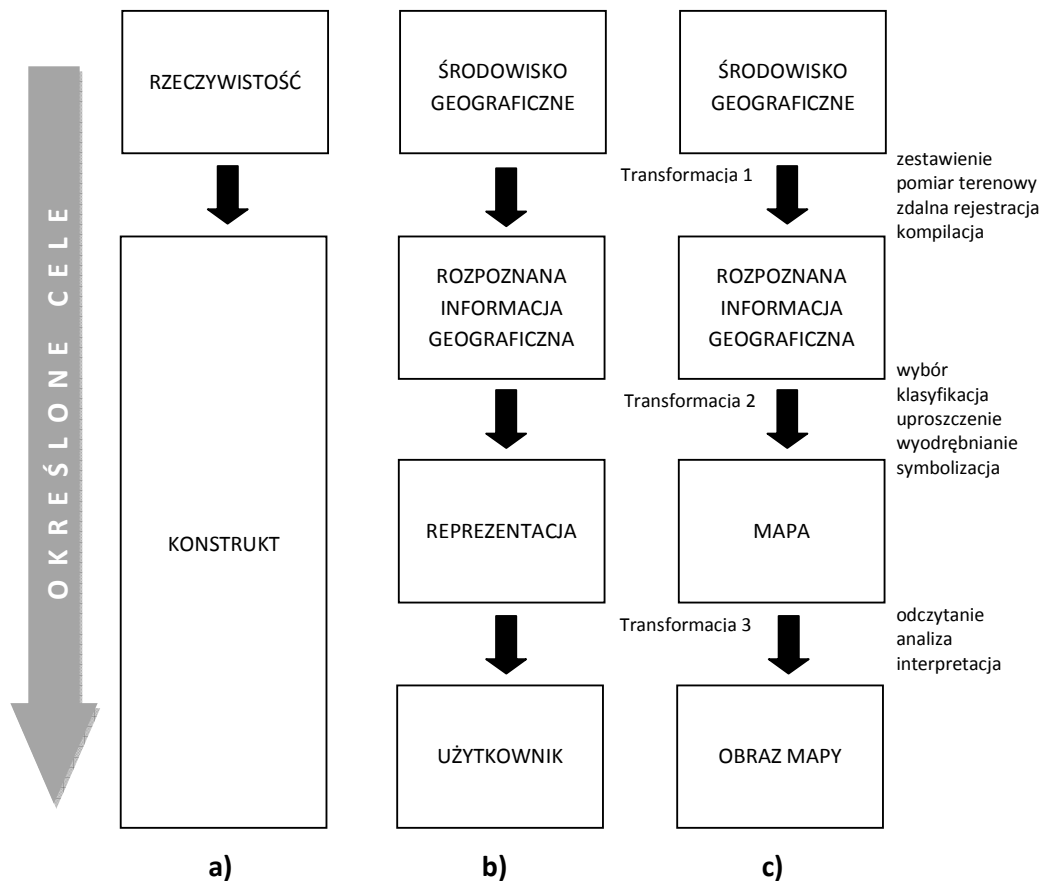


**Ryc. 2.2. Kartograficzna sieć komunikacyjna na tle trójstopniowego modelu kartograficznego (Hake i in. 2002)**

Wąskie ujęcie modelowania kartograficznego obejmuje proces transformacji kartograficznej. Proces modelowania rzeczywistości zdaniem Ratajskiego (1973) ma na celu określenie zależności semantycznej między znakiem na mapie a jego treścią znaczeniową. Żyszkowska (2000) traktuje modelowanie kartograficzne jako proces wizualizacji rzeczywistości. Autorka określa trzy etapy modelowania: budowanie pojęciowego modelu przestrzeni geograficznej (rozumianego jako logiczna struktura mapy), tworzenie systemu symboli (model syntaktyczny) oraz transformacja modelu syntaktycznego na obraz kartograficzny.

Wizualizacja geograficzna (Slocum i in. 2010, Dykes i in. 2005; Kraak, Ormelling 2009) jest nurtem w kartografii rozwijającym się w nawiązaniu do wykorzystywania form

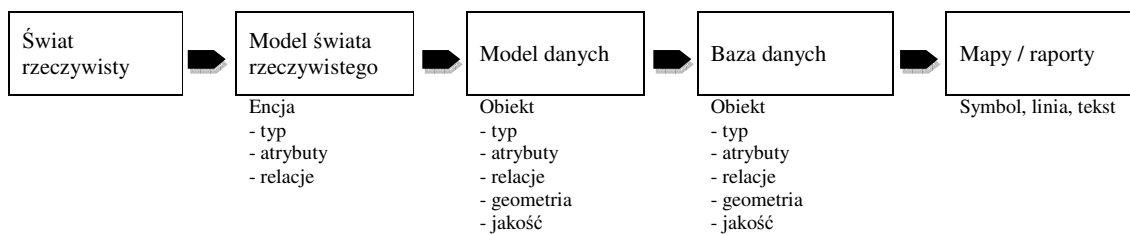
kartograficznej prezentacji danych przestrzennych. Kluczowym pojęciem geowizualizacji staje się interaktywne ujawnianie zjawisk przestrzennych zapisanych przede wszystkim w bazach danych geografii (Slocum i in. 2010, MacEarchen 1994).



Ryc. 2.3. Schematyczne reprezentacje wirtualnej rzeczywistości a transformacja kartograficzna (Brodie i in 2001); a) rzeczywistość wirtualna (model podstawowy), b) rzeczywistość wirtualna (model rozszerzony), c) transformacyjne podejście kartografii

Odmienne w treści są koncepcje opierające się na **modelowaniu danych przestrzennych**. Podejście to wynika z włączenia do modeli procesu budowania cyfrowych baz danych, a nie tworzenia mapy w tradycyjnym ujęciu. Zdaniem Zeilera (1999) modelowanie informacji geograficznej przebiega przez trzy etapy: konceptualny, logiczny i fizyczny. Na etapie konceptualnym (pojęciowym) następuje tworzenie modelu perspektywy użytkownika, które polega na określeniu przeznaczenia bazy danych, zakresu niezbędnych danych oraz sposobu ich organizacji. Na tym etapie następuje w szczególności: zdefiniowanie obiektów (klas obiektów), relacji między obiektami oraz wybór reprezentacji geometrycznej. Przy definiowaniu geometrycznej reprezentacji następuje podział na obiekty dyskretne, o

wyraźnych granicach przestrzennych i zjawiska o charakterze ciągłym, realizowane najczęściej w postaci pól. Etap logiczny obejmuje dobór typów danych w bazie oraz organizację struktury bazy danych przestrzennych. W etapie fizycznym następuje zdefiniowanie fizycznego schematu bazy danych, który doprowadza do podziału opisywanej rzeczywistości na klasy obiektów. Proces modelowania zdaniem Peuquet (1984) wiąże się z przechodzeniem na kolejne poziomy abstrakcji: od rzeczywistości poprzez model danych, strukturę danych aż do struktury fizycznego pliku komputerowego. Bernhardsen (2002) traktuje modelowanie rzeczywistości z punktu widzenia GIS jako przejście z rzeczywistości poprzez jej model, model danych, bazę danych aż do mapy (ryc. 2.4.).



**Ryc. 2.4. Przejście ze świata rzeczywistego do GIS poprzez tworzenie kolejnych modeli upraszczających rzeczywistość (Bernhardsen 2002)**

W tworzeniu modelu świata rzeczywistego według Bernhardsena (2002) zasadniczym nośnikiem informacji jest encja określana jako reprezentacja przedmiotu lub zjawiska świata rzeczywistego. Encja należy do określonego typu klasyfikacyjnego, posiada atrybuty oraz charakteryzuje się określonymi relacjami. Jeśli wyobrażenie świata ma być zrozumiałe przez ogół ludzi, konieczne jest utworzenie modelu danych, w którym encje zostają przetworzone na obiekty geometryczne, przy czym jedna encja może być opisana przez wiele obiektów geometrycznych. W wyniku przyjętego modelu danych tworzona jest fizyczna baza danych, w której obiekty charakteryzują się odpowiednią jakością przechowywanych informacji.

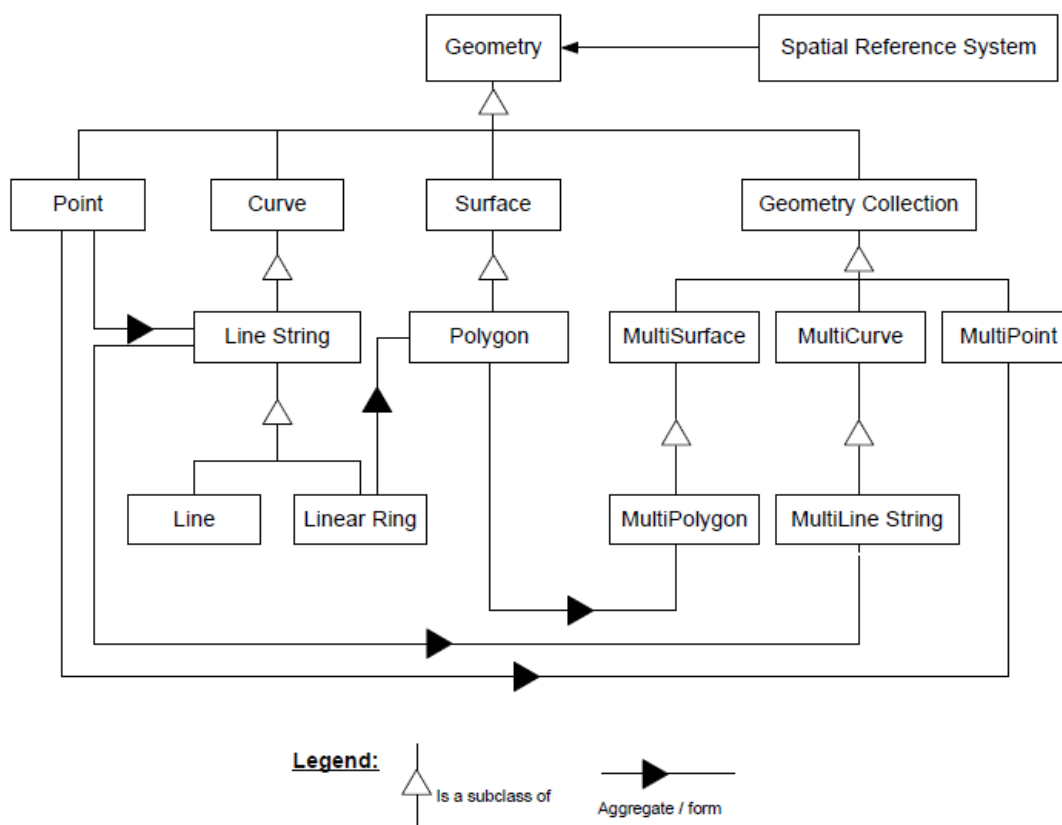
Podsumowując można stwierdzić, iż modelowanie rzeczywistości charakteryzuje się dualizmem podejścia: bazodanowego i kartograficznego. Na gruncie kartograficznym powstało pojęcie cyfrowego modelu krajobrazu oraz cyfrowego modelu kartograficznego (Kraak, Ormeling 1998, Makowski 2005, Głazewski 2006) jako współczesnych odpowiedników modelowania rzeczywistości, właściwego dla kartografii cyfrowej. Cyfrowy model krajobrazu (ang. *digital landscape model*, DLM) stanowi zbiór danych odnoszących się do wybranego obszaru, stąd często nazywany jest modelem topograficznym (Głazewski 2006). Cyfrowy model kartograficzny (ang. *digital cartographic model*, DCM) powstaje w wyniku transformacji kartograficznej cyfrowego modelu krajobrazu.

### 2.2.2. Reprezentacja obiektów przestrzennych

W modelowaniu rzeczywistości na poziomie cyfrowych modeli krajobrazu i cyfrowych modeli kartograficznych obiekty i zjawiska są przedstawiane jako reprezentacje (Longley i in. 2006). Podstawowe modele danych przestrzennych będące wynikiem modelowania i mające powszechne zastosowanie w bazach danych topograficznych i tematycznych to raster i wektor. Rastrowy model danych wykorzystuje macierz elementów – pikseli do przechowywania informacji o obiektach. W komórkach rastra mogą być przechowywane wartości atrybutów (kategorie obiektów, liczby całkowite reprezentujące np. barwę lub liczby zmiennoprzecinkowe określające np. rzędną terenu). W klasycznym przypadku dane rastrowe są gromadzone w pliku jako tablica wartości przypisanej poszczególnym oczkom regularnej siatki (ang. *grid*). W wektorowym modelu danych każdy element środowiska przyrodniczego jest najpierw klasyfikowany do odpowiedniego typu geometrycznego: punktów, linii lub wieloboków. Poszczególne elementy tego samego typu tworzą klasę elementów.

Obiekty w modelu danych mogą mieć charakter dyskretny lub ciągły. Obiekty dyskretny mają wyraźnie określone granice określające jednoznacznie ich przestrzenny zasięg. Cechą charakterystyczną tych obiektów jest ich policzalność. Reprezentacje o charakterze ciągłym dają możliwość określenia wartości zmiennej obiektu w dowolnym punkcie przestrzeni, którą zajmują. Ten typ obiektów najczęściej określa się mianem pola (Longley i in., 2006).

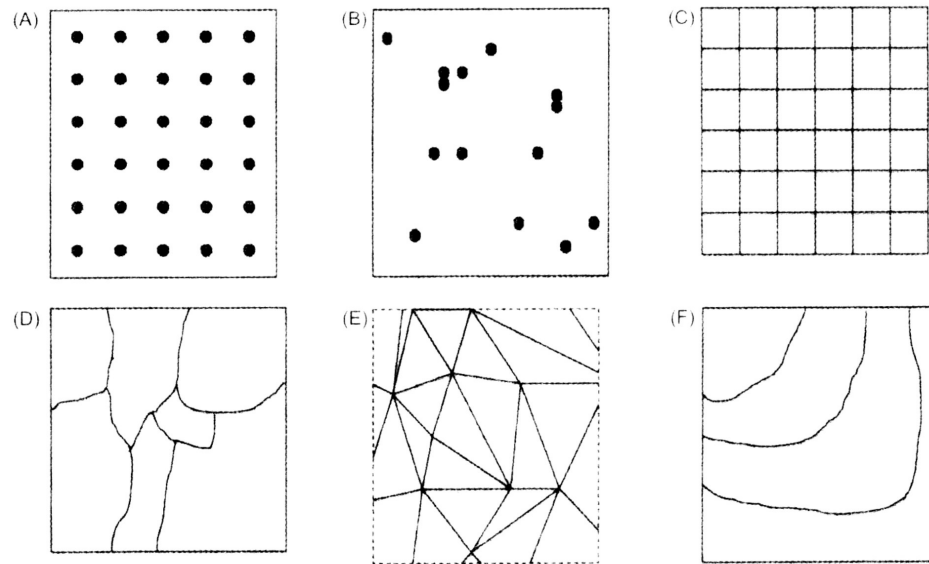
Zbiór obiektów o charakterze dyskretnym może być reprezentowany w modelu środowiska przez punkty, linie, powierzchnie i objętości (Kraak, Ormeling 1998). Są to terminy ogólne (Chang 2004), określane m.in. przez Yeunga i Brent Halla (2007) jako graficzne prymitywy (ang. *graphical primitives*), a ich synonimami są: dla punktów – węzeł (ang. *node*), wierzchołek (ang. *vertex*), zerowymiarowa komórka (ang. *0-cell*); dla linii – krawędź (ang. *edge*), ogniwo (ang. *link*), łańcuch (ang. *chain*), jednowymiarowa komórka (ang. *1-cell*); dla powierzchni – wielobok (ang. *polygon*), powierzchnia (ang. *face*), strefa (ang. *zone*), dwuwymiarowa komórka (ang. *2-cell*) (Laurini, Thompson 1992). W podziale przyjętym przez *Open Geospatial Consortium* (OGC, 1999) głównymi obiektami geometrycznymi są: punkt, krzywa, powierzchnia oraz zbiór elementów geometrycznych (ryc. 2.5.). Według podziału Gomasca (2009: Fig. 9.13.) wektorowe reprezentacje obiektów geometrycznych obejmują: węzeł, linię lub segment, krzywą, poligon, powierzchnię oraz objętość.



**Ryc. 2.5. Model obiektów geometrycznych wg OGC (1999)**

W obrębie głównych typów obiektów wyróżniane są podtypy. Podtypy można wyróżniać w odniesieniu do struktury organizacji danych przestrzennych oraz charakteru geometrii. Ze względu na strukturę organizacji danych przestrzennych można wyróżnić np. w przypadku obiektów liniowych strukturę drzewiastą (np. system rzeczny), niepołączone linie (np. uskoki) oraz strukturę sieciową (np. sieć dróg) (Peuquet 1984). Ze względu na charakter geometrii obiekty liniowe mogą być traktowane jako: odcinek, połączenie, wektor, łamana otwarta, krzywa lub łamana zamknięta (Bielecka 2006). Obiekty przestrzenne mogą przyjmować postać prostą lub złożoną (Longley i in. 2006, Chang 2004, Yeung, Brent Hall 2007). O ile postać prostą mogą stanowić obiekty jednego typu geometrycznego (punkt, linia lub powierzchnia) o tyle obiekt złożony składać się będzie ze zbioru obiektów prostych. Na przykład sieć rzeczna będzie składać się z obiektów powierzchniowych (rzeki o znacznej szerokości), liniowych (pozostałe rzeki i ciekł wodne) oraz obiektów punktowych (węzły łączące dopływy z rzekami wyższego rzędu).

Dane o charakterze dyskretnym ze względu na swój charakter najlepiej są realizowane w wektorowym modelu danych (Longley i in. 2006). Natomiast cechy środowiska mające charakter ciągły mogą być reprezentowane na wiele różnych sposobów (Longley i in. 2006): jako regularna siatka wartości, siatka punktów rozmieszczonych nieregularnie, regularna siatka pól (pikseli), układ pól o nieregularnych kształtach, nieregularna sieć trójkątów czy też izolinie (ryc. 2.6.). Do reprezentacji danych o charakterze ciągłym ma więc zastosowanie zarówno model wektorowy, jak i rastrowy.



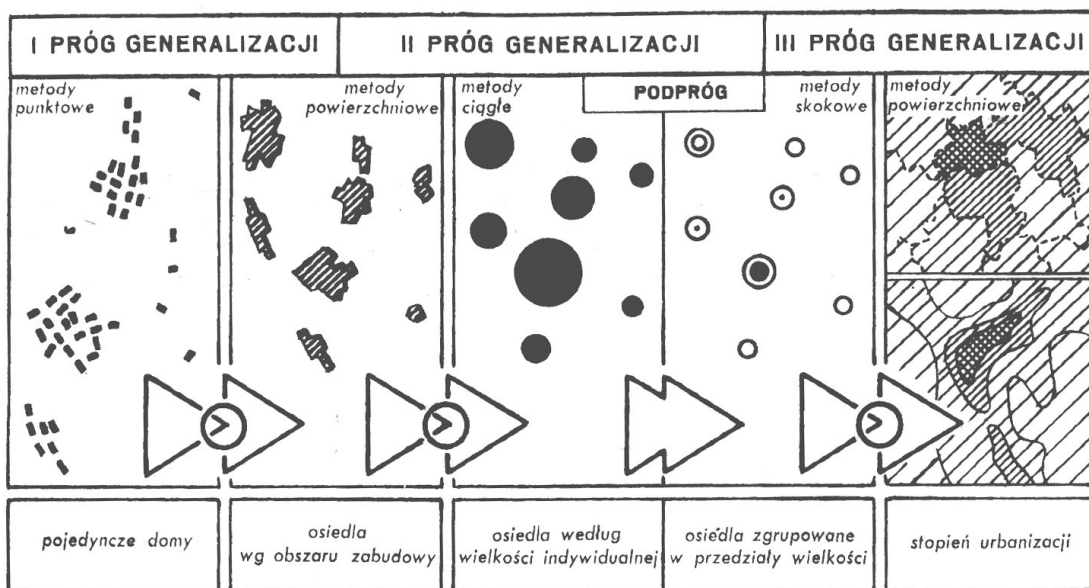
**Ryc. 2.6. Sposoby reprezentacji pola wg Longley i in. (2006): (A) regularna siatka punktów, (B) nierównomiernie rozmieszczone punkty, (C) prostokątne komórki (piksele), (D) wielokąty reprezentujące pola o nieregularnych kształtach, (E) nieregularna sieć trójkątów (powierzchnia TIN), (F) linie łamane reprezentujące poziomicę.**

### 2.2.3. Generalizacja kartograficzna i konceptualna

Modelowanie informacji geograficznej nie kończy się na definicji obiektów dyskretnych i ciągłych oraz ich reprezentacji cyfrowych: rastrowych i wektorowych. Jak już wcześniej wspomniano problemy pojawiają się na etapie dostosowania reprezentacji zjawisk do sposobu fizycznego zapisu w systemach komputerowych. Przy rozpatrywaniu zagadnień modelowania informacji geograficznej istotne jest pojęcie generalizacji.

Modelowy charakter map geograficznych i innych przedstawień kartograficznych wynika z zdaniem Saliszczewa (1998) z selektywnego podejścia oraz generalizacji. Zdaniem Ostrowskiego (2008) generalizacja stanowi istotę modelowania kartograficznego. Ratajski (1973) dzieli generalizację na jakościową, która polega na uogólnieniu pojęć prezentowanych

na mapie, oraz ilościową, z którą wiąże się z redukcją liczby sygnałów. Ratajski ponadto wysuwa teorię progów generalizacyjnych w kartografii, które występują w momencie zmiany kategorii pojęciowej obiektów do nadrzędnej w sytuacji przekroczenia progu pojemności mapy. Przy przejściu do kolejnego etapu generalizacji następuje także zmiana metody prezentacji (ryc. 2.7.).



**Ryc. 2.7. Schemat progów generalizacji według Ratajskiego (1973)**

Bernhardsen (2002) odróżnia generalizację konceptualną, zwaną też strukturyzacją danych, od generalizacji kartograficznej, powstałą na etapie tworzenia modelu danych świata rzeczywistego, która ma miejsce na etapie modelowania kartograficznego z bazy danych. Stankiewicz (2000) wymienia wybór i klasyfikację oraz generalizację jako podstawowe procesy towarzyszące modelowaniu pojęciowemu. Longley i in. (2006) wskazują, że nie ma możliwości stworzenia idealnego opisu świata w bazie danych. Konieczne jest stosowanie technik generalizacji danych, które McMaster i Shea (1992) połączyli w dziesięć grup czynności: upraszczanie, wygładzanie, dekompozycja, agregacja, łączenie, scalanie, wybór obiektów, przewiększanie, wzmocnienie i przemieszczanie. Podział generalizacji według Kraaka i Ormellinga (1998) nawiązuje do klasyfikacji generalizacji, którą przedstawił Ratajski (1973). Autorzy ci wyróżniają generalizację pojęciową (odpowiednik jakościowej wg Ratajskiego) i graficzną (odpowiednik ilościowej). Przypisują oni każdemu z tych typów odpowiednie procedury: w przypadku generalizacji pojęciowej łączenie, wybór, symbolizację i wypuklenie, a dla generalizacji graficznej uproszczenie, przewiększenie, przesunięcie,

łączenie i wybór. Eckes (2006) wymienia trzy grupy działań generalizacyjnych związanych z modelowaniem rzeczywistości geograficznej: grupowanie obiektów w klasy, wybór reprezentacji obiektów wraz z uproszczeniem kształtów oraz zmniejszenie liczby wymiarów. Zdaniem Harvey'a (2008) pięć typów pospolitych operacji ma znaczącą rolę w generalizacji kartograficznej: agregacja, przesunięcia, rozszerzanie, wybór i upraszczanie.

### **2.3. Czynniki wpływające na jakość danych przestrzennych**

Zagadnieniem fundamentalnym przy korzystaniu ze źródłowych danych przestrzennych jest ich jakość. Pojęcie jakości danych ma charakter subiektywny i zawsze musi być rozpatrywane w odniesieniu do celu ich wykorzystania (Bielecka 2006). Obejmuje ono kompleks cech i charakterystyk danych, które wpływają na zaspokojenie przez te dane wymagań użytkowników. Zdaniem Bieleckiej (2006) miarą jakości danych jest oszacowanie rodzajów i wielkości błędów.

Główne błędy w danych powstają na etapie modelowania rzeczywistości (por. rozdz. 2.2.). Wielkość błędów jest więc zależna od dokładności lokalizacji obiektów w modelu, określenia atrybutów oraz od precyzji fizycznego zapisu danych (Bielecka 2006). Litwin i Myrda (2005: 101) zwracają uwagę na odróżnienie określenia dokładności i precyzji danych. To pierwsze pojęcie wiąże z teorią rozprzestrzeniania się błędów, natomiast precyzję odnoszą do poziomu, do jakiego schodzi operator wprowadzający dane lokalizacyjne i atrybutowe. Chang (2004) oraz Longley i in. (2006) rozpatrując dane przestrzenne rozróżniają dokładność, która określa zależność między zarejestrowaną i rzeczywistą lokalizacją obiektu oraz precyzję jako stopień szczegółowości zarejestrowania lokalizacji. Magnuszewski (1999) określając dokładność danych wyróżnia błędy obciążające geometrię oraz atrybuty obiektów. Ponieważ w wielu przypadkach baz danych źródłem informacji są mapy, więc upatruje on czynników wpływających na wielkość błędu np. w aktualności i jakości materiału kartograficznego, stopniu generalizacji map czy błędzie dyskretyzacji wektoryzowanych linii (Magnuszewski 1999). Błędy w atrybutach mogą być wynikiem niewłaściwego zdefiniowania obiektu na etapie budowania modelu pojęciowego lub wynikać z niewłaściwej procedury agregacji, klasyfikacji lub interpolacji danych (Magnuszewski 1999).

Na ocenę jakości danych geograficznych składają się (Bielecka 2006):

- kompletność - stosunek danych zebranych w bazie do danych, które powinny być zgromadzone (zarówno obiektów, jak i atrybutów),

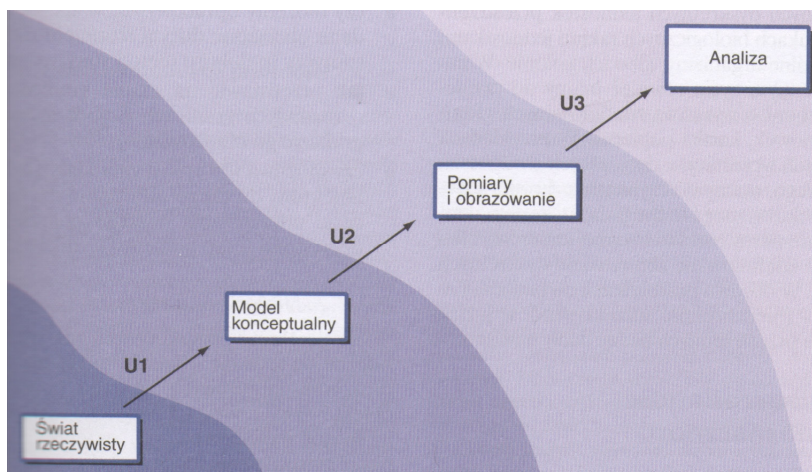
- spójność logiczna – zgodność między modelem rzeczywistości a danymi zgromadzonymi w bazie danych w zakresie semantyki, dziedziny, topologii i formatu,
- dokładność położenia – zależności lokalizacyjne obiektów przedstawiane w sposób bezwzględny, względny lub w stosunku do regularnej siatki,
- dokładność czasu – m. in. aktualność danych, dokładność pomiaru czasu,
- dokładność tematyczna – mierzona liczbą atrybutów oraz dokładnością wartości.

Longley i in. (2006) uważają, że danym przestrzennym towarzyszy niepewność, ponieważ nie można uzyskać nieskończenie dokładnej reprezentacji środowiska przyrodniczego. Autorzy ci wymieniają na drodze od rzeczywistości do jej reprezentacji trzy etapy, na których dochodzi do redukcji strumienia informacji, a przez to do przekształcenia lub zniekształcenia obrazu świata (ryc. 2.8). Przejawia się ona w:

- 1) niepewności modelu konceptualnego (U1 na ryc. 2.8.) – szczególnie w zakresie: a) nieokreśloności granic obiektów i zjawisk rzeczywistych, które w procesie modelowania stają się obiektami dyskretnymi oraz b) niejednoznaczności określeń i nazw geograficznych, niewłaściwego doboru wskaźników opisujących zjawiska oraz błędnej definicji i klasyfikacji obiektów i zjawisk;
- 2) niepewności pomiaru i reprezentacji zjawisk geograficznych (U2 na ryc. 2.8.) – głównie w zakresie błędu pomiaru lokalizacji oraz stopnia niepewności, która towarzyszy wektorowej i rastrowej reprezentacji zjawisk;
- 3) niepewności rezultatów analizy danych przestrzennych (U3 na ryc. 2.8.) – określana przez miarę propagacji błędów danych poddanych analizie.

Szczegółowo problem niepewności danych przestrzennych opisują Zhang i Goodchild (2002). Autorzy wskazują, że do opisu dokładności modelowanych obiektów geograficznych potrzebnych jest pięć składowych (2002): pochodzenie, dokładność atrybutów, zawartość logiczna i kompletność. Pochodzenie wynika z jakości materiału, z którego dane zostały pozyskane oraz dokładności zastosowanych metod pozyskania. Dokładność atrybutów określa stopień odzwierciedlenia rzeczywistych wartości ilościowych lub jakościowych cech. Zawartość logiczna przejawia się między innymi w realizacji reguł topologii międzyobiektowej. Kompletność jest miarą obecności w bazie danych istotnych i aktualnych dla danego obiektu.

Zdaniem Bernhardsena (2002) jakość danych przestrzennych dotyczy trzech aspektów: geometrii, atrybutów oraz relacji, a szacowanie jakości powinno uwzględniać: dokładność lokalizacyjną, dokładność atrybutów, dokładność czasu, zawartość logiczną i kompletność danych.



**Ryc. 2.8. Kolejne etapy uproszczenia rzeczywistości wpływające na niepewność zobrazowań zjawisk i obiektów (Longley i in. 2006)**

Medyńska-Gulij (2009) w ocenie jakościowych i ilościowych informacji historyczno-kartograficznych zawartych na dawnych mapach zastosowała pojęcie „potencjału informacyjnego niekartometrycznej mapy”. Zdaniem autorki można go rozpatrywać w odniesieniu do: 1) liczby znaków punktowych, liniowych i powierzchniowych zaprojektowanych na mapie, sposobów przedstawiania rzeźby terenu oraz napisów, 2) poziomów pomiarowych wykorzystanych przy opisie zjawisk na mapie, 3) charakteru liczbowego lub opisowego cech, 4) względnej dokładności lokalizacyjnej.

Według *US Federal Geographic Data Committee* (FGDC) na jakość danych przestrzennych składają się: poprawność atrybutów i geometrii obiektów, logiczna spójność, kompletność i stopień pokrycia przestrzennego (Longley i in. 2006).

W standardzie SDTS (*Spatial Data Transfer Standard*) jakość danych jest definiowana przez pięć wartości (Altheide 2008):

- pochodzenie - opisuje rodzaj materiału źródłowego oraz proces pozyskania danych;
- dokładność lokalizacyjna - określana przez dokładność poziomą i pionową oraz zastosowaną metodę pomiaru;
- dokładność atrybutowa - odpowiada ufności do wartości atrybutów oraz metod weryfikacji;
- zawartość logiczna - dotyczy poprawności relacji przestrzennych jak np. zależności topologicznych;
- kompletność - odpowiada wyborowi kryteriów użytych w trakcie gromadzenia danych i metod wykorzystanych do gromadzenia danych.

Jakość danych przestrzennych jest przedmiotem unormowań. W standardach ISO (*International Standards Organization*) trzy dokumenty normalizacyjne:

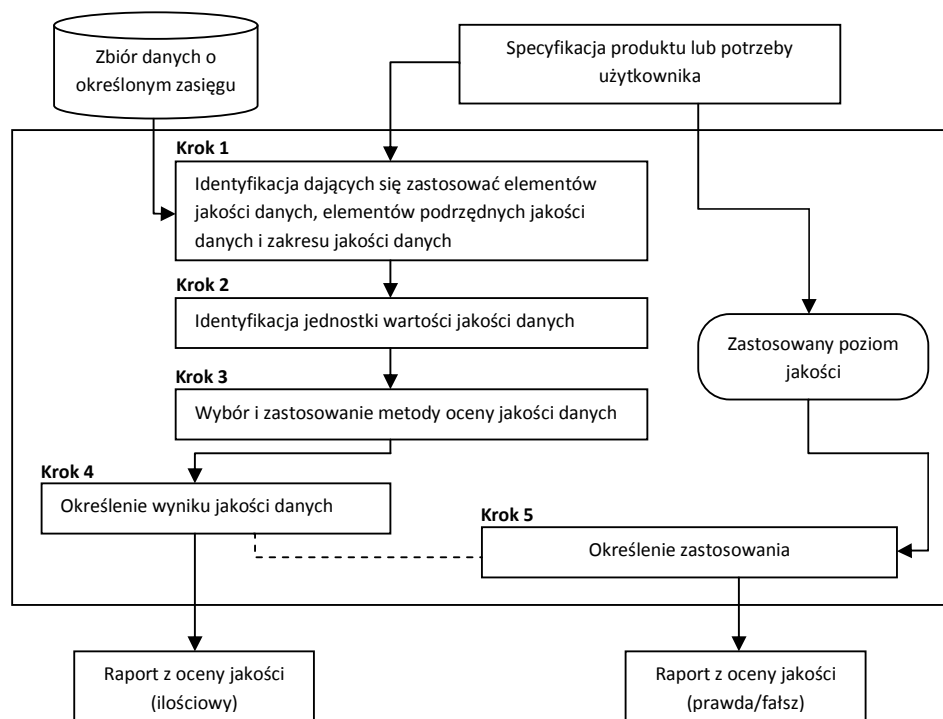
- ISO 19113:2002 (EN ISO 19113:2005) *Geographic information – Quality principles*
- ISO 19114:2002 (EN ISO 19114:2005) *Geographic information – Quality evaluation procedures*
- ISO /AWI 19138:2002 *Geographic information – Data quality measures*

Dwie pierwsze z tych norm w wyniku prac Polskiego Komitetu Normalizacyjnego, Komitet Techniczny 297 ds. Informatyki geodezyjnej doczekały się zatwierdzenia na gruncie polskim:

- PN-EN ISO 19113:2005 Informacja geograficzna – Podstawy opisu jakości
- PN-EN ISO 19114:2005 Informacja geograficzna – Procedury oceny jakości

Ponadto norma ISO 19113 została w całości przetłumaczona na język polski.

Problematyka jakości danych jest unormowana w tych przepisach pod kątem organizacji opisu jakości, sposobu określania i oceny modelu jakości dla danych, procedur i metod oceny jakości danych, stosowanych miar jakości oraz formatu zapisu jakości w postaci raportu oraz metadanych (zgodnie z normą ISO 19115). Normy ISO określają pięć etapów oceny jakości danych geograficznych (ryc. 2.9.).



Ryc. 2.9. Etapy oceny jakości danych geograficznych wg ISO 19114

Problematyka jakości krajowych baz danych przestrzennych była podnoszona w literaturze wielokrotnie, choć nie zawsze kompleksowo. Najczęściej sygnalizowano potrzebę zbudowania referencyjnych baz w sytuacji niezgodności topologicznej dotychczasowych danych (Berus, Kołodziej, Olszewski 2007, Nałęcz 2007, Olszewski 2005, Gotlib 2005). Stosunkowo kompletne kompendium informacji o krajowych bazach danych stanowi opracowanie Gotliba, Iwaniaka i Olszewskiego (2007a, 2007b).

#### **2.4. Koncepcje skal pomiarowych w kartografii**

Obok przestrzennej reprezentacji danych geograficznych, w tworzeniu danych przestrzennych istotne są atrybuty (Longley i in 2006: 71). Z atrybutami oraz możliwościami wizualizacyjnymi atrybutów związane jest pojęcie skali pomiarowej.

Skala pomiarowa (Frączek 1981) jest pojęciem teorii pomiaru i oznacza szereg postaci  $\langle E, L, f \rangle$ , gdzie pierwsze dwa wyrazy stanowią elementy dwóch odpowiadających sobie systemów relacyjnych – empirycznego (E) i liczbowego (L), natomiast  $f$  jest funkcją pomiarową, która odwzorowuje te dwa systemy (przyporządkowuje liczby zdarzeniom empirycznym). System relacyjny jest pewnym zbiorem elementów (czyli wartości atrybutu) oraz relacji zachodzących między nimi. Rodzaj skali pomiarowej zależy od funkcji pomiarowej (Frączek 1981). Ostrowski (2008: 76) wskazuje, że skale pomiarowe zwane też są w literaturze poziomami pomiarowymi, a Żyszkowska (2000: 48) określa je mianem poziomów ujęcia atrybutów, przez co wyraźnie podkreśla wizualizacyjną istotę tego zagadnienia. Atrybuty zgromadzone w cyfrowej bazie danych są podstawą tworzenia wizualizacji na drodze transformacji kartograficznej (Longley i in. 2006) z zastosowaniem odpowiednio dobranych skal pomiarowych.

Podział informacji atrybutowej na ilościową i jakościową ma swoje odzwierciedlenie w wielu pracach kartograficznych dotyczących skal pomiarowych (Bertin 1983, Kraak, Ormeling 1998, Żyszkowska 2000, Korycka-Skorupa 2002a, 2002b, Chang 2004, Longley i in. 2006). Jako pośredni między tymi skalami został zaadaptowany w szeregu pracach kartograficznych poziom porządkowy (Ostrowski 2008, Korycka-Skorupa 2002a). Frączek w swoim artykule (1981) poddaje w wątpliwość wydzielenia charakterystyk jakościowych i ilościowych w skalach pomiarowych. Autorka przytacza przykłady prezentacji cechy ilościowej, która przy zmianie w legendzie wartości liczbowych na rangi przyjmuje charakter prezentacji cechy jakościowej, a także przykład odwrotny. Jej zdaniem skale pomiarowe tworzą szereg od najsłabszej do najmocniejszej, różniące się szczegółowością informacji.

Jeśliby przyjąć powyższe podejście, to w tab. 2.1. „najślabszą” skalą pomiarową będzie **dychotomiczny poziom ujęcia atrybutów** (Żyszkowska 2000). Ta kategoria została wydzielona jako informacja o występowaniu lub niewystępowaniu obiektu w przestrzeni geograficznej. Ponieważ zastosowanie tego poziomu nie wymaga istnienia atrybutów (chyba, że zjawisko ma zdefiniowane w bazie danych pole typu logicznego), wymyka się on spod podziału na skale jakościowe i ilościowe. Kolejne poziomy są już przyporządkowywane przez większość autorów do jakościowej lub ilościowej kategorii, w zależności od charakteru atrybutów (tab. 2.1.).

Wśród skal właściwych atrybutom jakościowym najczęściej wymieniane są skale nominalne, klasyfikacyjne i porządkowe. W **skali nominalnej** każdy z obiektów jest identyfikowany indywidualnie na podstawie nazwy własnej lub identyfikatora. **Skala klasyfikacyjna** wykorzystuje funkcję pomiarową, która każdemu obiektowi przypisuje określoną wartość, a obiekty geograficzne grupuje się zgodnie z tym podziałem. Ta skala w podziale Żyszkowskiej (2000) ma swój odpowiednik w **poziomie morfologicznym**. Wykorzystując **skalę porządkową** tworzy się klasy obiektów wzajemnie wykluczające się i stanowiące grupy obiektów uporządkowanych ze względu na określoną własność. Jak już wspomniano niektórzy autorzy (Ostrowski 2008, Korycka-Skorupa 2002a) wyodrębniają skalę porządkową jako pośrednią między grupą skal jakościowych i ilościowych. Natomiast Bernhardsen (2002) oraz Harvey (2008) klasyfikują poziom porządkowy jako najmniej dokładny poziom grupy ilościowej. Żyszkowska (2000) w grupie jakościowej poziomów ujęcia atrybutów umieszcza jeszcze **poziom hierarchiczny**, w którym klasy tworzy się w oparciu o zasady klasyfikacji lub typologii. Hierarchia polega na tym, że każdy obiekt niższego rzędu należy do obiektu wyższego rzędu.

Skale pomiarowe związane z cechami ilościowymi również nie są zgodnie opisywane przez różnych autorów (tab. 2.1.). Największa zbieżność dotyczy występowania **poziomu interwałowego**, w którym wartości liczbowe są grupowane w klasy zdefiniowane jako przedziały liczbowe. Powszechnie w klasyfikacjach występuje również **poziom ilorazowy**. Poziom ten posiada naturalny punkt zerowy oznaczający brak zjawiska, przez co różnice między poszczególnymi wartościami mają charakter bezwzględny. Tak też należy rozumieć wymienioną przez Bielecką (2006) **skalę bezwzględną**, wskazany przez Żyszkowską (2000) **skalarny poziom ujęcia atrybutów** oraz **wartości bezwzględne** w ujęciu poziomu ilościowego według Ostrowskiego (2008). Trzecią skalą liczbową, wyróżnioną przez Kraaka i Ormelinga (1998) jest **skala wskaźnikowa** wykorzystująca pewne miary wartości, jak np. gęstość czy średnia. Wydaje się, że odpowiednikiem tej skali jest u Żyszkowskiej (2000)

**poziom wagowy**, a u Ostrowskiego (2008) atrybuty o charakterze wskaźnikowym zawarte w poziomie ujęcia ilościowego – **wartościach względnych**. Zakwalifikowanie przez Bernhardsena (2002) i Harveya (2008) poziomu porządkowego do grupy ilościowej zasygnalizowano już przy omawianiu skal jakościowych.

**Tabela 2.1. Zestawienie skal pomiarowych (poziomów pomiarowych, poziomów ujęcia atrybutów) według wybranych autorów**  
**Źródło: opracowanie własne**

ATRYBUTY			AUTOR (uwagi)	
IŁOŚCIOWE	JAKOŚCIOWE			
Skale pomiarowe (s)/poziomy pomiarowe (p)/poziomy ujęcia atrybutów (pu)				
- s. absolutna - s. ilorazowa - s. interwałowa - s. porządkowa - s. klasyfikacyjna - s. nominalna			Frączek (1981) (autorka neguje podział skal na jakościowe i ilościowe, wskazuje na możliwość określenia skal "mocniejszych" i „słabszych”)	
- s. interwałowa - s. wskaźnikowa		- s. nominalna - s. porządkowa	Kraak, Ormeling (1998)	
- pu. wagowy - pu. porządkowy - pu. interwałowy - pu. skalarny		- pu. nominalny - pu. hierarchiczny - pu. porządkowy - pu. morfologiczny	Ujęcie atrybutów	Żyszkowska (2000)
- pu. dychotomiczny			Obecność obiektów	
- p. ilościowy	- p. porządkowy	- p. jakościowy	Korycka-Skorupa (2002)	
- s. interwałowa - s. ilorazowa		- s. nominalna - s. porządkowa	Chang (2004)	
- p. porządkowy - p. interwałowy - p. wskaźnikowy		- p. nominalny	Bernhardsen (2002) Harvey (2008)	
- s. interwałowa - s. bezwzględna		- s. nominalna - s. porządkowa	Bielecka (2006)	
- pu. ilościowy - wartości absolutne - pu. ilościowy - wartości względne		- pu. porządkowy	- pu. jakościowy	Ostrowski (2008)
- pu. ilościowy		- pu. jakościowy	Obiekty indywidualne	

Zagadnienie skal pomiarowych, wyrosłe na gruncie tradycyjnej kartografii, w nowszych pracach jest konfrontowane z atrybutami nieprzestrzennymi bazy danych (Bielecka 2006, Ostrowski 2008). Nowoczesne podejścia klasyfikacyjne (Żyszkowska 2000, Ostrowski 2008) wymagają dyskusji i podjęcia nowych badań (Ostrowski 2008).

Skale pomiarowe w dobie kartografii cyfrowej są ściśle związane z danymi atrybutowymi. Dane opisowe można sklasyfikować ogólnie w cztery typy: logiczne, znakowe, numeryczne oraz daty i czasu (Stones, Matthew 2003). **Typ logiczny** przechowuje

informacje dotyczące występowania lub niewystępowania zjawiska/cechy. W przypadku braku informacji pojawia się wartość NULL. **Typ danych znakowych** może być pojedynczym znakiem, ciągiem znaków o stałej długości lub ciągiem znaków o określonej maksymalnej ilości znaków. W standardowych bazach ilość znaków nie przekracza liczby 255. **Dane typu numerycznego** mogą być zapisane w różny sposób, w zależności od charakteru informacji. Mogą być to liczby całkowite, zmiennoprzecinkowe (rzeczywiste) lub dziesiętne. **Dane daty i czasu** mogą przechowywać w ściśle zdefiniowanym formacie informacje dotyczące dat lub czasu albo obu danych łącznie. Podtypy tych czterech podstawowych typów danych są tworzone w poszczególnych aplikacjach bazodanowych oraz programach geoinformacyjnych, a ich obecność zależy od celu tworzenia bazy i/lub aplikacyjnego profilu oprogramowania (por. tab. 2.2.).

**Tabela 2.2. Podtypy danych atrybutowych w wybranych aplikacjach bazodanowych oraz geoinformacyjnych**  
*Źródło: opracowanie własne*

Typ danych	MySQL	Ms Access (PL)	SQL Server	ArcGIS	MapInfo Pro
<b>LOGICZNE</b>	BOOL	Tak/nie	BIT	-	Logiczne
<b>ZNAKOWE</b>	CHAR CHAR(N) VARCHAR (N)	Tekst Memo Hiperłącze Obiekt OLE	BINARY VARBINARY CHART VARCHAR IMAGE SYSNAME TEXT	TEXT	Tekstowe
<b>NUMERYCZNE</b>	TINYINT SMALLINT MEDIUMINT INT BIGINT FLOAT (P) DOUBLE (D, N) DECIMAL (D, N)	Waluta Liczba (bajt) Liczba (liczba całkowita) Liczba (liczba całkowita długa) Liczba (pojedyncza precyzja) Liczba (podwójna precyzja) Liczba (dziesiętna)	TINYINT SMALLINT INT SMALLMONEY MONEY FLOAT REAL	SHORT LONG FLOAT DOUBLE	Małe liczby Liczby całkowite Zmiennopozycyjne Dziesiętne
<b>DATY I CZASU</b>	DATE TIME TIMESTAMP DATETIME	Data/godzina (data: ogólna, długa, średnia, krótka; godzina: długa, średnia, krótka)	DATETIME SMALLDATETIME TIMESTAMP	DATE	Daty Czasu Daty/czasu
<b>INNE</b>	BLOOB	Autonumeracja	-	BLOOB OBJECTID GUID GEOMETRY RASTER (HYPERLINK)	

## **2.5. Uwarunkowania tworzenia polskich baz danych przestrzennych**

Bodźcem rozwoju baz danych przestrzennych w Polsce było coraz szersze wykorzystywanie technik komputerowych przy produkcji seryjnych map tematycznych. Prekursorami tego kierunku były Główny Urząd Geodezji i Kartografii oraz Państwowy Instytut Geologiczny. Od roku 1994 równoległe do analogowej powstawała cyfrowa postać Mapy Sozologicznej Polski (Żynda 2004), a w 1997 rozpoczęto również produkcję seryjnych map hydrograficznych w skali 1:50000 w oparciu o technologie geoinformacyjne (Kaniecki 2004). Mapy te wcześniej powstawały wyłącznie w technologii analogowej. Znowelizowane w roku 1997 wytyczne techniczne K-3.4 (mapa hydrograficzna) i K-3.6 (mapa sozologiczna) obligowały do przygotowywania przez wykonawców map w kroju arkuszowym w oparciu o oprogramowanie MapInfo Professional. W połowie lat 90. XX wieku w Państwowym Instytucie Geologicznym zaczęto realizować cyfrowe wersje Szczegółowej Mapy Geologicznej Polski (1994; Gogołek 2007) oraz Mapy Hydrogeologicznej Polski (1996; Herbich i in 2004), a także Mapy Geologiczno-Gospodarczej Polski (1997; Sikorska-Maykowska 2004). Wcześniej mapy te były produkowane metodami analogowymi.

Niezależnie od wczesnych produkcji cyfrowych map tematycznych, rozpoczęto prace koncepcyjne i wdrożeniowe nad wielkoskalową mapą zasadniczą w systemie geoinformacyjnym. Powstała w GUGiK nowelizacja instrukcji technicznej K-1 (1995), która poprzez kolejne modyfikacje przyczyniła się do ustalenia standardu mapy zasadniczej.

W latach 90. XX wieku szeroko prowadzono prace przy innej bazie danych wielkoskalowych – Leśnej Mapie Numerycznej. Powstała ona jako przestrzenne rozwinięcie tabel Systemu Informatycznego Lasów Państwowych (Olenderek 2000). W roku 1998 opracowano wstępne założenia techniczne dla wykonawców Leśnej Mapy Numerycznej (Zarządzenie nr 23 Dyrektora Generalnego Lasów Państwowych), jednak pełny standard zatwierdzono w 2001 roku (Zarządzenie nr 74 Dyrektora Generalnego Lasów Państwowych z dnia 23 sierpnia 2001 r. w sprawie zdefiniowania standardu leśnej mapy numerycznej dla poziomu nadleśnictwa oraz wdrażania systemu informacji przestrzennej w nadleśnictwach).

W połowie lat 90. XX wieku w Instytucie Geodezji i Kartografii powstała cyfrowa baza danych pokrycia terenu dla obszaru Polski. Została ona zrealizowana w ramach europejskiego projektu CORINE Land Cover 90 (Ciołkosz, Bielecka 2005). Projekt polegał na opracowaniu wektorowej mapy form pokrycia terenu na podstawie interpretacji zdjęć satelitarnych z lat 1989-1993. Nominalna skala opracowania wynosi 1:100 000. Kartowanie zostało powtórzone w oparciu o zdjęcia satelitarne z lat 1999-2001 (CLC 2000) oraz ostatnio

w 2006 roku (CLC 2006). Bazy CLC2000 i CLC2006 są uzupełnione warstwami zmian w pokryciu terenu w badanych przedziałach czasowych.

Wśród inicjatyw tworzenia baz tematycznych należy wspomnieć o realizowanym w Instytucie Meteorologii i Gospodarki Wodnej projekcie tworzenia nowej mapy podziału hydrograficznego. Produkt cyfrowy tego przedsięwzięcia został wdrożony pod nazwą Mapy Podziału Hydrograficznego Polski. Od wielu lat prowadzona jest także sukcesywna cyfryzacja Mapy Glebowo-Rolniczej w skali 1:5 000 i 1:25 000 w Instytucie Upraw i Nawożenia w Puławach, a także w Urzędach Marszałkowskich niektórych województw.

Bazy danych topograficznych zaczęły w Polsce powstawać po roku 2000. W wyniku akcesji Polski do NATO Zarząd Geografii Wojskowej Wojska Polskiego zrealizował w latach 2000-2004 Mapę Wektorową Poziomu Drugiego (Sobczyński 2004, Bac-Bronowicz i in. 2007). Powstała ona w wyniku cyfrowania dotychczasowej mapy wojskowej w skali 1:50 000 zgodnie ze standardem DIGEST (Digital Geographic Information Exchange Standard) będącymi ogólnym standardem map cyfrowych NATO. Na podstawie Mapy Wektorowej Poziomu Pierwszego (1:250000) GUGiK we współpracy z UNEP GRID oraz kilkoma resortami centralnymi opracował w roku 2003 jednolitą dla całego kraju Bazę Danych Ogólnogeograficznych. Powstała ona w czterech szeregach skalowych: 1:250 000, 1:500 000, 1:1 000 000 i 1:4 000 000. Od 2003 roku realizowane jest ponadto w kroju arkuszowym przedsięwzięcie tworzenia Bazy Danych Topograficznych (TBD) będącej odpowiednikiem mapy topograficznej 1:10000 (Gotlib, Olszewski, Iwaniak 2007a).

Przedstawiony powyżej rys historyczny obejmuje początkowe etapy realizacji głównych projektów krajowych, które dostarczyły najciekawszych, z punktu widzenia badań gośrodowiskowych, danych. Opracowania te są przygotowywane w układzie arkuszowym (tak, jak mapy analogowe) albo w zakresie jednostek przestrzennych (np. obręb geodezyjny lub nadleśnictwo) lub też jako bazy ciągłe, obejmujące większe obszary (województwo, kraj). Większość z tych opracowań po etapie przygotowania danych w kroju arkuszowym doczekała się scalenia na poziomie województw (np. Baza HYDRO i SOZO, mapa gośrodowiskowa) lub kraju (np. VML2). Stan pokrycia terenu poprzez poszczególne bazy danych przestrzennych jest zróżnicowany i zmienny w czasie. Część z tych opracowań zostało zrealizowanych dla całego kraju (BDO, VML2, MPHP), niektóre są nadal w trakcie realizacji lub reambulacji, inne znów są zawieszono, jeszcze inne doczekały się kontynuacji w postaci nowego produktu z uzupełnionymi dodatkowymi danymi (MGPP – MGP) lub danymi znacznie rozszerzającymi dotychczasowy zakres (MHgP nowej edycji) albo też zostały przekształcone w uproszczone kompilacje z dotychczasowych opracowań (Mapa

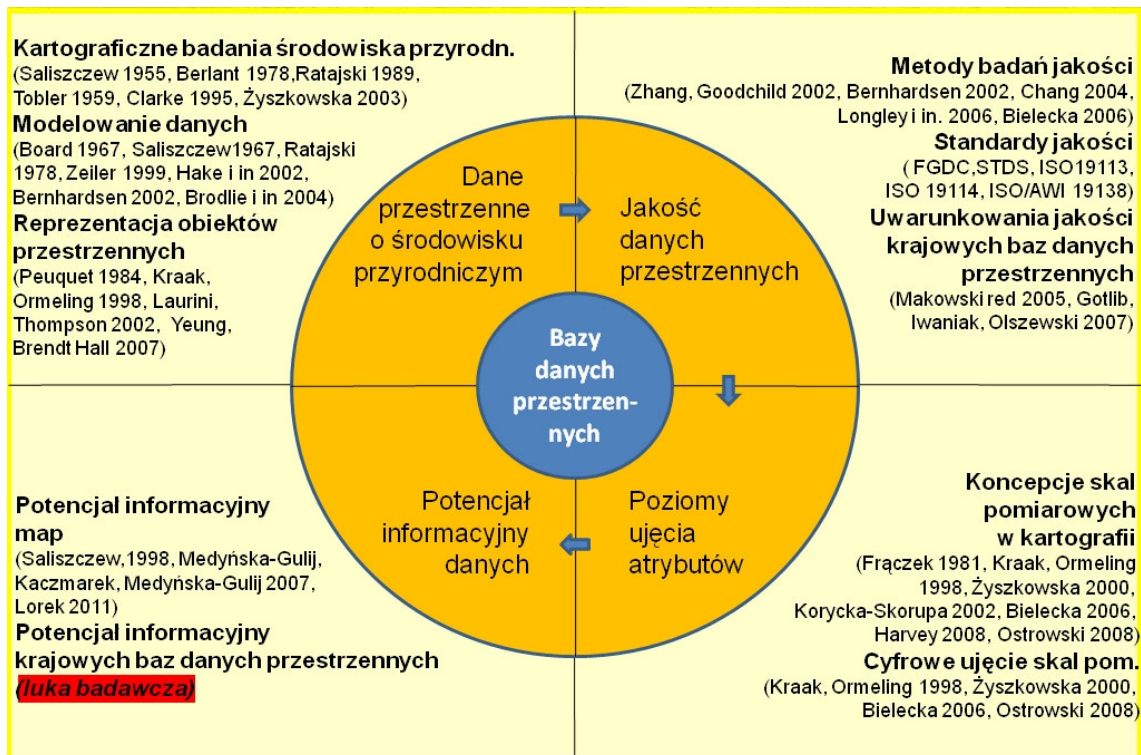
Litostratygraficzna Polski z SMGP). W efekcie ponad 15. letniej historii wdrażania cyfrowych baz danych ugruntował się jednak pewien model danych bazowych, częściowo zachodzących na siebie semantycznie oraz niespójnych znaczeniowo i topologicznie.

## 2.6. Identyfikacja problemów badawczych

Na tle zaprezentowanych powyżej teorii naukowych zarysowują się problemy badawcze wymagające rozwiązania w ramach niniejszej rozprawy. Dotyczą one następujących zagadnień (ryc. 2.10.):

1. Roli baz danych przestrzennych jako potencjalnego źródła informacji geograficznej dla opracowań naukowych i użytkarnych.
2. Metodyki badania jakości danych przestrzennych.
3. Propozycji klasyfikacji skal pomiarowych w kartografii uwzględniających specyfikę cyfrowych danych przestrzennych.
4. Koncepcji potencjału informacyjnego baz danych przestrzennych.

Podjęcie tej problematyki w niniejszej rozprawie zmierza do wypełnienia luki badawczej w stosunku do stanu wiedzy prezentowanego w dotychczasowych publikacjach.



Ryc. 2.10. Problematyka badawcza rozprawy  
Źródło: opracowanie własne

### 2.6.1. Dane cyfrowe jako źródło informacji o środowisku przyrodniczym

Z zaprezentowanego przeglądu literatury wynika, że zagadnienie stosowania kartograficznych metod badawczych w badaniach środowiska przyrodniczego jest obecnie silnie powiązane z technologiami geoinformacyjnymi. Sprzyja temu coraz większa dostępność cyfrowych baz danych tworzonych w kraju na poziomie rządowym, samorządowym oraz resortowym. Szczególnie obiecujące są działania w zakresie tworzenia ogólnonarodowej infrastruktury informacji przestrzennej w ramach Dyrektywy INSPIRE.

Modelowanie informacji przestrzennej, zdefiniowane w dokumentach ogólnych (Dyrektywa INSPIRE, Ustawa IIP), jak i przepisach wykonawczych (Instrukcje i Wytyczne Techniczne) wpisuje się poprzez definicje obiektów w koncepcję geokomponentów środowiska (Richling 1992, 2007). Każdą warstwę cyfrowych danych przestrzennych można przyporządkować do jednego z geokomponentów: atmosfery, hydrosfery, morfogeosfery, biosfery, pedosfery, litosfery, hydrogeosfery czy też antroposfery. Zatem dane cyfrowe służące badaniu środowiska przyrodniczego można porządkować do poszczególnych geokomponentów. Na ogół sprzyja temu podejściu struktura fizyczna dostępnych baz danych topograficznych i tematycznych.

Bazy danych przestrzennych, o których była mowa w rozdziale 2.5., w przeważającej większości powstawały jako źródło docelowych opracowań kartograficznych, udostępnianych analogowo (np. Mapa Hydrograficzna, Mapa Sozologiczna, Szczegółowa Mapa Geologiczna, Mapa Hydrogeologiczna, Mapa Geośrodowiskowa). Przyjęty w tych bazach model fizyczny danych pozwala na traktowanie warstw tematycznych w kategorii elementów cyfrowego modelu krajobrazu, stanowiącego przedkartograficzną reprezentację środowiska przyrodniczego. Świadczy o tym choćby niezgodność struktury warstw tematycznych baz z semantyczną strukturą map powstałych w oparciu o te bazy czy też powszechna nadmiarowość atrybutów względem potrzeb wizualizacyjnych produktów mapowych. Wydaje się również, że cechy generalizacji obiektów w tych bazach cyfrowych w niewielkim stopniu dotyczą zagadnień generalizacji kartograficznej (np. sztucznego przesunięcia obiektów). Warto wspomnieć, że niektóre krajowe cyfrowe bazy danych powstają z założenia jako bazy geoinformacyjne, dla których aspekt kartograficzny opracowania nie był celem głównym. Wśród tych opracowań należy wymienić Mapę Podziału Hydrograficznego czy Leśną Mapę Numeryczną z bazą SILP, Ewidencję Gruntów i Budynków czy bazę Corine Land Cover. Te bazy z pewnością reprezentują cyfrowy model krajobrazu. Pomiedzy bazami pierwszego typu („kartograficznymi”) a drugiego typu („geoinformacyjnymi”) występuje grupa baz pośrednich, które mają spełniać funkcje kartograficzne, natomiast są rozbudowane

pod względem struktury bazodanowej. Do tej grupy należy zaliczyć np. Bazę Danych Ogólnogeograficznych i Mapę Wektorową Poziomu Drugiego.

Traktowanie wszystkich rodzajów baz danych przestrzennych powstałych w Polsce jako modelu danych krajobrazowych jest obarczone pewną nieścisłością związaną ze szcążkowym zastosowaniem metod generalizacji kartograficznej w części baz, niemniej pozwoli spojrzeć na ten zasób jako na potencjalne źródło informacji przestrzennej do kartograficznych badań środowiska przyrodniczego. W literaturze niejednokrotnie wskazywano na możliwość wykorzystania polskich baz danych przestrzennych w kartograficznych badaniach środowiska przyrodniczego (m. in. Bródka 2007, 2010, Dzikowska i in. 2007, Graf 2007, Kaczmarek 2007, 2010, Łowicki, Stępniewska 2007, Nałęcz 2007, Rudowicz-Nawrocka 2006, Kaczmarek, Medyńska-Gulij 2007, 2008, Medyńska-Gulij, Kaczmarek red 2007). Do tej pory nie pojawiła się jednak propozycja szerokiego wykorzystania tego zasobu, poparta systematyczną analizą materiału źródłowego.

#### 2.6.2. Jakość danych przestrzennych

Uwarunkowania jakości danych przestrzennych wynikają z przyjętego podejścia do modelowania obiektów przestrzennych, a w szczególności założeń modelu konceptualnego, przyjętej reprezentacji obiektów przestrzennych oraz stopnia generalizacji danych w bazie (por. rozdz. 2.2.). Czynniki mające wpływ na określenie jakości danych przestrzennych wiązać należy z takimi miarami, jak kompletność, spójność logiczna, dokładność położenia, dokładność czasu i dokładność tematyczna (Bielecka 2006). Dotyczą więc one trzech aspektów: geometrii, relacji oraz atrybutów (Bernhardsen 2002). Metodyka określania jakości danych przestrzennych była przedmiotem opracowań naukowych (Zhang, Goodchild 2002) oraz unormowań (EN ISO: 19114:2005 i ISO/AWI 19138:2002). Na tym tle brak jest opracowania podejścia do kompleksowego badania jakości danych przestrzennych zawartych w polskich bazach danych przestrzennych. Autor podjął próbę opracowania kompleksowych miar jakości, generowanych stosunkowo łatwo dla dużych zbiorów danych na drodze przetwarzania geoinformacyjnego.

#### 2.6.3. Skale pomiarowe w ujęciu danych cyfrowych

Zaprezentowane w przeglądzie literatury poglądy na stosowanie skal pomiarowych w większości odnoszą się do aspektu kartograficznego. Nieliczni autorzy (Żyszkowska 2000, Ostrowski 2008) odnoszą tę cechę bezpośrednio do baz danych. Tradycyjne jednak

klasyfikacje poziomów pomiarowych wiążą się bezpośrednio z metodami prezentacji kartograficznej (Medyńska-Gulij 2010).

Zastosowanie skal pomiarowych w bazach danych przestrzennych odnosi się do poszczególnych atrybutów określonych warstw tematycznych. W tym przypadku, oprócz znaczenia atrybutu w opisie cech obiektów, istotne są takie właściwości danych atrybutowych, jak: typ danych, dokładność wartości liczbowych, zastosowany słownik atrybutowy czy możliwości wtórnego przetwarzania danych. Jednocześnie należy zauważyć, że jedno pole atrybutowe może dać możliwość zastosowania różnych skal pomiarowych dla prezentacji kartograficznej, w zależności od potrzeb. Dlatego też zagadnienie klasyfikacji skal pomiarowych skłoniły autora do wypracowania innego podejścia, związanego bezpośrednio z danymi atrybutowymi.

#### 2.6.4. Potencjał informacyjny baz danych

Cyfrowy model krajobrazu zapisany w poszczególnych bazach w układzie komponentowym stanowi źródło informacji o środowisku przyrodniczym. Fizyczna reprezentacja danych, w przeważającej większości jako dyskretnych obiektów wektorowych, pozwala na szerokie stosowanie kartograficznych metod badawczych z wykorzystaniem technologii geoinformacyjnych. Problemem jest jednak określenie poziomu ufności do danych źródłowych (Zhang, Goodchild 2002), a więc pewnej miary ich dokładności. W sytuacji powszechnej redundancji danych w rozpatrywanych bazach (por. rozdz. 2.5.) powyższa kwestia wymaga ponadto uwzględnienia wzajemnego porównania danych tematycznie zbieżnych, a pochodzących z różnych baz danych. Celem takiego porównania powinna być optymalizacja wyboru danych źródłowych (Kaczmarek, Medyńska-Gulij 2008), a w szczególności nie tylko dokładność danych, ale i stopień spełniania oczekiwań w zakresie zawartości informacyjnej. Taka ocena ma zawsze charakter subiektywny (Bielecka 2006) i jest realizowana w kontekście konkretnego zapotrzebowania na dane przestrzenne.

W rozprawie został podjęty problem wyznaczania potencjału informacyjnego baz danych. Jest to zagadnienie, które nie doczekało się jeszcze kompleksowego opracowania. Potencjał informacyjny danych przestrzennych jest rozumiany w tej pracy jako miara zasobności modeli obiektów przestrzennych w atrybuty istotne z punktu widzenia użytkownika danych. Teoretyczne podstawy potencjału informacyjnego mogą stanowić podstawę praktycznej waloryzacji baz danych przestrzennych.

### **3. ANALIZA ZAWARTOŚCI BAZ DANYCH PRZESTRZENNYCH W KONTEKŚCIE MODELOWANIA RZECZYWISTOŚCI**

Uwarunkowania tworzenia krajowych baz danych przestrzennych zaprezentowano w rozdziale 2.5. Kontekst, w jakim powstawał ten zasób ma istotne znaczenie przy rozpatrywaniu jakości danych, stąd te uwarunkowania zostaną uwzględnione przy analizie zawartości baz oraz jakości danych.

Bazy danych przestrzennych, traktowane w niniejszej pracy jako przedmiot rozważań, stanowią modelowe ujęcie środowiska przyrodniczego w wybranym zakresie tematycznym. Dlatego analiza zawartości baz danych musi uwzględnić w pierwszej kolejności zagadnienia modelowania rzeczywistości. Modelowanie rzeczywistości przedstawiane jest w literaturze w różny sposób:

- a) jako odniesienie do transformacji kartograficznej oraz wyobrażenia produktu modelowania u odbiorcy (Board 1967, Ratajski 1983, Żyszkowska 2000);
- b) w szerokim kontekście, obejmującym poza modelem kartograficznym również model pierwotny oraz wyobrażenie efektu modelowania kartograficznego u odbiorcy (Hake i in. 2002);
- c) jako proces etapowego konstruowania bazy danych (Peuquet 1984, Zeiler 1999, Bernhardsen 2002).

Koncepcja Boarda (1967; por. ryc. 2.1., rozdział 2.2.1.) wskazuje na rolę specjalisty w cyklu modelowania oraz na analogię etapowego procesu tworzenia mapy do prowadzenia analizy (filtrowanie-selekcja rzeczywistości-przetwarzanie danych).

Z punktu widzenia analizowanych w pracy baz danych oraz ogólnego podejścia do tworzenia i przetwarzania informacji przestrzennej najbardziej przydatna do opisu baz danych jest koncepcja Zeilera (1999). Podejście to skupia się na określeniu cech fizycznego modelu danych, odpowiedniego zastosowanemu w badanych bazach danych przestrzennych. Wynikiem modelowania rzeczywistości będzie cyfrowy model krajobrazu (DLM) stanowiący przedkartograficzną reprezentację środowiska przyrodniczego. Obiekty powstałe na tym etapie będą charakteryzowały się jakością, która wynika z zastosowania generalizacji konceptualnej (Bernhardsen 2002). Wszystkie trzy etapy modelowania: konceptualny, logiczny i fizyczny są z reguły zdefiniowane w wytycznych technicznych i instrukcjach będących podstawą produkcji przedmiotowych baz danych. Stąd stały się one podstawą przeprowadzenia analizy tych baz danych przestrzennych w rozdziale 3.1. Określenie cech

modelu danych pozwoli na ustalenie poziomu teoretycznej, czyli zakładanej w wytycznych i instrukcjach technicznych, jakości danych przechowywanych w poszczególnych bazach danych. Jest to punkt wyjścia do oceny rzeczywistej jakości danych i ich waloryzacji.

Właściwa ocena baz danych pod kątem ich wykorzystania w badaniach środowiska przyrodniczego wymaga traktowania ich jako źródła informacji o komponentach środowiska przyrodniczego. Takie podejście ma na celu wskazanie roli, jaką poszczególne bazy mogą potencjalnie pełnić w przestrzennej charakterystyce komponentów środowiskowych. Rozdział 3 stanowi więc punkt wyjścia do określenia potencjalnej roli elementów baz danych (warstw i grup warstw tematycznych) w charakterystyce geograficznej komponentów środowiska przyrodniczego.

### **3.1. Założenia dotyczące analizy zawartości merytorycznej baz danych przestrzennych**

Analiza zawartości baz danych będzie poprawna, jeśli zapewni porównywalność poszczególnych produktów, szczególnie w odniesieniu do metod modelowania rzeczywistości, zakresu tematycznego oraz dokładności. Stąd głównym celem tego rozdziału jest stworzenie porównywalnego opisu baz danych pod względem założeń realizacyjnych, struktury tematycznej i jakości danych.

Przedstawiona poniżej analiza obejmuje produkty najistotniejsze z punktu widzenia komponentów środowiska przyrodniczego (tab. 1.1.). Ich istotność wynika zarówno z zakresu tematycznego warstw, jak i z wysokiego stopnia lub wręcz pełnego pokrycia danymi dla obszaru kraju.

Indywidualny opis każdej bazy danych przestrzennych ujęto jednolity, dwunastopunktowy schemat metadanych:

- I. **Nazwa bazy** – oficjalna nazwa bazy danych oraz skrót nazwy wykorzystywany w niniejszej publikacji
- II. **Merytoryczny cel tworzenia bazy** („przeznaczenie” wg EN ISO 19113) – ogólne zdefiniowanie treści bazy zawarte w przepisach wykonawczych dotyczących bazy lub w innym opisie źródłowym
- III. **Utylitarny cel tworzenia bazy** („wykorzystanie” wg EN ISO 19113) – wskazanie rzeczywistego wykorzystania bazy lub potencjalnego przeznaczenia bazy
- IV. **Właściciel i dysponent** – instytucja, która zleciła wykonanie bazy, była (jest) odpowiedzialna za realizację bazy i / lub dysponuje bazą (w tym udostępnia jej zasób)
- V. **Źródło opisu standardu bazy** („specyfikacja produktu” wg EN ISO 19113) – przepisy wykonawcze definiujące zasady opracowywania bazy danych

- VI. **Odniesienie przestrzenne** – wykorzystany układ poziomego odniesienia przestrzennego oraz odniesienie wysokościowe (jeśli zostało zdefiniowane), a także podstawowa jednostka przestrzenna opracowania bazy
- VII. **Skala, zasięg przestrzenny i okres realizacji** – skala liczbowa towarzysząca bazie opracowania kartograficznego lub, jeśli baza nie zawiera produktu kartograficznego, nominalna skala bazy wynikająca ze skali wykorzystanych materiałów źródłowych oraz poziomu generalizacji danych; zakres pokrycia powierzchni kraju opracowaniem oraz okres opracowania bazy
- VIII. **Format źródłowy i dystrybucyjny danych** – format komputerowy bazy wynikający z przyjętej technologii opracowania oraz powszechnie stosowany format komputerowy udostępnianych danych
- IX. **Grupy tematyczne danych** – zdefiniowane w przepisach wykonawczych zjawiska i obiekty posegregowane z punktu widzenia realizacji merytorycznego celu tworzenia bazy
- X. **Systemy identyfikacji obiektów** – zastosowane w bazie identyfikatory przestrzenne, pozwalające na przyłączanie danych z innych baz danych przestrzennych lub z zewnętrznych źródeł danych
- XI. **Źródła danych geometrycznych i atrybutowych** („pochodzenie” wg EN ISO 19113) – zdefiniowane w źródle opisu standardu danych rekomendowane materiały źródłowe
- XII. **Uwagi** – uzupełniające informacje o bazie istotne z punktu widzenia dokładności danych oraz obecności dodatkowych informacji opisowych, a także istotne publikacje związane z bazą.

Warto zauważyć, że opis taki zawiera w punkcie 2, 3 i 11 przeglądowe elementy jakości danych, ujętych w normie EN ISO19113, rozumianych jako składniki nieliczbowe dokumentujące jakość zbioru danych.

W poniższych opisach baz danych szczególnie istotne informacje znajdują się w punkcie 10 i 11. Dotyczą one systemu identyfikacji obiektów oraz źródeł danych geometrycznych i atrybutowych. Zastosowanie w bazie danych systemu identyfikatorów jest miarą otwartości bazy na inne bazy oraz rejestry urzędowe, które potencjalnie można przyłączać do obiektów przedmiotowej bazy. Natomiast opisane źródła danych geometrycznych i atrybutowych wskazują teoretyczny poziom jakości bazy w zakresie dokładności lokalizacyjnej obiektów oraz dokładności tematycznej danych.

### **3.2. Zawartość merytoryczna baz danych przestrzennych w ujęciu modelu danych**

Informacje o bazach danych przestrzennych zostały pozyskane na drodze studiowania literatury oraz szczegółowej analizy założeń do baz danych (instrukcje, wytyczne techniczne oraz inne przepisy wykonawcze). Spośród znacznej liczby opracowań literaturowych omawiających poszczególne bazy danych przestrzennych wybrano następujące opracowania: Baranowski (2004a i b), Borowicz, Krupski (2009), Gotlib, Iwaniak, Olszewski (2007a), Gotlib, Olszewski (2005), Graf (2004), Herbich (2004), Kaczmarek (2010), Kaczmarek, Medyńska-Gulij (2006, 2007a, 2008), Kaniecki (2004), Kurzeja, Szłapczyńska (2007), Okonek (2005), Olenderek (2010), Olszewski (2007, 2009), Olszewski, Stankiewicz (2004), Sikorska-Maykowska (2004, 2007), Sobczyński (2007), Żynda (2004). W pozyskaniu informacji o bazach korzystano także z oficjalnych informacji zamieszczonych w Internecie ([www.gugik.gov.pl](http://www.gugik.gov.pl), [www.codgik.gov.pl](http://www.codgik.gov.pl), [www.geoportal.gov.pl](http://www.geoportal.gov.pl), [www.wodgik.poznan.pl](http://www.wodgik.poznan.pl), [www.pgi.gov.pl](http://www.pgi.gov.pl), [www.imgw.pl](http://www.imgw.pl), [www.iung.pulawy.pl](http://www.iung.pulawy.pl) i inne). W analizie wykorzystano również następujące normy realizacyjne baz: DIGEST (2000), Wytyczne Techniczne TBD (2003), Wytyczne Techniczne K-3.4. (1997), Wytyczne Techniczne K-3.6. (1997), Wytyczne Techniczne GIS-3 (2005), Wytyczne Techniczne GIS-4 (2005), Instrukcja wydania i opracowania SMGP (1996), Instrukcja opracowania i komputerowej edycji MHgP (1999), Instrukcja opracowania i aktualizacji MGGP (2002), Instrukcja opracowania MGP (2005), Rozporządzenie MRRiB (2001), Instrukcja Techniczna G-5 (2003), Standard LMN (2005), CORINE LandCover Technical Guide (1994).

#### **BAZA DANYCH OGÓLNOGEOGRAFICZNYCH**

##### I. Nazwa bazy

Baza Danych Ogólnogeograficznych (BDO), poziom skalowy 1:250 000

##### II. Merytoryczny cel tworzenia bazy

Prowadzenie zasobu danych geograficznych w postaci spójnych baz danych

##### III. Utylitarny cel tworzenia bazy

Cyfrowe dane referencyjne do opracowań GIS na poziomie regionalnym i krajowym, produkcja kartograficzna

##### IV. Właściciel i dysponent

Główny Geodeta Kraju

#### V. Źródło opisu standardu bazy

Użytkownikom udostępniono wyłącznie strukturę atrybutów i system kodowania (w postaci tabel MS Excel)

#### VI. Odniesienie przestrzenne

Układ współrzędnych PUWG 1992/19, układ odniesień wysokościowych Kronsztadt; baza opracowywana w cięciu siatki geograficznej 1° x 1° (ryc. 4.1.) albo jako baza ciągła na poziomie województwa lub kraju

#### VII. Skala, zasięg przestrzenny i okres realizacji

Opracowanie jest wykonane w czterech szeregach skalowych: 1:250 000, 1:500 000, 1:1 000 000 i 1:4 000 000. Przedmiotem zainteresowania jest opracowanie w skali 1:250 000 obejmujące obszar całego kraju; bazę zrealizowano w okresie 2001-2003 z planem aktualizacji, do której nie doszło

#### VIII. Format źródłowy i dystrybucyjny danych

Dane są zapisane w strukturze warstw informacyjnych ArcInfo. Ponadto materiał jest dostępny w formacie .eps oraz jako wizualizacja kartograficzna w formacie .tif

#### IX. Grupy tematyczne danych

Podział administracyjny

Osadnictwo i obiekty antropogeniczne

Hydrografia

Rzeźba terenu

Transport

Pokrycie terenu

Obszary i obiekty chronione

Nazwy

Siatka geograficzna

#### X. Systemy identyfikacji obiektów

Baza jest sprzężona z branżowymi bazami danych: NTS (jednostki administracyjne), TERYT (jednostki administracyjne, miejscowości), ID\_IMGW (cieki i zbiorniki wodne), ID\_GDDP (drogi) oraz ID\_KSOCH (obiekty chronione powierzchniowe i punktowe)

#### XI. Źródła danych geometrycznych i atrybutowych

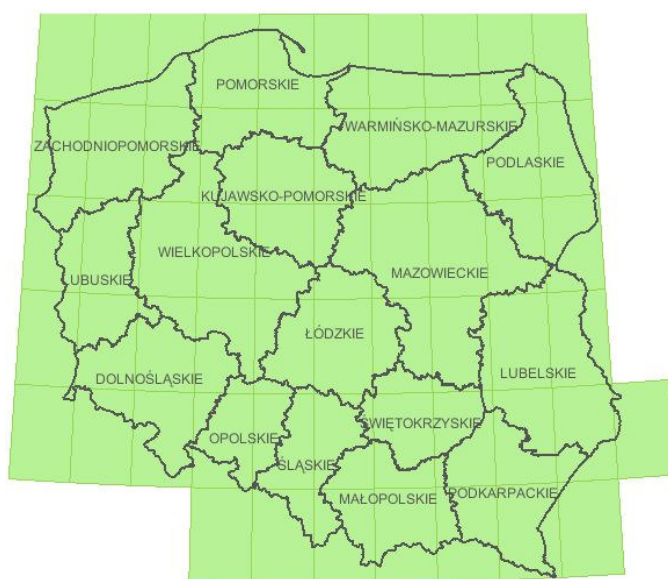
Dane geometryczne w podstawowym zakresie są przejęte z georelacyjnej bazy danych Mapy Wektorowej Poziomu Pierwszego (VML1), w zakresie granic

administracyjnych z Państwowego Rejestru Granic (PRG), a w odniesieniu do użytkowania ziemi z bazy CORINE Land Cover

Dane atrybutowe zostały pozyskane ze zintegrowanych z BDO (poprzez identyfikatory) baz takim, jak: Krajowy Rejestr Urzędowego Podziału Terytorialnego Kraju (TERYT, Główny Urząd Statystyczny), Komputerowa Mapa Podziału Hydrograficznego Polski (KMPHP, Instytut Meteorologii i Gospodarki Wodnej), Baza Danych Dróg Krajowych (BDDK, Generalna Dyrekcja Dróg Krajowych i Autostrad) oraz Krajowy System Obszarów Chronionych (KSOCh, Ministerstwo Środowiska)

## XII. Uwagi

Baza została opracowana w latach 2001-2003 w Centrum Informacji o Środowisku UNEP/GRID – Warszawa z założeniem regularnej aktualizacji . Niestety aktualizacja ta nie jest prowadzona.



*Ryc. 3.1. Podstawowe jednostki opracowania Bazy Danych Ogólnogeograficznych w skali 1:250 000  
Źródło: serwisy.codgik.gov.pl, stan na 06.2011*

## **MAPA WEKTOROWA POZIOMU DRUGIEGO**

I. Nazwa bazy

Mapa Wektorowa Poziomu Drugiego (VML2)

II. Merytoryczny cel tworzenia bazy

Stworzenie spójnej bazy obiektów topograficznych dla całego kraju

III. Utylitarny cel tworzenia bazy

Produkcja kartograficzna, referencja baz cyfrowych, zastosowania wojskowe

#### IV. Właściciel i dysponent

Zarząd Geografii Wojskowej Wojska Polskiego; współwłaścicielem oraz dysponentem jest również Główny Geodeta Kraju

#### V. Źródło opisu standardu bazy

DIGEST – *Digital Geographic Information Exchange Standard*, a w zakresie realizacji produktu w warunkach polskich również szczegółowe opisy zasad wprowadzania i kodowania oraz struktury atrybutowej warstw tematycznych, dostępne w postaci elektronicznej (.doc).

#### VI. Odniesienie przestrzenne

Układ współrzędnych WGS84, Układ odniesień wysokościowych Kronsztadt; baza opracowywana w cięciu arkuszowym Międzynarodowej Mapy Świata dla skali 1:50 000 (po dwa sąsiadujące arkusze), dostępna również jako baza ciągła na poziomie województwa lub mniejszych jednostek terytorialnych

#### VII. Skala, zasięg przestrzenny i okres realizacji

1:50 000; baza została zrealizowana w okresie 2002-2004 dla obszaru całego kraju, planowana aktualizacja w postaci VML2+ nie została dotąd zrealizowana

#### VIII. Format źródłowy i dystrybucyjny danych

Źródłowy format VPF (*vector product format*); spotykany format dystrybucyjny: VPF, shp

#### IX. Grupy tematyczne danych

AERO (informacje lotnicze)  
BND (granice)  
CUL (obiekty socjalno-kulturalne)  
ELEV (rzeźba terenu)  
GEN (ogólne)  
HYDRO (hydrografia)  
IND (przemysł)  
PHYS (fizjografia)  
TRANS (transport)  
VEG (roślinność)

#### X. Systemy identyfikacji obiektów

Jedyny identyfikator zgodny z bazami krajowymi to atrybut IDN warstwy „obszar administracyjny” zgodny z numeracją TERYT; pozostałe zastosowane identyfikatory są zgodne z systemem DIGEST

## XI. Źródła danych geometrycznych i atrybutowych

Dane geometryczne. Źródłem danych geometrycznych były podkłady topograficzne wojskowych map w skali 1:50 000. Izobaty są pozyskane dla większych zbiorników wodnych z Atlasu jezior Polski (IMGW). Lokalizacja punktów osnowy geodezyjnej jest pozyskana bezpośrednio z wartości współrzędnych płaskich przypisanych im w katalogu współrzędnych punktów geodezyjnych. Przebieg granic kraju jest zgodny z Państwowym Rejestrem Granic (PRG).

Dane atrybutowe. Atrybuty form ochrony przyrody są pozyskane z Krajowego Systemu Obszarów Chronionych (KSOCh); parametry zbiorników wodnych są pozyskane z Atlasu jezior Polski (IMGW) oraz z powstałej w wojsku Bazy danych jezior Polski; charakterystyki rzek i brodów zostały pozyskane lub zaktualizowane w terenie; opisy cech drzewostanu wymagały aktualizacji terenowej; dane opisowe dotyczące dróg i mostów są częściowo uzupełnione z bazy BDDK Generalnej Dyrekcji Dróg Krajowych i Autostrad, a po części pozyskane z map wojskowych i uaktualnione w terenie; sieci energetyczne zostały uzupełnione atrybutami z Albumu Sieci Elektroenergetycznych (1996-97), a przy opisie linii kolejowych korzystano z Albumu Sieci Kolejowej RP wykonanej przez służby topograficzne w latach 1996-97.

## XII. Uwagi

VML2 charakteryzuje się spójnością topologiczną oraz bogactwem atrybutów. Istotne byłoby wykonanie aktualizacji bazy w oparciu o materiały teledetekcyjne, planowanej od 2004 roku jako VML2+.

## **BAZA DANYCH TOPOGRAFICZNYCH**

### I. Nazwa bazy

Baza Danych Topograficznych (Topograficzna Baza Danych; TBD)

### II. Merytoryczny cel tworzenia bazy

Źródło podstawowych danych o obiektach topograficznych o charakterze referencyjnym

### III. Utylitarny cel tworzenia bazy

Referencyjna baza obiektów topograficznych dla opracowań GIS.

Produkcja kartograficzna

### IV. Właściciel i dysponent

Główny Geodeta Kraju

#### V. Źródło opisu standardu bazy

Wytyczne Techniczne. Baza danych topograficznych (TBD). GUGiK, Warszawa 2003

#### VI. Odniesienie przestrzenne

Układ współrzędnych 1992, poziom odniesienia Kronsztadt, baza jest realizowana w cięciu arkuszowym 1:10 000 Międzynarodowej Mapy Świata

#### VII. Skala, zasięg przestrzenny i okres realizacji

1:10000, baza obejmuje znaczne obszary woj. mazowieckiego oraz kujawsko-pomorskiego, na pozostałych obszarach TBD ma jeszcze bardzo ograniczony zasięg (ryc. 4.2.)

#### VIII. Format źródłowy i dystrybucyjny danych

Dla bazy danych topograficznych jest to format .gml.

Dla numerycznego modelu terenu format ASCII (dane źródłowe), TIN i TTN (dane archiwalne) oraz GRID ASCII (dane użytkowe).

Dla ortofotomapy stosuje się format geoTIFF bez kompresji.

Format dystrybucyjny danych jest zależny od zapotrzebowania odbiorcy i możliwości technicznych WODGiK.

#### IX. Grupy tematyczne danych

##### Wektorowa baza danych topograficznych

Hydrografia

Drogi

Koleje

Uzbrojenie terenu

Roślinność

Zabudowa

Granice (jednostki podziału terytorialnego)

Rzeźba terenu

##### Baza NMT (numerycznego modelu rzeźby terenu)

##### Baza ORTOFOTO (ortofotomapa)

#### X. Systemy identyfikacji obiektów

Wytyczne techniczne nie wskazują bezpośrednio na obligatoryjność identyfikatorów z wyjątkiem numeru TERYT na poziomie gminy i wyżej. Brak dostępu do Opracowań TBD nie pozwolił na zweryfikowanie tej informacji

## XI. Źródła danych geometrycznych i atrybutowych

Głównym źródłem danych jest ortofotomapa cyfrowa uzupełniana informacjami pozyskanymi z materiałów źródłowych i kartowania terenowego

## XII. Uwagi

Marszałkowie z szeregu województw zlecają w ostatnim okresie elementy TBD dla większych obszarów (np. warstwę dróg, warstwę budynków itp.). Koncepcja tworzenia uproszczonego TBD nosi nazwę Bazy Danych Obiektów Topograficznych (BDOT)



*Ryc. 3.2. Stopień pokrycia Topograficzną Bazą Danych – pełna treść  
Źródło: serwisy.codgik.gov.pl, stan na 06.2011*

**A****B****C****D****E**

**Ryc. 3.3. Stopień pokrycia wybranymi elementami Bazy Danych Obiektów Topograficznych**  
**A – ciekі wodne, B – obszary wodne, C – drogi, D – linie kolejowe, E - budynki.**  
**Źródło: serwisy.codgik.gov.pl, stan na 06.2011**

## **MAPA HYDROGRAFICZNA POLSKI**

### I. Nazwa bazy

Mapa Hydrograficzna Polski (MHP)

### II. Merytoryczny cel tworzenia bazy

Opracowanie tematyczne opisujące przyrodnicze uwarunkowania obiegu wody w powiązaniu ze środowiskiem geograficznym, jego zainwestowaniem i przekształceniem

### III. Utylitarny cel tworzenia bazy

Produkcja kartograficzna

### IV. Właściciel i dysponent

Główny Geodeta Kraju

### V. Źródło opisu standardu bazy

Wytyczne Techniczne K-3.4. System Informacji o Terenie. Mapa Hydrograficzna Polski, skala 1:50 000 w formie analogowej i numerycznej. GUGiK 1997

Wytyczne Techniczne GIS-3 Mapa Hydrograficzna Polski, skala 1:50000 w formie analogowej i numerycznej. GUGiK 2005

### VI. Odniesienie przestrzenne

Układ współrzędnych PUWG 1992/19 (niektóre opracowania wykonane w układzie 1942; por. ryc. 4.3. A), układ odniesień wysokościowych Kronsztadt; baza opracowywania w cięciu arkuszowym Międzynarodowej Mapy Świata dla skali 1:50 000, dla niektórych województw dostępna również jako baza ciągła na poziomie województwa lub mniejszych jednostek terytorialnych

### VII. Skala, zasięg przestrzenny i okres realizacji

1:50 000; okres realizacji od 1997 do 2008; do dziś brak pełnego pokrycia kraju opracowaniem; MHP jest obecnie opracowana dla woj. zachodniopomorskiego, lubuskiego, wielkopolskiego, dolnośląskiego, śląskiego, małopolskiego i świętokrzyskiego oraz dla części woj. pomorskiego, lubelskiego, łódzkiego i pojedynczych arkuszy dla woj. kujawsko-pomorskiego (ryc. 4.3.).

### VIII. Format źródłowy i dystrybucyjny danych

Źródłowy format: MapInfo TAB (pozwala na przechowywanie obiektów punktowych, liniowych i powierzchniowych w jednej warstwie); spotkane formaty dystrybucyjne: MapInfo TAB, ESRI shape

## IX. Grupy tematyczne danych

Topograficzne działy wodne,  
Wody powierzchniowe,  
Wyływy wód podziemnych,  
Wody podziemne pierwszego poziomu,  
Przepuszczalność gruntów,  
Zjawiska i obiekty gospodarki wodnej,  
Punkty hydrometryczne pomiarów stacjonarnych.

## X. Systemy identyfikacji obiektów

W odniesieniu do Wytycznych z 1997 r. – wyłącznie wewnętrzny system identyfikacji obiektów zgodny z czterema bazami tabelarycznymi prowadzonymi dla MHP na poziomie krajowym: DBARKUSZ (nazw arkuszy mapy), DBMIEJSC (nazwy miejscowości), DB\_OB\_UC (nazwy zakładów stanowiących uciążliwość dla wód), DBWODY (nazwy wód powierzchniowych); ponadto dla zrzutów ścieków oraz miejsc pomiaru przepływów obiekty na tych warstwach posiadają identyfikatory (na poziomie arkusza) zgodne z numeracją zastosowaną w komentarzu do arkusza

W odniesieniu do Wytycznych z 2005 r. – zlewnie, ciekie nazwane oraz zbiorniki wodne posiadają identyfikator zgodny z tabelą BAZA\_WODY, oczyszczalnie ścieków posiadają REGON użytkownika oraz TERYT miejscowości, zrzuty ścieków posiadają identyfikator ciekie do którego ścieki są zrzucane (ID z tabeli BAZA\_WODY) oraz REGON zakładu zrzucającego ścieki; gminy posiadają 7-cyfrowy TERYT, powiaty posiadają 4-cyfrowy TERYT, województwa posiadają 2-cyfrowy TERYT; do tego prowadzone są trzy tabelaryczne bazy uzupełniające: BAZA\_TERYT, BAZA\_REGON i BAZA\_WODY; ta ostatnia jest kontynuacją bazy wg wytycznych z 1997 r., lecz jest sukcesywnie uzupełniana identyfikatorami hydrograficznymi wg IMGW; ponadto dla zrzutów ścieków oraz miejsc pomiaru przepływów obiekty na tych warstwach posiadają identyfikatory (na poziomie arkusza) zgodne z numeracją zastosowaną w komentarzu do arkusza

## XI. Źródła danych geometrycznych i atrybutowych

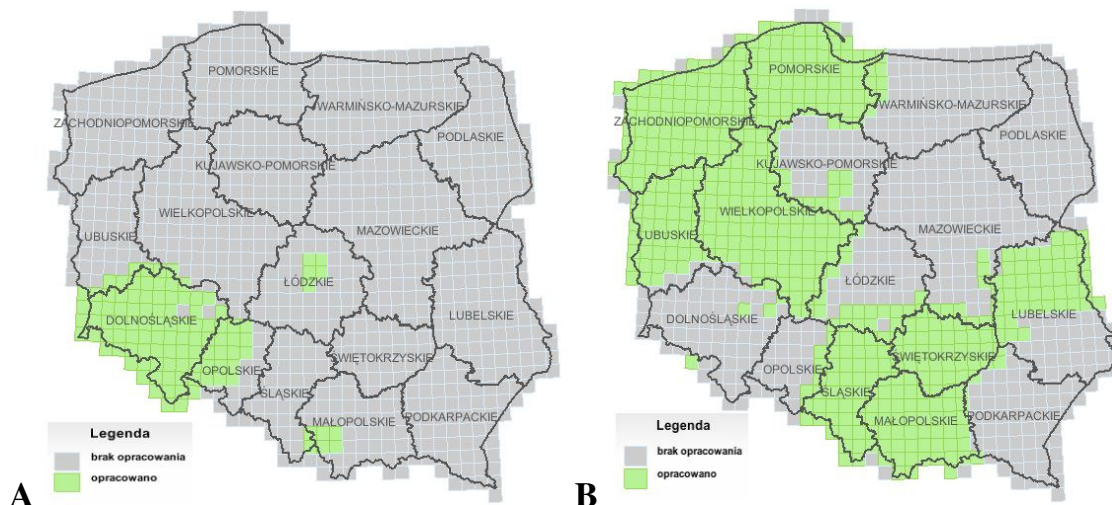
Wytyczne definiują ogólnie dane źródłowe w pięciu grupach: dotychczas wykonane mapy i operaty hydrologiczne, mapy topograficzne oraz tematyczne (publikowane i niepublikowane), materiały teledetekcyjne, materiały archiwalne dotyczące wcześniejszych opracowań mapy hydrograficznej w skali 1:50 000 oraz rejestry i bazy danych, szczególnie PIOŚ, IMGW, PIG, IRŚ i Sanepid-u.

Dane geometryczne. Wytyczne nie precyzują źródeł danych geometrycznych; z reguły wymieniane są mapy topograficzne w skalach 1:10 000, 1:25 000 i 1:50 000, mapa glebowo-rolnicza w skali 1:25 000 oraz ogólnie określone zdjęcia lotnicze; jedynie w przypadku wyznaczania działów wodnych wskazuje się jako źródło mapę topograficzną 1:25 000, a przy wyznaczaniu obszarów przepuszczalności gruntów - mapę glebowo-rolniczą 1:25 000; przy określaniu zasięgów obszarów zdrenowanych oraz lejów depresyjnych jako źródło uzupełniające wskazuje się dokumentację odpowiednich urzędów państwowych.

Dane atrybutowe. Wytyczne zakładają pozyskanie znacznej części danych atrybutowych w trakcie hydrograficznego zdjęcia polowego; są to np. bezpośrednie pomiary poziomu wód podziemnych w studniach, wydajności źródeł, rozróżnienie cieków stałych od okresowych, zakwalifikowanie cieków do odpowiedniego przedziału szerokości koryta; rząd działu wodnego należy pozyskać z mapy podziału hydrograficznego Polski (IMGW); charakterystyki zbiorników wodnych należy pozyskać z danych IRS lub z katalogu jezior (Choiński 1991, 1992); niektóre dane (np. rzędną zwierciadła wody zbiorników) należy pobrać z mapy topograficznej.

## XII. Uwagi

Do każdego arkusza mapy jest dołączony komentarz, który zawiera charakterystykę warunków przyrodniczych, budowy geologicznej, topograficznych działów wodnych, opadów atmosferycznych, wód powierzchniowych, warunków hydrologicznych, wód podziemnych, warunków hydroklimatycznych towarzyszących pracom terenowym, przeobrażenia stosunków wodnych i innych informacji. Komentarz ma charakter ogólny i pozwala na uszczegółowienie informacji atrybutowej bazy cyfrowej tylko w ograniczonym zakresie dotyczącym: miejsc zrzutu ścieków oraz punktów pomiaru przepływu wody.



**Ryc. 3.4. Zakres przestrzenny opracowania Mapy Hydrograficznej Polski**  
**A – w układzie współrzędnych 1942, B – w układzie współrzędnych 1992**  
**Źródło: serwisy.codgik.gov.pl, stan na 06.2011**

## MAPA SOZOLOGICZNA POLSKI

### I. Nazwa bazy

Mapa Sozologiczna Polski (MSP)

### II. Merytoryczny cel tworzenia bazy

Opracowanie tematyczne przedstawiające stan środowiska przyrodniczego oraz przyczyny i skutki przemian zachodzących w środowisku pod wpływem różnych procesów, a także sposoby ochrony naturalnych wartości tego środowiska

### III. Utylitarny cel tworzenia bazy

Produkcja kartograficzna

### IV. Właściciel i dysponent

Główny Geodeta Kraju

### V. Źródło opisu standardu bazy

Wytyczne Techniczne K-3.6. System Informacji o Terenie. Mapa Sozologiczna Polski, skala 1:50 000 w formie analogowej i numerycznej. GUGiK 1997

Wytyczne Techniczne GIS-4 Mapa Sozologiczna Polski, skala 1:50000 w formie analogowej i numerycznej. GUGiK 2005

### VI. Odniesienie przestrzenne

Układ współrzędnych PUWG 1992/19 (niektóre opracowania wykonane w układzie 1942; por. ryc. 4.4. A), Układ odniesień wysokościowych Kronsztadt; baza opracowywania w cięciu arkuszowym Międzynarodowej Mapy Świata dla skali

1:50 000, dla niektórych województw dostępna również jako baza ciągła na poziomie województwa lub mniejszych jednostek terytorialnych

#### VII. Skala, zasięg przestrzenny i okres realizacji

1:50 000; okres realizacji od 1997 do 2008; do dziś brak pełnego pokrycia kraju opracowaniem; MSP jest obecnie opracowana dla woj. zachodniopomorskiego, lubuskiego, wielkopolskiego, dolnośląskiego, opolskiego, śląskiego, małopolskiego i podkarpackiego, prawie całego woj. świętokrzyskiego oraz dla części woj. lubelskiego, a także dla pojedynczych arkuszy woj. pomorskiego, kujawsko-pomorskiego i łódzkiego (ryc. 4.4.).

#### VIII. Format źródłowy i dystrybucyjny danych

Źródłowy format: MapInfo TAB (pozwala na przechowywanie obiektów punktowych, liniowych i powierzchniowych w jednej warstwie); spotkane formaty dystrybucyjne: MapInfo TAB, ESRI shape

#### IX. Grupy tematyczne danych

Formy ochrony środowiska przyrodniczego

Degradacja komponentów środowiska przyrodniczego

Przeciwdziałanie degradacji środowiska przyrodniczego

Rekultywacja środowiska przyrodniczego

Nieużytki

Oznaczenia uzupełniające

#### X. Systemy identyfikacji obiektów

W odniesieniu do Wytycznych z 1997 r. – wyłącznie wewnętrzny system identyfikacji obiektów zgodny z pięcioma bazami tabelarycznymi prowadzonymi dla MSP na poziomie krajowym: DBARKUSZ (nazw arkuszy mapy), DBMIEJSC (nazwy miejscowości), DB\_OB\_UC (nazwy zakładów stanowiących uciążliwość dla wód), DBWODY (nazwy wód powierzchniowych) i DBPOMNIK (rodzaje pomników przyrody); ponadto dla pomników przyrody, zrzutów ścieków, emitorów, obszarów szczególnie szkodliwych oraz oczyszczalni ścieków obiekty na tych warstwach posiadają identyfikatory (na poziomie arkusza) zgodne z numeracją zastosowaną w komentarzu do arkusza

W odniesieniu do Wytycznych z 2005 r. – pomniki przyrody posiadają TERYT miejscowości, cieki z nawą i zbiorniki wodne posiadają identyfikator zgodny z tabelą BAZA\_WODY, emitery posiadają REGON właściciela emitora, oczyszczalnie ścieków posiadają REGON użytkownika oraz TERYT miejscowości, zrzuty ścieków

posiadają identyfikator cieku do którego ścieki są zrzucane (ID z tabeli DBWODY) oraz REGON zakładu zrzucającego ścieki; warstwy granic gmin, powiatów i województw oraz miejscowości posiadają TERYT; do tego prowadzone są trzy tabelaryczne bazy uzupełniające: BAZA\_TERYT, BAZA\_REGON i BAZA\_WODY; ta ostatnia jest kontynuacją bazy wg wytycznych z 1997 r., lecz jest sukcesywnie uzupełniana identyfikatorami hydrograficznymi według IMGW; ponadto dla pomników przyrody, zrzutów ścieków, emitorów, obszarów szczególnie szkodliwych oraz oczyszczalni ścieków obiekty na tych warstwach posiadają identyfikatory (na poziomie arkusza) zgodne z numeracją zastosowaną w komentarzu do arkusza

#### XI. Źródła danych geometrycznych i atrybutowych

Wytyczne definiują ogólnie dane źródłowe w ośmiu grupach: mapy topograficzne, mapy tematyczne (publikowane i niepublikowane), materiały teledetekcyjne, studia uwarunkowań i kierunków zagospodarowania przestrzennego gmin, ekofizjografie planistyczne, plany zagospodarowania przestrzennego, inne opracowania dot. ochrony środowiska i gospodarki przestrzennej (publikowane i niepublikowane) oraz rejestry i bazy danych, w szczególności wyniki Państwowego Monitoringu Środowiska.

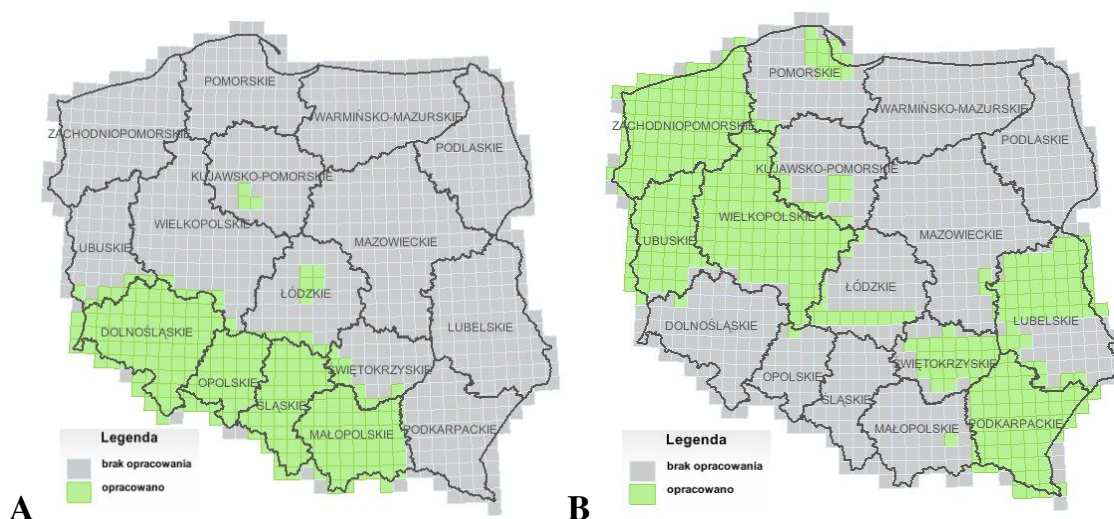
Dane geometryczne. Wytyczne nie precyzują źródeł danych geometrycznych; przy opisie procedury kartowania terenowego zaleca się stosowanie map podkładowych w skali 1:50 000 lub większej, a pierworys mapy wykonuje się w skali 1:50 000. Szereg obiektów należy wprowadzić na etapie terenowego kartowania sozologicznego. Są to w szczególności dzikie źródła skażeń i zanieczyszczeń środowiska przyrodniczego, wyrobisk, zwałowisk, deformacji poeksploatacyjnych terenu, gruntów podatnych na denudację naturogeniczną i uprawową, oczyszczalni ścieków, pasów wiatrochronnych, ekranów akustycznych itp.

Dane atrybutowe. Wytyczne zakładają pozyskanie znacznej części danych atrybutowych poprzez analizę źródeł informacji (wymienionych powyżej), wprowadzenie ich na podkłady mapowe oraz weryfikację terenową. Szereg z tych danych musi być jednak uwiarygodniona na etapie analizy danych źródłowych.

#### XII. Uwagi

Do każdego arkusza mapy jest dołączony komentarz, który zawiera charakterystykę podstawowych komponentów środowiska przyrodniczego, rozszerzające dane dotyczące poszczególnych grup warstw mapy, ogólną ocenę stanu środowiska przyrodniczego i stopnia jego degradacji, wskazania dot. kształtowania i ochrony środowiska oraz inne istotne informacje i oceny. Komentarz ma charakter ogólny i

pozwała na uszczegółowienie informacji atrybutowej bazy cyfrowej tylko w ograniczonym zakresie dotyczącym: pomników przyrody, zrzutów ścieków, emitorów, obszarów szczególnie szkodliwych oraz oczyszczalni ścieków.



**Ryc. 3.5. Zakres przestrzenny opracowania Mapy Sozologicznej Polski  
A – w układzie współrzędnych 1942, B – w układzie współrzędnych 1992  
Źródło: serwisy.codgik.gov.pl, stan na 06.2011**

## **SZCZEGÓŁOWA MAPA GEOLOGICZNA POLSKI**

### I. Nazwa bazy

Szczegółowa Mapa Geologiczna Polski (SMGP)

### II. Merytoryczny cel tworzenia bazy

Odzwierciedlenie budowy geologicznej powierzchni terenu przy równoczesnym rozpoznaniu budowy geologicznej w profilu pionowym.

### III. Utylitarny cel tworzenia bazy

Produkcja kartograficzna

### IV. Właściciel i dysponent

Ministerstwo Środowiska, Państwowy Instytut Geologiczny

### V. Źródło opisu standardu bazy

Instrukcja opracowania i wydania Szczegółowej Mapy Geologicznej Polski w skali 1:50000. Państwowy Instytut Geologiczny, Warszawa 1996

### VI. Odniesienie przestrzenne

Układ współrzędnych „1942”; baza opracowywania w cięciu arkuszowym Międzynarodowej Mapy Świata dla skali 1:50 000; w miejsce godeł przyjęto własną numerację arkuszy mapy

## VII. Skala, zasięg przestrzenny i okres realizacji

1:50 000; mapa w komputerowej postaci jest realizowana od 1994 roku; opracowano już większość arkuszy (ryc. 4.5.)

## VIII. Format źródłowy i dystrybucyjny danych

Baza jest realizowana w oprogramowaniu ArcINFO; najpopularniejszym formatem dystrybucyjnym jest ESRI shp

## IX. Grupy tematyczne danych

Mapa zawiera warstwę wydzieleni genetyczno-litologiczno-stratygraficznych, elementy dokumentacyjne oraz linie przekrojów geologicznych

## X. Systemy identyfikacji obiektów

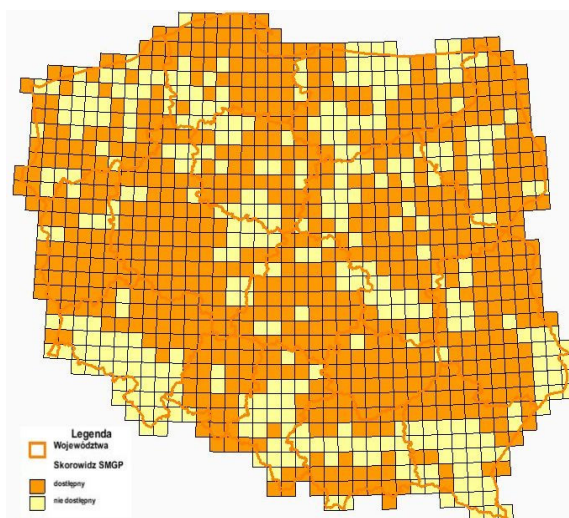
Baza zawiera wewnętrzny system identyfikatorów słownikowych. Brak identyfikatorów o charakterze przestrzennym

## XI. Źródła danych geometrycznych i atrybutowych

Przy sporządzaniu prac geologicznych nad mapą autor jest zobowiązany do wykorzystania wszystkich geologicznych i geofizycznych materiałów dokumentacyjnych oraz baz danych gromadzących dane archiwalne z różnych dziedzin, szczególnie bazy danych Państwowego Instytutu Geologicznego: MIDAS, HYDRO, SMGP oraz Bank Danych Elektrooporowych. Głównym źródłem danych są wyniki szczegółowych prac terenowych wykonawcy arkusza.

## XII. Uwagi

Dla każdego arkusza mapy są dołączone objaśnienia, które zawierają informacje na temat ukształtowania powierzchni terenu i budowy geologicznej, w tym: stratygrafii, tektoniki i rzeźby podłoża czwartorzędu oraz rozwoju budowy geologicznej. Uzupełnieniem tekstu jest szkic geomorfologiczny, szkic geologiczny odkryty, tabela litologiczno-stratygraficzna oraz inne tabele i zestawienia.



*Ryc. 3.6. Dostępność bazy Szczegółowej Mapy Geologicznej Polski  
Źródło: baza.pgi.gov.pl, stan na 06.2011*

## **MAPA HYDROGEOLOGICZNA POLSKI**

### I. Nazwa bazy

Mapa Hydrogeologiczna Polski (MHP), główny poziom wodonośny

### II. Merytoryczny cel tworzenia bazy

Kartograficzne odwzorowanie warunków hydrogeologicznych, w tym wskazanie głównego piętra / poziomu wodonośnego i jego charakterystyki jakościowej, ilościowej oraz zagrożeń zasobów wód podziemnych

### III. Utylitarny cel tworzenia bazy

Produkcja kartograficzna

### IV. Właściciel i dysponent

Ministerstwo Środowiska, Państwowy Instytut Geologiczny

### V. Źródło opisu standardu bazy

Instrukcja opracowania i komputerowej edycji Mapy Hydrogeologicznej Polski w skali 1:50 000. Część I Opracowanie autorskie Część II Opracowanie komputerowe. Państwowy Instytut Geologiczny, Warszawa 1999

### VI. Odniesienie przestrzenne

Układ współrzędnych „1942”; baza opracowywania w cięciu arkuszowym Międzynarodowej Mapy Świata dla skali 1:50 000; w miejsce godeł przyjęto własną numerację arkuszy mapy

## VII. Skala, zasięg przestrzenny i okres realizacji

1:50 000; mapa w komputerowej postaci jest realizowana od 1994 roku; opracowano już niemal wszystkie arkusze (ryc. 4.6.)

## VIII. Format źródłowy i dystrybucyjny danych

Format źródłowy związany jest ze środowiskiem MGE Intergraph; najpopularniejszy format dystrybucyjny – ESRI shp

## IX. Grupy tematyczne danych

Rozmieszczenie zwykłych wód podziemnych

Charakterystyka ilościowa wód podziemnych

Hydrodynamika

Jakość wód podziemnych

Stopień zagrożenia wód podziemnych

## X. Systemy identyfikacji obiektów

Mapa posiada wyłącznie wewnętrzny system identyfikacji obiektów w ramach sekcji mapy, pozwalający na przyłączanie danych tabelarycznych zawartych w objaśnieniach.

## XI. Źródła danych geometrycznych i atrybutowych

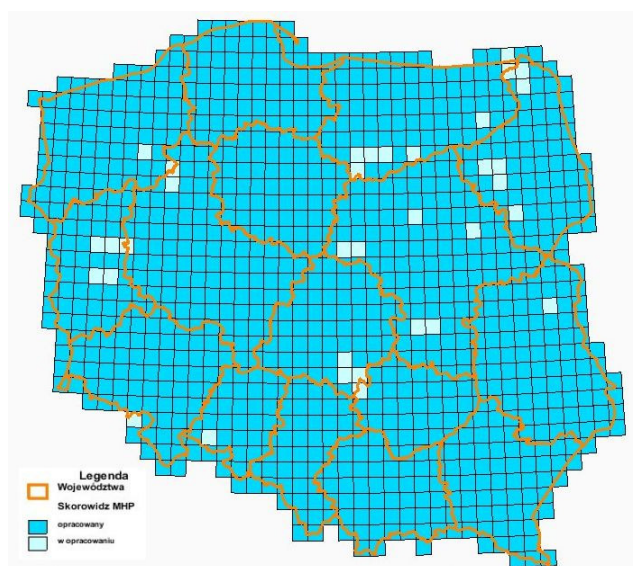
Mapa Hydrogeologiczna Polski jest opracowaniem autorskim, opracowanym z wykorzystaniem dużej ilości danych źródłowych (reinterpretacja). Materiały źródłowe to m. in.: atlas zasobów zwykłych wód podziemnych i ich wykorzystanie w Polsce (1:500 000), atlas (1:500 000) i mapa (1:200 000 i 1:50 000) hydrogeologiczna Polski, mapa obszarów głównych zbiorników wód podziemnych (GZWP; 1:500 000), mapy geologiczne Polski (1:200 000 i 1:50 000), Szczegółowa mapa geologiczna Sudetów (1:25 000), mapy glebowo-rolnicze (1:25 000 i 1:50 000), podział hydrograficzny Polski (1:200 000) i inne; ponadto w opracowaniu wykorzystywane są dane z otworów wiertniczych, ujęć wodnych, studni kopanych, dokumentacje hydrogeologiczne, aktualne dane z istniejących i potencjalnych źródeł zanieczyszczeń i szereg innych.

## XII. Uwagi

Dla każdego arkusza mapy są dołączone objaśnienia, które zawierają następujące działy: warunki geologiczne, klimat i wody powierzchniowe, warunki hydrogeologiczne (w tym: użytkowe piętra / poziomy wód zwykłych oraz regionalizacja hydrogeologiczna), jakość wód podziemnych oraz zagrożenia i ochrona wód. Integralną częścią tekstu są przekroje hydrogeologiczne, mapy: głębokości

występowania głównego piętra / poziomu wodonośnego, miąższości i przewodności głównego piętra / poziomu wodonośnego oraz mapa dokumentacyjna, a ponadto znaczna liczba tabel z opisami otworów, studni, źródeł, wynikami analiz chemicznych i in.

Omówiona mapa hydrogeologiczna obejmuje główny poziom wodonośny, stanowiący zasadnicze źródło zaopatrzenia w wodę. W roku 2004 powstała instrukcja reambulacji tej mapy oraz utworzenia drugiej edycji Mapy Hydrogeologicznej Polski, obejmującej pierwszy poziom wodonośny (Mapa Hydrogeologiczna Polski w skali 1:50 000. Udostępnianie, weryfikacja, aktualizacja i rozwój. Państwowy Instytut Geologiczny, Warszawa 2004).



*Ryc. 3.7. Stan opracowania Mapy Hydrogeologicznej Polski  
Źródło: baza.pgi.gov.pl, stan na 06.2011*

## MAPA GEOŚRODOWISKOWA POLSKI

### I. Nazwa bazy

Mapa Geośrodowiskowa Polski (MGP)

### II. Merytoryczny cel tworzenia bazy

Mapa prezentuje obszary występowania kopalin, gospodarkę złożami, wybrane elementy górnictwa i przetwórstwa kopalin, hydrogeologii, geologii inżynierskiej, ochrony przyrody, krajobrazu i zabytków kultury, stan geochemiczny powierzchni ziemi oraz możliwość składowania odpadów

### III. Utylitarny cel tworzenia bazy

Produkcja kartograficzna

#### IV. Właściciel i dysponent

Ministerstwo Środowiska, Państwowy Instytut Geologiczny

#### V. Źródło opisu standardu bazy

Instrukcja opracowania i aktualizacji Mapy geologiczno-gospodarczej Polski w skali 1:50 000. Państwowy Instytut Geologiczny, Warszawa 2002.

Instrukcja opracowania Mapy geośrodowiskowej Polski w skali 1:50000. Państwowy Instytut Geologiczny, Warszawa 2005.

#### VI. Odniesienie przestrzenne

Układ współrzędnych „1942”; baza opracowywania w cięciu arkuszowym Międzynarodowej Mapy Świata dla skali 1:50 000; w miejsce godeł przyjęto własną numerację arkuszy mapy

#### VII. Skala, zasięg przestrzenny i okres realizacji

1:50 000; mapa jest realizowana od 2002 roku jako rozszerzona wersja Mapy Geologiczno-Gospodarczej Polski; podstawowa wersja Mapy Geologiczno-Gospodarczej Polski jest opracowana dla prawie całego kraju, z czego część jest zaktualizowana do Mapy Geośrodowiskowej Polski (ryc. 4.7.)

#### VIII. Format źródłowy i dystrybucyjny danych

Format źródłowy związany jest ze środowiskiem MGE Intergraph; najpopularniejszy format dystrybucyjny – ESRI shp

#### IX. Grupy tematyczne danych

Złóża kopalin

Wody

Zagrożenia powierzchni ziemi

Warunki podłoża

Ochrona środowiska

#### X. Systemy identyfikacji obiektów

W bazie brak identyfikatorów pozwalających przyłączać dane zewnętrzne

#### XI. Źródła danych geometrycznych i atrybutowych

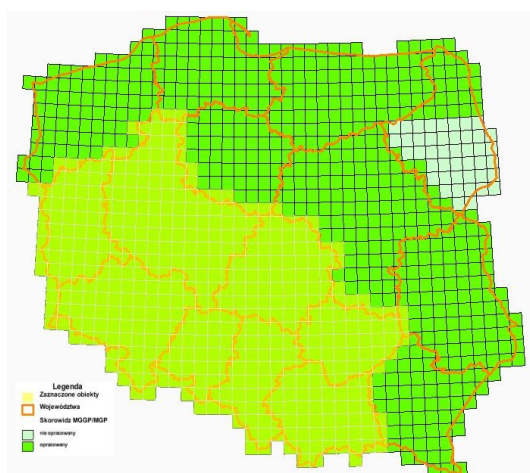
Mapa Geośrodowiskowa Polski jest autorskim opracowaniem z wykorzystaniem znacznej ilości danych źródłowych (reinterpretacja). Wśród źródeł należy wymienić m. in.: mapę topograficzną 1:50 000 układ 1992 oraz skale większe, Szczegółową Mapę Geologiczną Polski (1:50 000), Szczegółową Mapę Geologiczną Sudetów (1:25 000), Mapę Geologiczną Polski (1:200 000), Mapę Hydrogeologiczną Polski (1:50 000 i 1:200 000), Mapę Geologiczno-Gospodarczą Polski (1:50 000), Mapę

Głównych Zbiorników Wód Podziemnych oraz inne; ponadto materiały z Regionalnych Zarządów Gospodarki Wodnej, Wojewódzkich Inspektoratów Ochrony Środowiska, Instytutu Meteorologii i Gospodarki Wodnej, Bazy danych geologicznych PIG, dokumenty planistyczne i inne.

## XII. Uwagi

Cała baza Mapy Geośrodowiskowej Polski składa się z trzech odrębnych baz danych: „Zasoby przyrody” (baza MGGP), „Składowanie odpadów” (baza SO) i „Geochemia środowiska” (baza GS). Pierwsza z wymienionych baz pozwala na wygenerowanie obrazu kartograficznego planszy A mapy, natomiast dwie pozostałe składają się na planszę B.

Dla każdego arkusza mapy są dołączone objaśnienia, które zawierają następujące działy: charakterystykę geograficzną i gospodarczą terenu, budowę geologiczną, złoża kopalin, górnictwo i przetwórstwo kopalin, perspektywy i prognozy występowania kopalin, warunki wodne, geochemię środowiska, składowanie odpadów, warunki podłoża budowlanego oraz ochronę przyrody i krajobrazu. Objasnienia wzbogacone są w schematyczne mapy geologiczne, geomorfologiczne, hydrogeologiczne, a także tabele i zestawienia złóż kopalin, czynnych kopalń, geochemii, warunków podłoża budowlanego itp.



**Ryc. 3.8. Stan opracowania Mapy Geologiczno-Gospodarczej Polski (kolor zielony z kratkowaniem) i Mapy Geośrodowiskowej Polski (kolor jasnozielony)  
Źródło: baza.pgi.gov.pl, stan na 06.2011**

## **MAPA PODZIAŁU HYDROGRAFICZNEGO POLSKI**

### I. Nazwa bazy

Mapa Podziału Hydrograficznego Polski (MPHP)

### II. Merytoryczny cel tworzenia bazy

Stworzenie pełnej charakterystyki geometrycznej i opisowej sieci wodnej oraz zlewni w układzie zlewniowym.

### III. Utylitarny cel tworzenia bazy

Opracowanie referencyjne dla działań związanych z gospodarką wodną i jej planowaniem zgodnie z Prawem Wodnym oraz Ramową Dyrektywą Wodną Unii Europejskiej.

Produkcja kartograficzna

### IV. Właściciel i dysponent

Ministerstwo Środowiska, Krajowy Zarząd Gospodarki Wodnej

### V. Źródło opisu standardu bazy

Powszechnie niedostępne (opracowanie Instytutu Meteorologii i Gospodarki Wodnej)

### VI. Odniesienie przestrzenne

Układ współrzędnych 1992; mapa jest opracowana jako baza ciągła dla terytorium całego kraju oraz dla części dorzecza Wisły i Odry, leżącymi poza granicami kraju. W postaci rastrowej mapa jest dostępna także w cięciu arkuszowym Międzynarodowej Mapy Świata

### VII. Skala, zasięg przestrzenny i okres realizacji

1:50 000, obszar całego kraju; mapę zrealizowano w roku 2007

### VIII. Format źródłowy i dystrybucyjny danych

ESRI shape; mapa jest dostępna również w postaci rastrowej (jpg + jgw)

### IX. Grupy tematyczne danych

Obszarowe jednostki hydrograficzne

Cieki wyróżnione

Zbiorniki wyróżnione

Pozostałe cieki

Podział sekcyjny mapy topograficznej 1:50 000

### X. Systemy identyfikacji obiektów

Zastosowano spójny system identyfikacji hydrograficznej obiektów (cieków, zbiorników wodnych i zlewni)

## XI. Źródła danych geometrycznych i atrybutowych

Dane geometryczne – mapy topograficzne 1:50 000 w układzie 1942

## XII. Uwagi

Baza stanowi referencję do wszelkich działań z zakresu gospodarki wodnej i ochrony środowiska w odniesieniu do wód powierzchniowych

## **MAPA GLEBOWO-ROLNICZA**

### I. Nazwa bazy

Mapa Glebowo-Rolnicza (MGR)

### II. Merytoryczny cel tworzenia bazy

Charakterystyka przestrzenna gleb gruntów rolnych pod względem kompleksów rolniczej przydatności, typów i podtypów genetycznych oraz gatunków gleb z punktu widzenia składu mechanicznego

### III. Utylitarny cel tworzenia bazy

Produkcja kartograficzna

### IV. Właściciel i dysponent

Instytut Upraw i Nawożenia w Puławach i / lub (niezależnie) Urzędy Marszałkowskie (Wojewódzkie Ośrodki Dokumentacji Geodezyjnej i Kartograficznej)

### V. Źródło opisu standardu bazy

Powszechnie niedostępne (opracowanie Instytut Upraw i Nawożenia w Puławach lub ustalenia własne innych inicjatorów utworzenia bazy); standard bazy opiera się na założeniach mapy glebowo-rolniczej w postaci analogowej

### VI. Odniesienie przestrzenne

Zróznicowane, zależne od inicjatora utworzenia bazy; najczęściej spotykany jest układ współrzędnych 1992, choć dla map 1:5 000 zdarza się układ 2000

### VII. Skala, zasięg przestrzenny i okres realizacji

1:25 000 i/lub 1:5 000, zasięg przestrzenny i okres realizacji trudne do określenia (brak precyzyjnej informacji)

### VIII. Format źródłowy i dystrybucyjny danych

Format źródłowy zróznicowany, zależny od inicjatora utworzenia bazy; najczęściej dane są udostępniane w formacie ESRI shape

### IX. Grupy tematyczne danych

Wydzielenia jednostek glebowo-rolniczych

#### X. Systemy identyfikacji obiektów

Nie stosuje się systemu identyfikacji przestrzennej

#### XI. Źródła danych geometrycznych i atrybutowych

Źródłem danych są najczęściej mapy glebowo-rolnicze w skali 1:25 000 i 1:5 000, które mogą być uzupełniane informacjami z operatów glebowo-rolniczych

#### XII. Uwagi

Brak szczegółowych unormowań w zakresie cyfrowania map glebowo-rolniczych powoduje regionalne zróżnicowanie w produktach cyfrowych Mapy Glebowo-Rolniczej; dotyczy ono głównie konstruowania struktury zapisu informacji atrybutowej.

### **EWIDENCJA GRUNTÓW I BUDYNKÓW**

#### I. Nazwa bazy

Ewidencja Gruntów i Budynków (EGiB)

#### II. Merytoryczny cel tworzenia bazy

Prowadzenie w sposób jednolity dla kraju, systematycznie aktualizowanego zbioru informacji o gruntach, budynkach i lokalach, ich właścicielach oraz o innych osobach fizycznych lub prawnych władających tymi gruntami, budynkami i lokalami.

#### III. Utylitarny cel tworzenia bazy

Stworzenie katastru gruntowego w Polsce.

Produkcja kartograficzna

#### IV. Właściciel i dysponent

Powiatowe Ośrodki Dokumentacji Geodezyjnej i Kartograficznej (PODGiK)

#### V. Źródło opisu standardu bazy

Rozporządzenie Ministra Rozwoju Regionalnego i Budownictwa z dnia 29 marca 2001 r. w sprawie ewidencji gruntów i budynków (Dz. U. Nr 38, poz. 454)

Instrukcja Techniczna G-5 Ewidencja gruntów i budynków z 3 listopada 2003

#### VI. Odniesienie przestrzenne

Zależne od regionu kraju (układy państwowe i lokalne), docelowo układ współrzędnych 2000

#### VII. Skala, zasięg przestrzenny i okres realizacji

Zwyczajowo 1:5 000, ale dokładność danych geometrycznych jest geodezyjna (zgodnie z Instrukcją O-1 współrzędne szczegółów sytuacyjnych I grupy dokładnościowej, np. znaków granicznych powinny być wyznaczone z dokładnością  $\pm$

0,10 m); zasięg przestrzenny i okres realizacji EGiB są trudne do określenia ze względu na decentralizację dysponentów (PODGiK-i)

#### VIII. Format źródłowy i dystrybucyjny danych

Format źródłowy jest zależny od wybranego przez Starostwo Powiatowe rozwiązania technologicznego; nie każde z tych rozwiązań pozwala na bezproblemową konwersję danych do formatu zgodnego ze standardem GIS. Przyjęty standard wymiany danych ewidencyjnych (SWDE) nie spełnił do końca swej roli. Dlatego na etapie dystrybucji danych do tych zastosowań konieczna jest często bezpośrednia współpraca między służbami informatycznymi PODGiK a odbiorcą danych

#### IX. Grupy tematyczne danych

Główne grupy tematyczne istotne dla celów badawczych i użytkarnych to:

Działki ewidencyjne (GESDZI)

Budynki (GESBZO)

Klasoużytki (GESKLU)

Pozostałe grupy tematyczne obejmują inne elementy ewidencji, np. punkty graniczne nieruchomości, zasięgi obrębów geodezyjnych i inne, a ponadto sieci osnów geodezyjnych i elementy uzbrojenia terenu. Ta ostatnia grupa warstw nie jest kompletna dla znacznych obszarów.

#### X. Systemy identyfikacji obiektów

Istnieje pełna identyfikacja przestrzenna obiektów za pomocą rozbudowanego numeru TERYT. W przypadku trzech głównych grup tematycznych (por. pkt 9) każdy obiekt geometryczny posiada identyfikator przestrzenny w postaci WWPPGG\_R.XXXX., gdzie człon przed kropką oznacza TERYT gminy, a po kropce – umowny numer obrębu geodezyjnego. Po tym numerze następuje numer działki ewidencyjnej (warstwa GESDZI), numer działki i nr porządkowy budynku (warstwa GESBZO) lub wyróżnik literowy użytku gruntowego i numer porządkowy konturu użytku (warstwa GESKLU). Przedstawiony powyżej system identyfikacyjny jest modyfikowany w sytuacjach szczególnych, np. jeśli numeracja działek ewidencyjnych jest niepowtarzania w granicach arkusza, a nie w granicach obrębu geodezyjnego.

#### XI. Źródła danych geometrycznych i atrybutowych

Dane geometryczne (część graficzna EGiB): pomiary terenowe, archiwalne operaty pomiarowe, szkice geodezyjne, czasami digitalizacja map analogowych. Szczegółową procedurę pozyskania tych danych opisuje Instrukcja G-5

Dane atrybutowe (część opisowa EGiB) gromadzone są w postaci rejestrów i kartotek: rejestr gruntów, rejestr budynków, rejestr lokali, kartoteka budynków, kartoteka lokali. Szczegółową procedurę pozyskania tych danych opisuje Instrukcja G-5

## XII. Uwagi

System EGiB stanowi najdokładniejsze źródło danych przestrzennych dla obszarów nieleśnych spośród prezentowanych w tym opracowaniu. Wszelkie dane geometryczne i opisowe są pozyskiwane zgodnie z ustalonymi procedurami (Instrukcja G-5) zapewniającymi zgodność ze stanem faktycznym i prawnym

## **LEŚNA MAPA NUMERYCZNA**

### I. Nazwa bazy

Leśna Mapa Numeryczna (LMN)

### II. Merytoryczny cel tworzenia bazy

Opisowo-graficzna baza danych, obejmującą całą działalność nadleśnictwa, w tym działalność gospodarczą w pododdziałach i wydzieleniach

### III. Utylitarny cel tworzenia bazy

Zarządzanie zasobami leśnymi.

Produkcja kartograficzna

### IV. Właściciel i dysponent

Dyrekcja Generalna Lasów Państwowych, Regionalne Dyrekcje Lasów Państwowych, Nadleśnictwa

### V. Źródło opisu standardu bazy

Leśna Mapa Numeryczna (LMN) – Standard Leśnej Mapy Numerycznej (SLMN) obowiązujący na 31 lipca 2005 roku ([www.lp.gov.pl](http://www.lp.gov.pl))

System Informatyczny Lasów Państwowych (SILP) – opis powszechnie niedostępny

### VI. Odniesienie przestrzenne

Układ współrzędnych 1992, baza jest realizowana dla nadleśnictw

### VII. Skala, zasięg przestrzenny i okres realizacji

1:5 000, baza jest prawdopodobnie opracowana dla całego kraju, okres realizacji – początek jednolitego standardu to rok 2001

### VIII. Format źródłowy i dystrybucyjny danych

Baza jest opracowywana w programie ArcINFO (LMN) oraz w formacie bazodanowym MS Access (SILP); dane są udostępniane jako warstwy ESRI shape oraz jako tabele .mdb

## IX. Grupy tematyczne danych

Leśna Mapa Numeryczna zawiera znaczną liczbę warstw tematycznych; istotna z punktu widzenia użytkownika oczekującego charakterystyki lasu jest warstwa pododdziałów leśnych (warstwa „wydz\_pol”) połączona z odpowiednimi tabelami atrybutowymi bazy SILP; najistotniejsza informacja uzupełniająca znajduje się na warstwach „syt\_pol” (powierzchniowe elementy sytuacyjne), „kom\_lin” (linie komunikacyjne) oraz „ciek\_lin” (obiekty liniowe rzek, kanałów i rowów). Pozostałe elementy bazy mają zdaniem autora niewielkie znaczenie dla zwykłego użytkownika tych danych.

## X. Systemy identyfikacji obiektów

Baza zawiera spójny system adresowy pododdziałów leśnych, co pozwala na lokalizację obiektów w przestrzeni po identyfikatorze oraz zapewnia spójność bazy wektorowej (LMN) i atrybutowej (SILP) w obrębie całego kraju

## XI. Źródła danych geometrycznych i atrybutowych

Dane geometryczne pochodzą z operatów pomiarowych, szkiców i map leśnych, natomiast dane atrybutowe są wynikiem kompleksowego opisu taksacyjnego lasu.

## XII. Uwagi

System LMN wraz z bazą SILP stanowią najdokładniejsze źródło danych przestrzennych dla obszarów leśnych spośród prezentowanych w tym opracowaniu.

## **NUMERYCZNY MODEL TERENU Z MAPY WEKTOROWEJ POZIOMU DRUGIEGO**

### I. Nazwa bazy

Numeryczny model terenu z mapy wektorowej poziomu drugiego (DTED2)

### II. Merytoryczny cel tworzenia bazy

Zasilanie systemów geoinformacyjnych wymagających danych wysokościowych

### III. Utylitarny cel tworzenia bazy

Tworzenie wizualizacji i analiz rzeźby terenu na średnim poziomie skalowym

### IV. Właściciel i dysponent

Zarząd Geografii Wojskowej Wojska Polskiego

### V. Źródło opisu standardu bazy

DIGEST – *Digital Geographic Information Exchange Standard*

#### VI. Odniesienie przestrzenne

Układ współrzędnych WGS84, Układ odniesień wysokościowych Kronsztadt; baza opracowywania w cięciu arkuszowym Międzynarodowej Mapy Świata dla skali 1:50 000 (po dwa sąsiadujące arkusze)

#### VII. Skala, zasięg przestrzenny i okres realizacji

1:50 000, dane obejmują cały obszar kraju, zostały opracowane w latach 1997-2003

#### VIII. Format źródłowy i dystrybucyjny danych

Dostępny jest format źródłowy .dt2, ASCII, ERDAS i ESRI

#### IX. Grupy tematyczne danych

Siatka numerycznego modelu rzeźby terenu

#### X. Systemy identyfikacji obiektów

Brak

#### XI. Źródła danych geometrycznych i atrybutowych

Podstawą opracowania były na podstawie poziomic z diapozytywów wojskowych map topograficznych 1:50 000. Dane źródłowe przetworzono do regularnej siatki GRID o rozdzielczości terenowej około 30m (odpowiadającej łukowi 1" długości geograficznej).

#### XII. Uwagi

Jakkolwiek DTED2 jest znakomitym modelem rzeźby, to powstała równoległa do niego i na podstawie tego samego materiału źródłowego wektorowa warstwa poziomic (VML2) pozwala na zbudowanie dokładniejszego modelu rzeźby terenu.

### **MAPA POKRYCIA TERENU CORINE LAND COVER**

#### I. Nazwa bazy

Mapa Pokrycia Terenu Corine Land Cover (CLC)

#### II. Merytoryczny cel tworzenia bazy

Dokumentacja stanu i zmian w pokryciu terenu na obszarze Polski,

#### III. Utylitarny cel tworzenia bazy

Określanie kierunków i zakresu zmian w pokryciu terenu, aktualizacja porównywalnych danych dla Europy

#### IV. Właściciel i dysponent

Główny Inspektorat Ochrony Środowiska

#### V. Źródło opisu standardu bazy

Standard CORINE Land Cover

#### VI. Odniesienie przestrzenne

Układ współrzędnych 1992, baza jest realizowana dla całego kraju

#### VII. Skala, zasięg przestrzenny i okres realizacji

Nominalna skala opracowania 1:100 000, baza obejmuje teren całego kraju. CLC była realizowana w latach 90. XX wieku (CLC1990), na pocz. XXI wieku (CLC2000) oraz w 2006 roku (CLC2006)

#### VIII. Format źródłowy i dystrybucyjny danych

Dane przygotowano w środowisku oprogramowania ESRI, a dostępne są w formacie ESRI shape lub .e00

#### IX. Grupy tematyczne danych

Warstwa pokrycia terenu

Warstwa zmian w pokryciu terenu

#### X. Systemy identyfikacji obiektów

Obiekty bazy nie posiadają identyfikacji przestrzennej

#### XI. Źródła danych geometrycznych i atrybutowych

W przypadku CLC1990 i CLC2000 źródłem informacji były zdjęcia satelitarne wykonane za pomocą skanera TM przez satelitę Landsat 7. Dla CLC2006 wybrano zdjęcia satelitarne wykonane przez indyjskiego satelitę IRS oraz francuskiego SPOT-4.

#### XII. Uwagi

Brak

### **3.3. Analiza porównawcza danych o środowisku przyrodniczym w bazach**

Czternaście krajowych baz danych przestrzennych, scharakteryzowanych w ujednolicony sposób w rozdziale 3.2. cechuje zróżnicowany zakres tematyczny i dokładnościowy. Rozpatrywanie aplikacyjności tego zasobu do zastosowań naukowych i użytecznych wymaga w pierwszej kolejności stworzenia tematycznego zestawienia grup warstw dostępnych w przedmiotowych bazach. Zabieg ten pozwoli wytypować dane szczególnie cenne z punktu widzenia celu badawczego pracy. Kryterium selekcji danych przy tworzeniu tego zestawienia jest spełnianie przez nie roli opisowej składowych poszczególnych komponentów środowiska przyrodniczego. Takie odniesienie ma na celu wskazanie składowych źródłowych baz danych przestrzennych, które mogą tworzyć charakterystykę poszczególnych komponentów środowiska przyrodniczego (tab. 3.1.).

**Tabela 3.1. Podział środowiska przyrodniczego na geokomponenty**  
**Źródło: opracowanie własne**

<b>GEOKOMPONENT</b>	<b>PRZYKŁADOWE CHARAKTERYSTYKI KOMPONENTÓW ŚRODOWISKA PRZYRODNICZEGO</b>
<b>Atmosfera (At)</b>	- uwarunkowania topoklimatu - zagrożenia atmosfery zanieczyszczeniami - mierzone parametry meteorologiczne
<b>Antroposfera (An)</b>	- zasięg obszarów zabudowy - linie komunikacyjne - obiekty przemysłowe zagrażające środowisku - eksploatacja surowców naturalnych
<b>Biosfera (Bi)</b>	- zasięgi formacji roślinnych - geobotaniczne warunki terenu - skład gatunkowy lasu - formy ochrony przyrody żywej
<b>Hydrosfera (Hy)</b>	- sieć hydrograficzna - zlewnie rzeczne - zagrożenia wód powierzchniowych - obszary narażone na wezbrania powodziowe
<b>Morfogeosfera (Mg)</b>	- typy morfometryczne terenu - typy genetyczne form terenu - nachylenie i orientacja stoków - wskaźniki energii rzeźby terenu
<b>Pedosfera (Pd)</b>	- typy genetyczne gleb - kompleksy rolniczej przydatności gleb - klasy bonitacyjne gruntu - typy użytków gruntowych
<b>Litosfera (Li)</b>	- budowa głębszego podłoża - litologia przypowierzchniowa - przepuszczalność gruntu - skała macierzysta gleb

Odnosząc się do podziału środowiska przyrodniczego na geokomponenty (Richling 2007) opracowano zestawienie trzydziestu przykładowych składników komponentów środowiska przyrodniczego (tab. 3.2.). Wyboru składników komponentowych dokonano w oparciu o umieszczony w rozdziale 3.1. opis baz danych przestrzennych, a w szczególności punkt 9 opisu („Grupy tematyczne danych”). Celem zestawienia było:

1. Wskazanie potencjalnych źródeł danych do charakterystyki poszczególnych składników komponentów środowiska przyrodniczego.
2. Określenie stopnia redundancji danych

Zestawienie tabelaryczne wskazuje obecność danej grupy warstw w bazach. Znak „+/-” oznacza niepełną wartość danych, przejawiającą się najczęściej w ograniczonym względem oczekiwanego zakresie tematycznym danych.

Jak widać z zestawienia (tab. 3.2.) bardziej uniwersalne w zapewnieniu danych źródłowych są bazy danych o charakterze topograficznym lub tematyczne o szerokim spectrum tematyki (np. mapa sozologiczna). Mówią o tym liczebności grup warstw występujących w ramach każdej z baz danych (wiersz w dolnej części tabeli). Wyjątek stanowi mapa hydrograficzna, która jest bazą o wąskim temacie, ale zawiera szereg warstw dość dobrze odzwierciedlających również inne komponenty środowiska. Bazy o wąskim zakresie tematycznym zapewniają z reguły dane do jednego, dwóch tematów. Wśród nich można wymienić: Szczegółową Mapę Geologiczną Polski, Mapę Glebowo-Rolniczą, Mapę Podziału Hydrograficznego Polski czy numeryczny model rzeźby DTED2. Oczywiście liczebność istotnych grup warstw nie stanowi jeszcze o atrakcyjności tej bazy bez rozpoznania jakości danych.

Z drugiej strony tab. 3.2. prezentuje stopień redundancji danych wyrażony liczebnością baz posiadających poszczególne grupy warstw tematycznych (ostatnia kolumna tabeli). Znaczny poziom redundancji dotyczy elementów środowiska zaliczanych zwyczajowo w kartografii do obiektów będących treścią map topograficznych. Bardziej szczegółowa informacja tematyczna (np. emitery, typy gleb) jest reprezentowana przez małą ilość baz.

Podsumowując można ocenić, że:

- większość z rozpatrywanych baz danych reprezentuje krajobrazowy model danych (DLM); wynika to z: dokładności danych geometrycznych (EGiB, LMN), zastosowanego źródła danych (BDOT) lub przyjętą metodę klasyfikacji rozpatrywanego zjawiska w przestrzeni geograficznej (MGR, SMGP, MPHP, CLC); bazy danych związane w założeniach z produkcją kartograficzną (np. MSP, MHP, MGP) lub powstałe w oparciu o materiał kartograficzny (VML2, BDO) zawierają elementy modelu kartograficznego (DCM);
- większość baz danych posiada dokładny opis standardu w postaci instrukcji i wytycznych technicznych (ryc. 3.9.). Jedynie bazie MPHP nie towarzyszy ogólnie dostępna instrukcja realizacyjna;
- cztery spośród rozpatrywanych baz (EGiB, LMN, BDOT i częściowo MGR) są opracowaniami nominalnie wielkoskalowymi (ryc. 3.9.), przez co znajdują praktyczne zastosowania w szczegółowych analizach przestrzennych na poziomie lokalnym. Pozostałe bazy znajdują zastosowanie z reguły w analizach regionalnych;

Tabela 3.2. Przykładowe 30 grup warstw charakteryzujących wybrane elementy geokomponentów: atmosfery (At), antroposfery (An), biosfery (Bi), hydrosfery (Hy), morfogeosfery (Mg), pedosfery (Pd) i litosfery (Li). Plus oznacza dostępność danych w bazie, „+/-”, oznacza dostępność niepełnych danych. Ostatnia kolumna określa stopień redundancji danych tematycznych (liczbę baz danych, w których występuje dana grupa warstw, a w nawiasie z uwzględnieniem warstw z niepełnymi danymi).

Źródło: opracowanie własne

Geokomponent	Wybrane składniki komponentów środowiska przyrodniczego	BDO	CLC	VML2	DTED2	MHP	MSP	SMGP	MHgP	MGP	MPHP	BDOT	MGR	EGiB	LMN	Redundancja danych
Nominalna skala opracowania		1:250 000	1:100 000	1:50 000								1:10 000	1:25 000 1:5 000	1:5 000		
At	(An) Emitory zanieczyszczeń, hałasu, odorów						+									1
	(An) Urządzenia odpylające, pasy wiatrochronne, ekrany akust.						+									1
An	Formy użytkowania terenu*)	+	+	+/-			+/-					+		+	+	5 (7)
	Linie komunikacyjne	+		+		+/-	+/-			+/-		+		+	+	5 (8)
	Obiekty przemysłowe	+	+/-	+			+			+		+		+	+	7 (8)
	Obszary zabudowy mieszkaniowej	+	+	+		+	+			+		+		+	+	9
	Granice jednostek administracyjnych	+		+		+	+					+		+	+	7
Bi	Zasięg lasów i roślinności nieleśnej	+	+	+		+	+			+		+		+	+	9
	Siedliska leśne														+	1
	(An) Stan sanitarny lasu						+								+	2
	(An) Formy ochrony przyrody	+/-		+/-			+/-			+/-		+/-				0 (5)
Hy	Wody powierzchniowe	+	+/-	+		+	+		+	+	+	+		+	+	10 (11)
	Działy wodne, zlewnie					+			+		+					3
	Bramy wodne					+										1
	Wpływy wód podziemnych			+/-		+				+		+				3 (4)
	(An) Oczyszczalnie ścieków			+		+/-	+		+			+				4 (5)
	(An) Zbiorniki wód przemysłowych			+		+	+			+		+				5
	(An) Klasy czystości wód powierzchniowych					+	+			+						3
	Głębokość zalegania wód podziemnych					+			+							2
	(An) GZWP						+			+						2
	(An) Ujęcia wód					+	+		+	+		+				5
	(An) Leje depresyjne					+	+		+	+						4
Mg	Ukształtowanie powierzchni	+			+											2
	Genetyczne formy terenu							+/-								0 (1)
	(An) Antropogeniczne formy rzeźby						+			+						2
	(An) Zagrożenia powierzchni terenu degradacją						+									1
Pd	Genetyczne i litologiczne typy gleb												+	+	+	3
	(An) Klasy bonitacyjne gleb													+		1
	(An) Kompleksy rolniczej przydatności gleb												+			1
Li	Wydzielenia litologiczno-genetyczno-stratygraficzne							+								1
		8 (9)	3 (5)	8 (11)	1	12 (14)	16 (19)	1 (2)	6	11 (13)	2	11 (12)	2	9	10	
		BDO	CLC	VML2	DTED2	MHP	MSP	SMGP	MHgP	MGP	MPHP	BDOT	MGR	EGiB	LMN	

\*) rozumiane jako obszary gospodarowania przestrzenią geograficzną przez człowieka

- jednorodność i rzetelność źródeł danych oraz względna kompletność atrybutów opisowych rozpatrywanych baz występuje w bazach referencyjnych (EGiB, LMN, BDOT, BDO) oraz w specjalistycznych bazach tematycznych (SMGP, MGR, DTED, CLC) (ryc. 3.9.). Pozostałe bazy powstawały w oparciu o niejednorodne i często nieprecyzyjne dane, co skutkuje niedostatkami danych i obniżoną ich precyzją;
- zasięg opracowania i przestrzenno-czasowa spójność danych są zróżnicowane i uzależnione warunkami tworzenia poszczególnych baz danych (ryc. 3.9.). Nowsze bazy obejmują obszar całego kraju i powstawały w krótkim okresie czasu (BDO, VML2, MPHP, DTED, CLC). Najstarsze opracowania bazodanowe, powstające w układach arkuszowych Międzynarodowej Mapy Świata są zadaniami długofalowymi i obecnie scala się je do opracowań regionalnych (np. MHP, MSP, MGP).

Należy podkreślić, że wszystkie rozpatrywane bazy danych wnoszą istotny wkład w opis środowiska przyrodniczego. Jedne z nich skupiają się na wybranych komponentach środowiska (SMGP, MGR, DTED, MPHP, CLC), podczas gdy inne prezentują większość komponentów w aspekcie topografii terenu.

A.

Cel tworzenia bazy	Dostępność standardu	Nominalna skala bazy	Obecność atrybutów	Źródła danych	Zasięg opracowania
DLM topologiczny	dostępny, precyzyjny	większa lub równa 1:10 000	pełna + system identyfikatorów	rejestry urzęd. I geodezja	>60%
DLM/DCM topologiczny	dostępny, nieprecyzyjny	równa skali 1:50 000 <sup>*)</sup>	niepełna	topografia, baza własna, ...	30-60%
model nietopologiczny	niedostępny dla użytkowników	mniejsza lub równa 1:100 000	baza uboga w atrybuty	nieznane, baza uboga	<30%

<sup>\*)</sup> w przypadku MGR skala 1:25000

B.

	BDO	VML2	BDOT	MHP	MSP	SMGP	MHgP	MGP	MPHP	MGR	EGIB	LMN	DTED	CLC
Cel tworzenia bazy														
Dostępność standardu														
Nominalna skala bazy														
Obecność atrybutów														
Źródła danych														
Zasięg opracowania										b.d.		b.d.		
	Baza Danych Ogólnogeograficznych	Mapa Wektorowa Poziomu Drugiego	Baza Danych Obiektów Topograficznych	Mapa Hydrograficzna Polski	Mapa Sozologiczna Polski	Szczegółowa Mapa Geologiczna Polski	Mapa Hydrogeologiczna Polski	Mapa Geośrodowiskowa Polski	Mapa Podziału Hydrograficznego Polski	Mapa Glebowo-Rolnicza	Ewidencja Gruntów i Budynków	Leśna Mapa Numeryczna	Numeryczny Model Terenu VML2	Mapa Pokrycia Terenu CORINELandCover

Ryc. 3.9. Zestawienie waloryzujące najważniejsze cechy baz danych przestrzennych:

A – skale ocen, B – wynik.

Źródło: opracowanie własne

## 4. BADANIE JAKOŚCI DANYCH PRZESTRZENNYCH

### 4.1. Zakres badania jakości danych

Przyporządkowanie w poprzednim rozdziale warstw tematycznych baz danych przestrzennych do komponentów środowiska przyrodniczego pozwoliło na wskazanie obszarów redundancji danych w badanych bazach (tab. 3.3.), a co za tym idzie – na identyfikację baz szczególnie wymagających waloryzacji. W przypadku niektórych elementów środowiska przyrodniczego (np. sieć hydrograficzna, sieć osadnicza, linie komunikacyjne, elementy pokrycia terenu) stopień nadmiarowości danych, wynikający z występowania tych samych klas obiektów w różnych bazach, jest bardzo wysoki. Dlatego w niniejszym rozdziale, w procedurze oceny jakości baz danych zastosowano metodę wstępnej oceny jakości danych wektorowych (Kaczmarek 2007, Kaczmarek, Medyńska-Gulij 2008).

Jakość danych przestrzennych powinna podlegać bezpośredniej weryfikacji, niezależnie od teoretycznych założeń, które wynikają z instrukcji i wytycznych technicznych. Takie miary jakości danych, jak kompletność, spójność logiczna, dokładność położenia, dokładność czasu i dokładność tematyczna (Bielecka 2006, 2010, EN ISO 19113:2005), wymagają empirycznego wyznaczenia bezwzględnych lub względnych wartości błędów danych.

Zakres badania jakości danych oraz stosowane metody zależą głównie od celu prowadzenia procedury. Badanie jakości można prowadzić z punktu widzenia producenta lub użytkownika danych (Bielecka 2010). Producent określa stopień zgodności bazy danych z przyjętymi założeniami produkcyjnymi (instrukcje, wytyczne techniczne). Natomiast użytkownik odnosi swoją ocenę do oczekiwań w zakresie tematycznej zawartości tej bazy oraz istotnych z jego punktu widzenia parametrów dokładnościowych.

- W tym rozdziale jakość baz danych przestrzennych rozpatruje się w dwóch ujęciach:
- ogólnych wskaźników oceny, pozwalających na zestawienia waloryzacyjne całych baz danych (por. tab. 4.1. i 4.2. w rozdz. 4.2.),
  - miar dokładności geometrycznej i lokalizacyjnej oraz kompletności obiektów w ramach konkretnych warstw tematycznych (rozdz. 4.3.).

## 4.2. Miary jakości danych

W dalszej części pracy wykorzystywane będą pojęcia elementów jakości danych oraz elementów podrzędnych jakości danych, zdefiniowanych w normie EN ISO 19913:2005, dlatego konieczne jest przytoczenie tych określeń.

Element jakości danych jest składnikiem ilościowym dokumentującym jakość zbioru danych.

W normie zdefiniowano pięć elementów jakości danych:

- kompletność – rozumianą jako „(...) *obecność i nieobecność obiektów, ich atrybutów i związków*”;
- spójność logiczną – określaną jako „(...) *stopień zgodności z logicznymi regułami struktury danych, atrybutów i związków*”;
- dokładność położenia – czyli „(...) *dokładność położenia obiektów*”;
- dokładność czasową – przyjętą jako „(...) *dokładność związanych z czasem atrybutów i związków obiektów*”;
- dokładność tematyczną – zdefiniowaną jako „(...) *dokładność atrybutów ilościowych i poprawność atrybutów nieilościowych oraz klasyfikacji obiektów i ich związków*”.

Norma definiuje ponadto elementy podrzędne jakości danych, czyli składniki wymienionych powyżej elementów jakości danych opisujące pewne aspekty tych elementów. W niniejszej pracy wykorzystuje się wyłącznie elementy podrzędne w zakresie spójności logicznej:

- „ - *spójność pojęciowa: zgodność z regułami schematu pojęciowego,*
- *spójność dziedziny: zgodność wartości z dziedzinami wartości,*
- *spójność formatu: do jakiego stopnia dane są zapisane zgodnie z fizyczną strukturą zbioru danych,*
- *spójność topologiczna: poprawność jawnie zapisanych cech topologicznych zbioru danych*”.

## 4.3. Wstępna ocena jakości baz danych

Znajomość ogólnej struktury grup warstw tematycznych w poszczególnych bazach pozwoliła na wskazanie baz, które w sposób pełny (znak + w tab. 3.3.) bądź niepełny charakteryzują dany składnik komponentowy (znak +/- w tab. 3.3.). Przy czym określenie „pełna charakterystyka składnika komponentowego” jest pojęciem umownym i oznacza prawdopodobną możliwość wykorzystania danej grupy warstw tematycznych określonej bazy danych w dalszej procedurze oceny. Warto zauważyć, że większość składników komponentowych wybranych w tab. 3.3. jest opisywanych przez więcej niż jedną, dwie bazy danych. Jednocześnie bazy danych, które opisują dany składnik komponentowy często należą

do różnych przedziałów skal nominalnych. Taka sytuacja powoduje pojawienie się trudności w wyborze bazy danych jako źródła informacji, właściwego z punktu widzenia celu zastosowania. Stąd zastosowano w kolejnym etapie badania przydatności baz danych, wykorzystanie zaproponowanego przez Kaczmarka i Medyńską-Gulij (2008) systemu wstępnej oceny jakości baz danych (tab. 4.1.). System opiera się na punktacji trzech głównych składowych, decydujących wg autorów o ogólnej jakości danych: kompletności obiektów przestrzennych, dokładności geometrycznej oraz dokładności tematycznej i czasowej. Każda składowa jest oceniana w czterostopniowej skali cyfrowej od 0 do 3 (tab. 4.1.).

*Tabela 4.1. System wstępnej oceny jakości cyfrowych baz danych  
(wg Kaczmarek, Medyńska-Gulij 2008, zmienione)*

<b>CHARAKTERYSTYKA</b>	<b>PUNKTY</b>
<b>1. Kompletność obiektów przestrzennych</b>	
akcesoryczna (tylko wybrane kategorie obiektów)	<b>0</b>
niepełna nietopologiczna lub wymagająca przetworzenia	<b>1</b>
niepełna topologiczna	<b>2</b>
pełna	<b>3</b>
<b>2. Dokładność geometryczna</b>	
bardzo niska (skala dokładności opracowania $\leq 1:250\ 000$ )	<b>0</b>
niska (skala dokładności opracowania $\leq 1:100\ 000$ i $> 1:250\ 000$ )	<b>1</b>
przeciętna (skala dokładności opracowania $\leq 1:50\ 000$ i $> 1:100\ 000$ )	<b>2</b>
wysoka (skala dokładności opracowania $> 1:50\ 000$ )	<b>3</b>
<b>3. Dokładność tematyczna i czasowa</b>	
trudno zdefiniować cechy bazy atrybutowej lub baza jest pusta	<b>0</b>
dane atrybutowe o różnej aktualności, częściowo z nieznanymi źródłami lub z wieloma błędami, niekompletna baza danych	<b>1</b>
dane atrybutowe pozyskane z rzetelnych źródeł o różnej aktualności i / lub niekompletną bazą danych	<b>2</b>
dane atrybutowe pozyskane z rzetelnych źródeł, spójne czasowo i w miarę kompletna baza danych	<b>3</b>

Spośród trzydziestu składników komponentów środowiska przyrodniczego, uwzględnionych w tab. 3.2. wybrano przykładowych osiem i dla nich zastosowano uproszczoną metodę wstępnej oceny jakości. Wynik oceny zaprezentowano w tab. 4.2.

Najwyższe oceny w zakresie kompletności, dokładności geometrycznej oraz dokładności tematycznej i czasowej uzyskiwały bazy danych o jakościach geodezyjnych. Są to bazy EGiB oraz LMN, a także korzystająca z geodezyjnych źródeł danych BDOT. Warto podkreślić, że wysokie oceny warstw wód powierzchniowych i zlewni w bazie Mapy Hydrograficznej Polski. Wyniki te przewyższają wartości uzyskane przez nowoczesną bazę

Mapy Podziału Hydrograficznego Polski. Zdecydowała o tym nominalna skala opracowań, korzystna dla MHP oraz niedomiar obiektów hydrograficznych MPHP wynikający z generalizacji sieci rzecznej. Pozostałe przykładowe dane analizowane metodą wstępnej oceny jakości danych wskazywały, z nielicznymi wyjątkami, znacznie niższe wartości

Zastosowanie tej uproszczonej metody oceny jakości baz danych pozwala na wytypowanie danych do dalszej procedury. Oczywiście do wyniku oceny należy podchodzić ostrożnie, mając świadomość, że zbliżoną do rzeczywistości ocenę jakości baz danych można dokonać wyłącznie w trybie badań szczegółowych. Jednakże zastosowana metoda wstępnej oceny jakości baz danych pozwala na prawidłowe wytypowanie zestawu danych do dalszej procedury oceny, co przy znacznej redundancji danych (por. tab. 3.3.) ma znaczenie fundamentalne.

#### **4.4. Badanie jakości wybranych danych przestrzennych**

Szczegółowej analizy wybranych warstw tematycznych dokonano w celu zweryfikowania wyników wstępnej oceny jakości baz danych przeprowadzonej metodą zaproponowaną przez Kaczmarka i Medyńską-Gulij (2008). Ponadto uzyskane wartości oceny jakościowej będą podstawą określania w rozdziale 5 cząstkowych miar jakości danych, które stanowią składowe potencjału informacyjnego w zakresie jakości danych. Do rozważań wybrano obiekty hydrograficzne, prezentowane w sposób kompleksowy w trzech bazach danych: Mapie Hydrograficznej Polski (MHP), Mapie Podziału Hydrograficznego Polski (MPHP) i Wektorowej Mapie Poziomu Drugiego (VML2). Przedmiotem zainteresowania były cechy geometryczne obiektów oraz kompletność obiektów. Dane porównywano ze sobą oraz odnoszono do materiałów referencyjnych, którymi były mapa topograficzna 1:10000 w układzie współrzędnych 1992 oraz pomiary terenowe zbiorników wodnych wykonane przez autora.

**Tabela 4.2. Zestawienie wybranych grup warstw baz danych przestrzennych z zastosowaniem systemu wstępnej oceny jakości opracowań numerycznych**  
**Źródło: opracowanie własne**

Wybrane składniki komponentów środowiska przyrodniczego		BDO	CLC	VML2	MHP	MSP	MHgP	MGP	MPHP	BDOT	MGR	EGiB	LMN
Nominalna skala opracowania		1:250 000	1:10 0000	1:50 000						1:10 000	1:25 000 1:5 000	1:5 000	
An	Formy użytkowania terenu*)	5 2/0/3	6 2/1/3	6 2/2/2		6 2/2/2				n.b.		8 3/3/2	8 3/3/2
	Obszary zabudowy mieszkalnej	5 2/0/3	6 2/1/3	6 2/2/2	6 2/2/2	6 2/2/2		6 2/2/2		9 3/3/3		9 3/3/3	9 3/3/3
	Granice jednostek administracyjnych	8 3/2/3		7 3/2/2	6 3/1/2	6 3/1/2				9 3/3/3		9 3/3/3	9 3/3/3
Bi	Zasięg lasów i roślinności nieleśnej	4 2/0/2	5 2/1/2	5 2/2/1	4 2/2/0	5 2/2/1		5 2/2/1		n.b.		7 3/3/1	9 3/3/3
	(An) Formy ochrony przyrody	6 2/1/3		5 2/2/1		6 2/2/2		6 2/2/2		n.b.			
Hy	Wody powierzchniowe	5 1/1/3	3 1/1/1	5 1/2/2	9 3/3/3	5 1/2/2	5 1/2/2	5 1/2/2	7 2/2/3	9 3/3/3		8 3/3/2	8 3/3/2
	Działy wodne, zlewnie				9 3/3/3		6 2/2/2		8 3/2/3				
Pd	Genetyczne i litologiczne typy gleb										8 3/3/2	8 3/3/2	8 3/3/2

\*) rozumiane jako obszary gospodarowania przestrzenią geograficzną przez człowieka

n.b. – nie badano (dane nie są jeszcze dostępne)

- W badaniach jakościowych wykorzystano szereg metod, takich, jak:
- analizy zakresu semantycznego danych,
  - porównanie wskaźników kształtu, lokalizacji i powierzchni obiektów z wykorzystaniem ich cech geometrycznych (obliczanie wartości cech geometrycznych obiektów oraz wskaźników tych cech, obliczenie odchyleń od wskaźników wzorcowych),
  - analiza cech geometrycznych obiektów w polach jednostkowych (generowanie wartości geometrycznych obiektów w polach 10 x 10 km, obliczanie wskaźników dla pól oraz odchyleń wartości między poszczególnymi bazami danych).

W realizacji tych metod wykorzystano głównie narzędzia geoinformacyjne, w szczególności operacje na danych (Kaczmarek 2010) oraz proste obliczenia statystyczne. Dla wyników poszczególnych analiz zastosowano ponadto metodę kartograficznych wizualizacji porównawczych oraz syntetyczne zestawienia tabelaryczne wyników.

Opis badania jakości danych wykonano osobno dla powierzchniowych wód płynących i powierzchniowych wód stojących. Empiryczne badania jakości baz danych pokazują tylko pewien wycinek zagadnienia, testujący zarówno warsztat badawczy, jak i konkretny zasób danych z obszaru województwa wielkopolskiego.

#### 4.4.1. Badanie jakości danych o wodach płynących

Powierzchniowe wody płynące są prezentowane w większości spośród rozpatrywanych baz danych, jednak najlepszej prezentacji można się spodziewać po cyfrowych opracowaniach kartograficznych (VML2) lub tematycznych związanych z hydrografią (MHP, MPHP). Najwyższe wskaźniki jakości we wstępnej ocenie jakości (tab. 4.2.) uzyskały jednak tylko te ostatnie.

Do procedury oceny w tej grupie źródeł danych wybrano następujące warstwy tematyczne:

- 1) dla Mapy Hydrograficznej Polski (nazwy wg Wytycznych Technicznych K-3.4., w nawiasie podano nazwy warstw wg wytycznych GIS-3):
  - ciek\_n (cieki\_z\_nazwa) – ciek z nazwą (obiekty liniowe)
  - ciek\_bn (cieki\_bez\_nazwy) – ciek bez nazwy (obiekty liniowe)
  - kanał (kanały) – kanały (obiekty liniowe)
  - baza uzupełniająca: bazawody.dbf (aktualna z Internetu)
- 2) dla Mapy Podziału Hydrograficznego Polski
  - r\_rzeki - ciek wyróżnione (obiekty liniowe)
  - rzekin – odcinki pozostałych cieków (obiekty liniowe)

- baza uzupełniająca: n\_dod\_cieki.dbf

3) dla Mapy Wektorowej Poziomu Drugiego (w nawiasie podano kody VPF2PL według aplikacji słownikowej VmapDict, ver. 4.02)

- rivera (ABH140) – rzeka/strumień\_c (obiekty powierzchniowe)
- riverl (LBH140) – rzeka/strumień\_l (obiekty liniowe)
- canala (ABH020) – kanał\_c (obiekty powierzchniowe)
- canall (LBH020) – kanał\_l (obiekty liniowe)
- ditcha (ABH030) – rów melioracyjny\_c (obiekty powierzchniowe - obiekty nie występują w badanym obszarze)
- ditchl (LBH030) – rów melioracyjny\_l (obiekty liniowe)
- dane uzupełniające:
  - \* rivcharc (PBH142) – charakterystyka rzeki/kanału\_p (obiekty punktowe)
  - \* rivcharl (LBH142) – charakterystyka rzeki/kanału\_l (obiekty liniowe)
  - \* curflowp (PBG010) – kierunek prądu\_p (obiekty punktowe)

Semantyczne zależności między powyższymi warstwami zaprezentowano na grafie (ryc. 4.1.).

Mapa Hydrograficzna Polski	Mapa Podziału Hydrograficznego Polski	Mapa Wektorowa Poziomu Drugiego
cieki_n	rzeki_r (zawiera również osie obiektów z warstwy s-rzeki)	rivera riverl (część)
kanal		canala canall
cieki_bn	rzekin	ditcha ditchl
		riverl (część)

**Ryc. 4.1. Znaczeniowy zakres poszczególnych warstw tematycznych płynących wód powierzchniowych trzech wybranych baz danych przestrzennych**  
**Źródło: opracowanie własne**

W strukturze VML2 występuje niezgodność opisu warstw z rzeczywistym grupowaniem kanałów melioracyjnych. Otóż dostępne są odrębne warstwy canall (LBH020) i canala (ABH020) gromadzące kanały żeglowne oraz o charakterze melioracyjnym (atrybut OPS=0), a odrębnie warstwy ditchl (LBH030) i ditcha (ABH030) zawierające rowy melioracyjne. Ten podział potwierdza słownik kodowania VML2 (VmapDict, ver. 4.02). Natomiast ogólnie dostępny słownik obiektów VML2 nie zawiera opisu do warstw ditchl i ditcha. Nie należy jednak traktować tej niezgodności jako uszczegółowienia bazy względem założeń, ponieważ nie ma żadnego wyraźnego rozgraniczenia obiektów występujących na warstwach opisujących kanały oraz rowy. Nie jest nim zarówno atrybut szerokości (WWR), ani statusu eksploatacji (OPS). W VML2 pominięto suche rowy – warstwa drditchl (LPL050) oraz parametry związane z przeprawami (brody – warstwa ford1 i ford2)

Warstwy wód płynących poddano analizom przestrzennym w kwadratach 10 x 10 km oraz analizom zasobności informacyjnej.

Analizy przestrzenne miały na celu:

- porównanie gęstości sieci rzecznej w poszczególnych bazach (woj. wielkopolskie),
- przeliczenie wartości długości cieków z baz wektorowych w regularnej siatce kwadratów,
- oszacowanie dokładności lokalizacyjnej cieków (niewielkie obszary testowe) – w tym celu wybrano sześć arkuszy mapy topograficznej w skali 1:10000 (PUWG1992) i na tych obszarach określono wartości wektorów przesunięć cieków w poszczególnych bazach względem cieków na mapie,
- określenie modelu topologicznego powiązań sieciowych.

Jakość danych źródłowych pod względem dokładności lokalizacyjnej i geometrii wynika głównie z zastosowanego źródła danych przestrzennych. W przypadku VML2 były to powojenne mapy wojskowe 1:50000 wykonane w układzie 1942. Prawdopodobnie taki sam materiał był wykorzystany przy wektoryzacji cieków dla MPHP (Gotlib i in. 2007a). Natomiast wytyczne techniczne Mapy Hydrograficznej Polski nakładają konieczność stosowania opracowań kartograficznych w skali 1:25000 i/lub 1:10000 jako materiału podstawowego.

#### *Analiza wskaźnika gęstości sieci rzecznej*

Celem tej analizy było porównanie stopnia zagęszczenia sieci rzecznej na obszarze województwa wielkopolskiego prezentowanej w bazach MHP, MPHP i VML2. Określono gęstości cieków: dla wszystkich warstw poszczególnych baz danych (ryc. 4.2. i tab. 4.3.) oraz dla wybranych warstw podobnych pod względem semantycznym (por. ryc. 4.4.).

Długości cieków zostały obliczone dla oczek regularnej siatki kwadratów 10 x 10 km skonstruowanej w układzie współrzędnych PUWG1992. Ze względu na zasięg bazy MHP pozostałe bazy również przycięto do granic administracyjnych województwa, a oczka siatek obszaru granicznego zamknięto granicą województwa.

Dla celów porównawczych wprowadzono wskaźnik gęstości sieci hydrograficznej, który należy rozumieć jako średniej wartości długości cieków przypadającej na jednostkę powierzchni (w poniższym przypadku w km/km<sup>2</sup>).

Ogólne wartości wskaźników gęstości sieci rzecznej dla poszczególnych baz danych prezentuje tab. 4.3.

Najkorzystniejsze wskaźniki gęstości sieci hydrograficznej w woj. wielkopolskim uzyskała MHP, a najniższe posiada VML2. Różnice sięgają aż 20% i są bezpośrednią miarą szczegółowości odzwierciedlenia sieci hydrograficznej (wartości w kolumnach „wszystkie”).

Wartości wskaźnika w odniesieniu do poszczególnych warstw tematycznych mogą być porównywane właściwie wyłącznie w odniesieniu do MHP i MPHP. Wynika to z podobnego zakresu semantycznego warstw tych opracowań (por. ryc. 4.1.). I tak, porównując wartości dotyczące warstwy ciek\_i\_n oraz kanal z bazy MHP z warstwą rzeki\_r z opracowania MPHP, widoczna jest sięgająca 15% przewaga MPHP, jednak baza MHP jest niemal o 20% bardziej zasobna w obiekty cieków bez nazwy (warstwa ciek\_i\_bn). Te różnice wydają się bezpośrednio wynikać z jednej strony z rzetelności i konsekwencji twórców MPHP w zakresie definiowania sieciowych powiązań między obiektami hydrograficznymi, a z drugiej szczegółowości zastosowanego materiału źródłowego do wektoryzacji obiektów hydrograficznych.

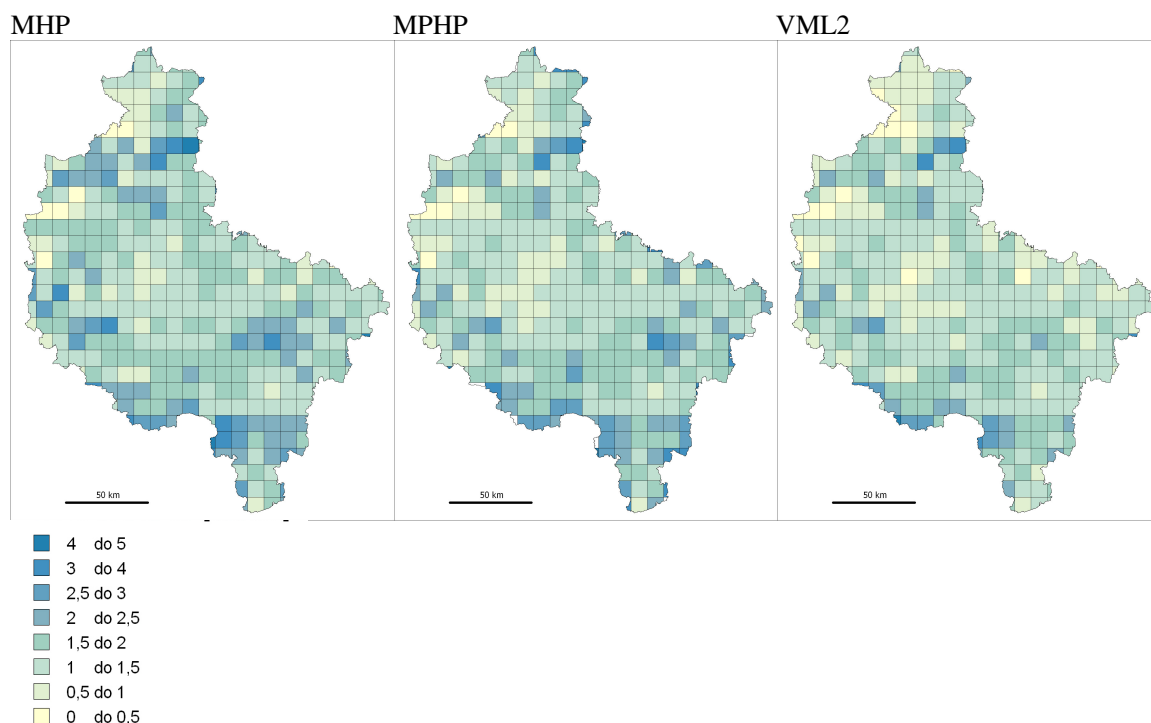
**Tabela 4.3. Wartości wskaźnika gęstości sieci hydrograficznej dla poszczególnych warstw tematycznych**

*Źródło: opracowanie własne*

Parametr	MHP				MPHP			VML2			
	Cieki_n	Cieki_bn	Kanal	WSZYSTKIE	Rzeki_r	Rzekin	WSZYSTKIE	Rivera, riverl	Ditchl	Canala, canall	WSZYSTKIE
Ogólna długość cieków [km]	103,62	464,61	6,26	<b>574,08</b>	127,72	395,98	<b>523,60</b>	302,10	169,57	6,24	<b>477,75</b>
Wartość maksymalna wskaźnika gęstości [km/km <sup>2</sup> ]	1,41	4,32	1,13	<b>4,98</b>	1,46	3,61	<b>4,44</b>	3,33	3,51	0,99	<b>4,06</b>
Wartość średnia wskaźnika gęstości [km/km <sup>2</sup> ]	0,29	1,30	0,02	<b>1,61</b>	0,36	1,11	<b>1,47</b>	0,85	0,47	0,02	<b>1,34</b>

Przestrzenne zróżnicowanie gęstości sieci rzecznej zaprezentowane na ryc. 4.2. wskazuje na duże podobieństwo w rozkładzie tego wskaźnika.

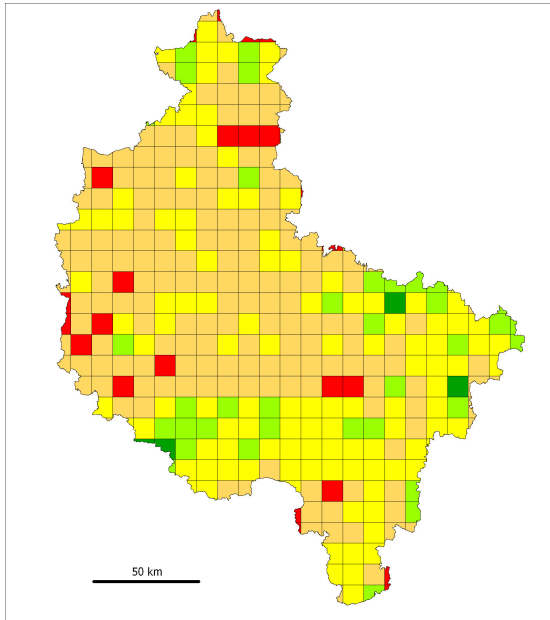
Wyraźnie widoczna jest ogólna przewaga wskaźnika gęstości sieci rzecznej MHP nad pozostałymi opracowaniami (por. ryc. 4.3.). Wartość różnicy wskaźników zaprezentowaną na powyższych rycinach można interpretować jako bezwzględną różnicę średniej wartości długości cieków przypadającą na jednostkę powierzchni ( $\text{km}^2$ ), a wyrażoną w kilometrach. W przypadku analizy porównawczej MHP i MPHP oraz MHP i VML2 ta różnica nierzadko sięga wartości 1 (odchylenie standardowe odpowiednio 0,25 i 0,23), co może być miarą niejednorodności kryteriów pozyskiwania danych źródłowych w jednym z opracowań. Zestawiając tę uwagę z dość jednorodnym obrazem różnic wskaźnika dla MPHP i VML2 (odchylenie standardowe 0,16) można wnioskować o niejednorodności pozyskiwania danych do MHP. Wynikać ona może z arkuszowej realizacji tej bazy przez różnych wykonawców.



**Ryc. 4.2. Gęstość sieci rzecznej w  $\text{km}/\text{km}^2$  w poszczególnych bazach danych**  
*Źródło: opracowanie własne*

Należy jednak zauważyć istotne różnice w wartościach tego wskaźnika przy porównaniu różnych źródeł danych. W bardziej przystępny sposób to zróżnicowanie prezentuje ryc. 4.3. oraz ryc. 4.4.

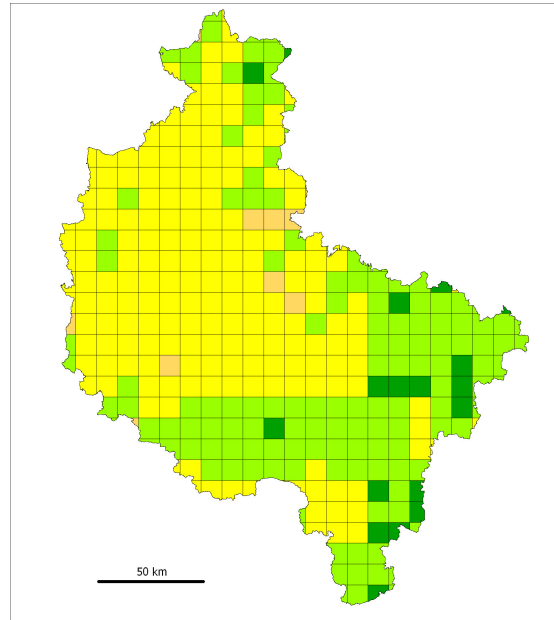
W zakresie warstw zawierających ciek nazwane zanotowano ogólną zgodność między MHP i MPHP przy ogólnej przewadze danych MPHP (ryc. 4.4.).



**Różnice w gęstości sieci rzecznej**

(wsk\_g\_MPHP-wsk\_g\_MHP)

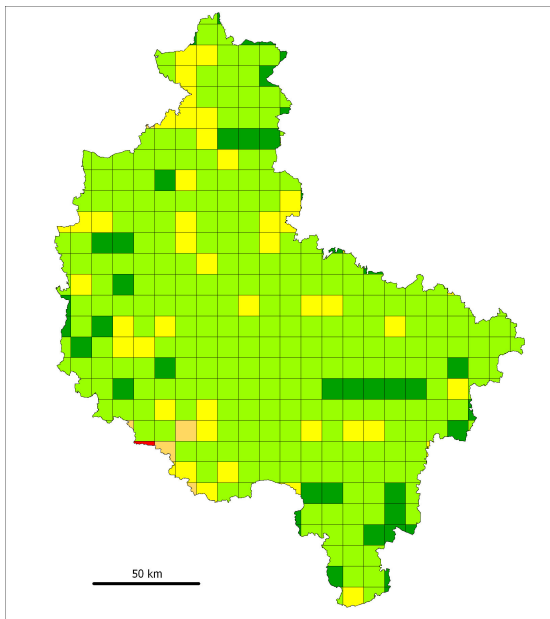
- 0,5 do 1,03 (istotna przewaga MPHP)
- 0,1 do 0,5 (przewaga MPHP)
- -0,1 do 0,1 (równowaga)
- -0,5 do -0,1 (przewaga MHP)
- -1,15 do -0,5 (istotna przewaga MHP)



**Różnice w gęstości sieci rzecznej**

(wsk\_g\_MPHP - wsk\_g\_VML2)

- 0,5 do 0,83 (istotna przewaga MPHP)
- 0,1 do 0,5 (przewaga MPHP)
- -0,1 do 0,1 (równowaga)
- -0,36 do -0,1 (przewaga VML2)



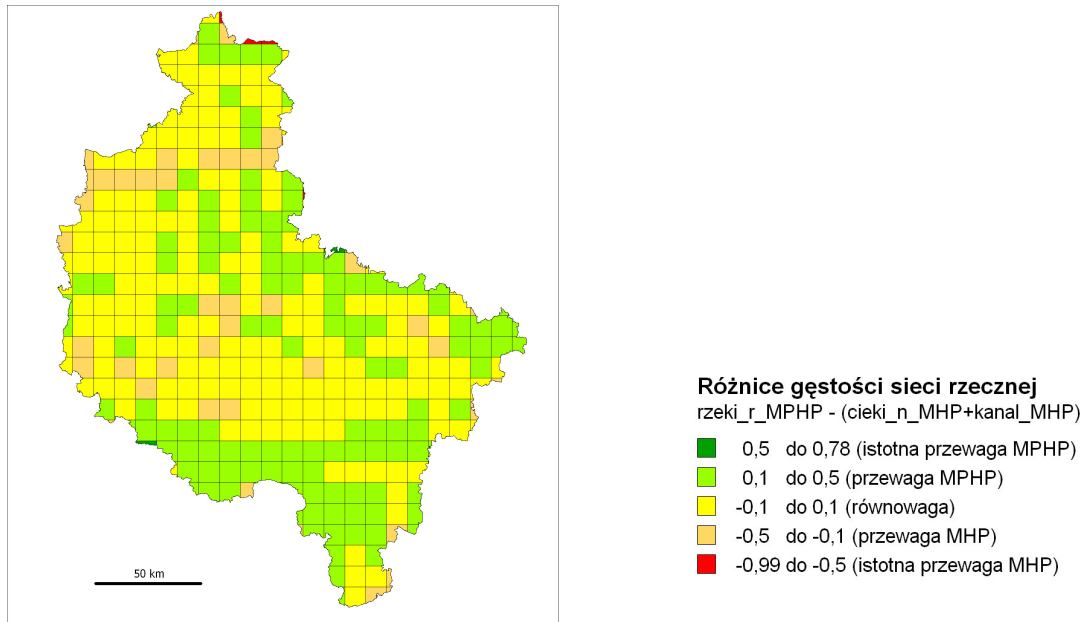
**Różnice w gęstości sieci rzecznej**

(wsk\_g\_MPH - wsk\_g\_VML2)

- 0,5 do 1,25 (istotna przewaga MPH)
- 0,1 do 0,5 (przewaga MPH)
- -0,1 do 0,1 (równowaga)
- -0,5 do -0,1 (przewaga VML2)
- -0,58 do -0,5 (istotna przewaga VML2)

**Ryc. 4.3. Matryca różnic wskaźnika gęstości sieci hydrograficznej poszczególnych baz danych w siatce 10 x 10 km**

**Źródło: opracowanie własne**



**Ryc. 4.4. Matryca różnic wskaźnika gęstości sieci hydrograficznej bazy MHP i MPHP w siatce 10 x 10 km w zakresie warstw cieków zawierających nazwę**  
**Źródło: opracowanie własne**

#### Analiza geometrii i lokalizacji

W celu przeanalizowania geometrii oraz dokładności lokalizacyjnej zastosowano dwa proste i obiektywne wskaźniki liczbowe:

-wskaźnik kształtu (WK) – stosunek długości odcinka cieków na badanej warstwie do długości wzorcowej tego odcinka:

$$WK = L_w / L_b$$

gdzie:

$L_w$  - dł. cieków wzorcowego [m],

$L_b$  - dł. cieków w bazie [m];

wskaźnik przyjmuje zatem wartości niemianowane (tab. 4.4.); zakładając wybór do analizy wyłącznie cieków, których przebieg jest zgodny ze wzorcem wartość wskaźnika oscylująca w pobliżu wartości 1 świadczyć może o braku generalizacji kształtu rzeki względem wzorca; wartości wskaźnika mniejsze od 1 świadczą o uproszczeniu geometrii przebiegu cieków na badanej warstwie względem wzorca; jako dobry wzorzec wybrano do badań mapę topograficzną 1:10 000 wykonaną w drugiej połowie lat 90. XX. wieku, aktualizowaną w oparciu o ortofotomapę;

- wskaźnik błędu lokalizacji (BL) – jako średnia odległość między wzorcowym odcinkiem cieków, a jego przebiegiem na badanej warstwie:

$$BL = P_b / L_w$$

gdzie:

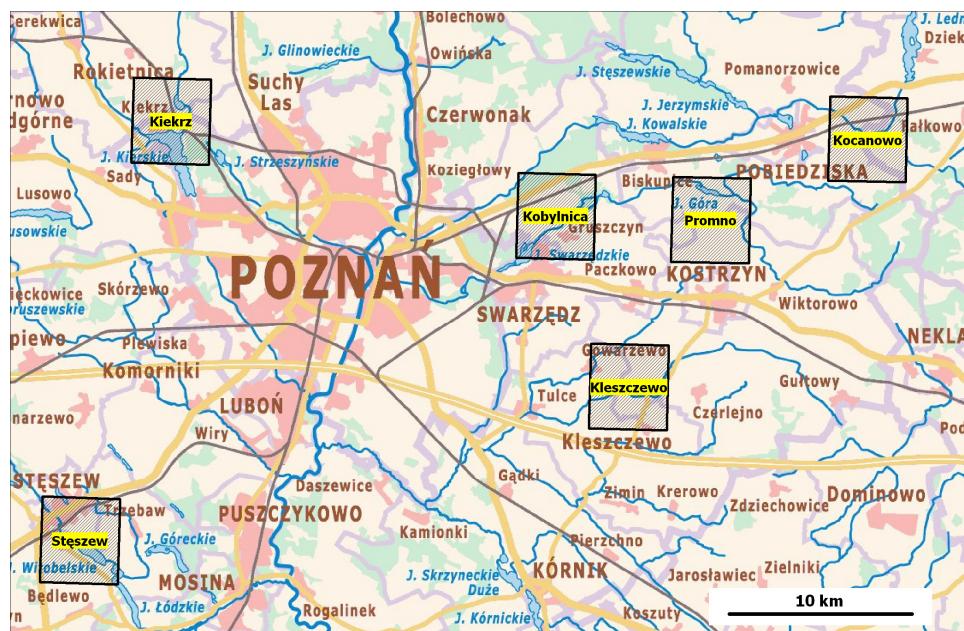
$P_b$  - powierzchnia obszaru między ciekami wzorcowymi a ciekami w bazie [m<sup>2</sup>],

$L_w$  - dł. cieków wzorcowych [m];

Wskaźnik ten uzyskano poprzez obliczenie powierzchni jaką zajmuje obszar ograniczony przez wzorcowy i badany przebieg odcinka cieków i podzielenie jej przez długość odcinka wzorcowego cieków; Wskaźnik błędu lokalizacji jest zatem wyrażony w jednostkach metrycznych (tab. 4.5.); wzorcem, jak przy wskaźniku kształtu jest tu mapa topograficzna w skali 1:10 000.

Do badań wybrano losowo obszary arkuszy mapy topograficznej 1:10 000 w układzie PUWG 1992 (lokalizacja na ryc. 4.5.):

Godło	Nazwa arkusza
N-33-130-D-a-1	Kiekrz
N-33-131-C-a-4	Kobylnica
N-33-131-C-b-4	Promno
N-33-131-C-d-3	Kleszczewo
N-33-131-D-a-2	Kocanowo
N-33-142-A-b-4	Stęszew



Ryc. 4.5. Lokalizacja pól testowych do badań wskaźników kształtu i błędu lokalizacji cieków  
Źródło: opracowanie własne

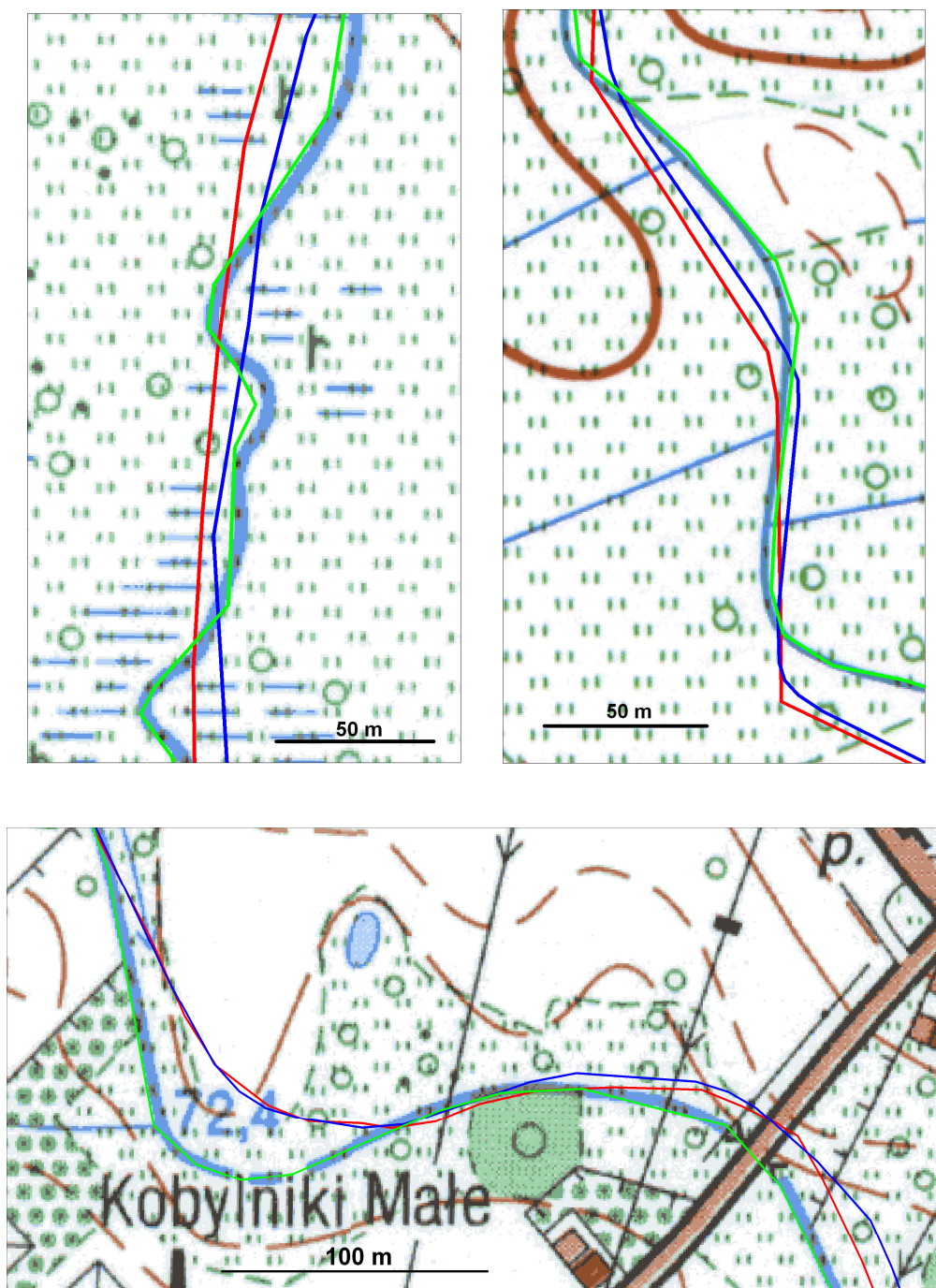
W wyniku nałożenia warstw numerycznych na zasięgi pól testowych wytypowano 19 odcinków testowych należących do dziewięciu cieków (por. tab. 4.4. i tab. 4.5.). Łączna długość badanych cieków wg mapy topograficznej 1:10 000 wynosi 38,18 km.

Wyniki analizy wskaźnika kształtu (tab. 4.4.) ukazują dużą zgodność kształtu cieków MHP ze wzorcem, co potwierdza materiał ilustracyjny (ryc. 4.6.). Wyniki te potwierdzają stosowanie map topograficznych w skalach 1:25 000 oraz 1:10 000 jako podkładu do wektoryzacji cieków przy produkcji MHP. Wyniki wskaźnika kształtu dla MPHP oraz VML2 wykazują znaczne (sięgające 5-10% długości cieków) uproszczenie przebiegu cieków względem wzorca. Wyniki te potwierdzają stosowanie map topograficznych 1:50 000 jako źródła przy wektoryzacji cieków do MPHP i VML2.

**Tabela 4.4. Wartości wskaźnika kształtu dla testowych odcinków cieków**  
*Źródło: opracowanie własne*

Nazwa cieków na MT10/92	Długość w m na MT10/92	wskaźnik kształtu		
		MHP	MPHP	VML2
Bogdanka	196,4	1,02	0,89	0,95
Cybina_1	4 440,1	0,99	0,95	0,95
Cybina_2	1 807,6	1,00	0,99	0,98
Cybina_3	624,3	0,99	0,96	0,96
Cybina_4	2 063,8	1,10	0,95	0,95
Główna_1	901,6	1,00	0,97	0,98
Główna_2	5 845,6	1,00	0,98	0,98
Kopla_1	3 179,3	0,99	0,99	0,98
Męcina_1	4 340,2	1,00	0,95	0,95
Męcina_2	1 512,8	1,00	0,98	brak
Przeźmierka_1	512,0	0,98	1,00	1,00
Przeźmierka_2	445,0	1,01	0,99	0,98
Przeźmierka_3	763,8	0,96	0,91	0,89
Samica Stęszewska_1	3 119,4	1,00	0,99	0,98
Samica Stęszewska_2	672,2	1,01	0,98	0,91
Samica_1	1 364,0	1,00	0,95	0,95
Samica_2	1 987,2	1,00	0,93	brak
Samica_3	1 013,1	1,00	0,96	0,97
Trzebawka	2 394,7	1,00	1,00	brak
<b>RAZEM</b>	<b>37 183,1</b>	<b>1,003</b>	<b>0,964</b>	<b>0,960</b>

Drugi wskaźnik – błąd lokalizacji określił średnie oddalenie obiektu cieków w badanych bazach danych od wzorcowego przebiegu koryta cieków wyrażone w metrach (tab. 4.5.). Ten wskaźnik potwierdza dużą precyzję lokalizacyjną obiektów zawartych w bazie MHP.



*Ryc. 4.6. Trzy przykłady obrazujące dokładność lokalizacyjną i geometryczną MHP (linia zielona), MPHP (linia niebieska) oraz VML2 (linia czerwona) na tle mapy topograficznej w skali 1:10000  
 Źródło: opracowanie własne*

**Tabela 4.5. Wskaźnik błędu lokalizacji cieków w poszczególnych bazach względem wzorca dla testowych odcinków cieków**  
**Źródło: opracowanie własne**

Nazwa ciek	dlugość	powierzchn.	wsk_błędu	powierzchn.	wsk_błędu	powierzchn.	wsk_błędu
	MT10	MHP		MPHP		VML2	
Bogdanka	196,4	658,9	3,35	1321,5	6,73	742,6	3,78
Cybina_1	4440,1	12899,4	2,91	78064,6	17,58	95476,3	21,5
Cybina_2	1807,6	4270,8	2,36	10869,5	6,01	15283,4	8,46
Cybina_3	624,3	665,1	1,07	3820,6	6,12	3582,3	5,74
Cybina_4	2063,8	6584,8	3,19	17212,9	8,34	19930,6	9,66
Główna_1	901,6	2399,4	2,66	4364,6	4,84	6930,8	7,69
Główna_2	5845,6	10652,6	1,82	53016,8	9,07	50588,3	8,65
Kopla_1	3179,3	13768,9	4,33	25737,9	8,1	26802,1	8,43
Męcina_1	4340,2	6940	1,6	29215,4	6,73	28331,9	6,53
Męcina_2	1512,8	4624,6	3,06	9169,7	6,06		
Przeźmierka_1	512	690,8	1,35	12194,5	23,82	9872,4	19,28
Przeźmierka_2	445	554,1	1,25	4051,6	9,1	4071,7	9,15
Przeźmierka_3	763,8	2037,7	2,67	9028,7	11,82	7266,4	9,51
Samica Sęszewska_1	3119,4	5743	1,84	15560,7	4,99	28510,1	9,14
Samica Sęszewska_2	672,2	2211,5	3,29	5244	7,8	983,5	1,46
Samica_1	1364	1649,6	1,21	13882,8	10,18	11698,1	8,58
Samica_2	1987,2	2047,9	1,03	54897,5	27,63		
Samica_3	1013,1	2339,6	2,31	8764	8,65	9252	9,13
Trzebawka	2394,7	511,3	0,21	7322,2	3,06		
<b>ŚREDNI BŁĄD LOKALIZACJI</b>			<b>2,18</b>		<b>9,82</b>		<b>9,17</b>

#### 4.4.2. Badanie jakości danych o wodach stojących

Jak wynika ze wstępnej oceny jakości baz danych (tab. 4.2.) najbardziej rzetelnym źródłem danych dotyczących wód powierzchniowych są bazy tematyczne (MHP, MPHP), a w mniejszym stopniu kartograficzne (VML2). Dane są zgromadzone w tych opracowaniach na następujących warstwach tematycznych:

- 1) dla Mapy Hydrograficznej Polski (nazwy wg Wytycznych Technicznych K-3.4., w nawiasie podano nazwy warstw wg wytycznych GIS-3):
  - zb\_wod (zbiorniki\_wodne) - zbiorniki wodne (dla uporządkowania informacji rozdzielono tę warstwę na obiekty powierzchniowe – zb\_wod\_region oraz obiekty punktowe – zb\_wod\_point)
  - zb\_izob (izobaty) – izobaty (obiekty liniowe)
  - baza uzupełniająca: bazawody.dbf (aktualna z Internetu)

2) dla Mapy Podziału Hydrograficznego Polski

- jeziora - zbiorniki wyróżnione (obiekty powierzchniowe)
- jez\_n – zbiorniki pozostałe (obiekty powierzchniowe)
- baza uzupełniająca: n\_dod\_jeziora.dbf (dot. warstwy jeziora)

3) dla Mapy Wektorowej Poziomu Drugiego (w nawiasie podano kody VPF2PL według aplikacji słownikowej VmapDict, ver. 4.02)

- lakea (ABH08) – jezioro/staw\_C (obiekty powierzchniowe)
- fishfrma (ABH050) – staw rybny/morska hod. ryb\_C (obiekty powierzchniowe)
- reservra (ABH130) – rezerwuuar/zbiornik\_C (obiekty powierzchniowe)
- islanda (ABA030) – wyspa\_C (obiekty powierzchniowe)
- depthcl (LBE015) – izobata\_L (obiekty liniowe)

(UWAGA: w przypadku obiektów jezior nie dających się przedstawić w skali mapy instrukcja wykonawcza zaleca umieszczać je na warstwie PBI010 – lockc; jednak w rzeczywistości na obszarze woj. wielkopolskiego umieszczono tylko trzy takie obiekty w związku z czym warstwę lock pominięto w analizach)

W zależnościach znaczeniowych warstw zachodzą trzy zasadnicze różnice między wybranymi źródłami danych:

- w MHP wszystkie obiekty wód powierzchniowych stojących znajdują się w jednej warstwie, ale obiekty mniejsze niż 1 250 m<sup>2</sup> przedstawiane są jako obiekty punktowe,
- w MPHP zbiorniki wodne podzielono na dwie warstwy, z których pierwsza („jeziora”) zawiera obiekty nazwane i skatalogowane,
- w VML2 z grupy zbiorników wodnych wydzielono grupę sztucznych zbiorników, najczęściej o charakterze retencyjnym (reservra) oraz stawy rybne (fishfrma).

Dla zbiorników wodnych województwa wielkopolskiego wykonano analizy w zakresie:

- wskaźnika jeziorności – dla porównania stopnia pokrycia oczek regularnej siatki pól podstawowych obszarami wód stojących dla poszczególnych baz danych; wykorzystano wszystkie dostępne w bazie obiekty powierzchniowych wód stojących dla obszaru woj. wielkopolskiego
- wskaźnika rozwinięcia linii brzegowej – w celu określenia dokładności odzwierciedlenia linii brzegowej w poszczególnych bazach danych; wykorzystano grupę zbiorników wodnych o powierzchni większej lub równej 10 ha,
- wskaźnika różnic powierzchni – aby określić stopień zróżnicowania wielkościowego poszczególnych zbiorników w różnych bazach danych; wykorzystano grupę zbiorników wodnych o powierzchni większej lub równej 10 ha.

Dokładność lokalizacyjna obiektów utworzonych w wybranych bazach danych wynika z jakości źródła wykorzystanego przy tworzeniu bazy. Dane te były pozyskiwane przez wykonawców tych baz z różnych źródeł, zgodnie z założeniami z instrukcji. W przypadku MHP (Wytyczne Techniczne K-3.4, 1997; Wytyczne Techniczne GIS-3, 2005) podstawę geometryczną determinowały cywilne mapy topograficzne (najczęściej w skali 1:25 000) i zdjęcia lotnicze. Dla VML2 podstawą digitalizacji były wojskowe mapy topograficzne 1:50 000 w układzie współrzędnych 1942. Baza MPHP nie posiada ogólnie dostępnej informacji precyzującej warunki pozyskiwania danych.

#### *Analiza wskaźnika jeziorności*

W celu porównania zasobności wybranych baz danych: MHP, MPHP i VML2 w zbiorniki wodne dokonano analizy w regularnej siatce kwadratów 10 x 10 km. Do analizy wybrano wszystkie obiekty wód powierzchniowych umieszczone w badanych bazach, a znajdujące się na obszarze województwa wielkopolskiego. Wybór dotyczy wyłącznie obiektów umieszczonych w bazach jako obiekty powierzchniowe.

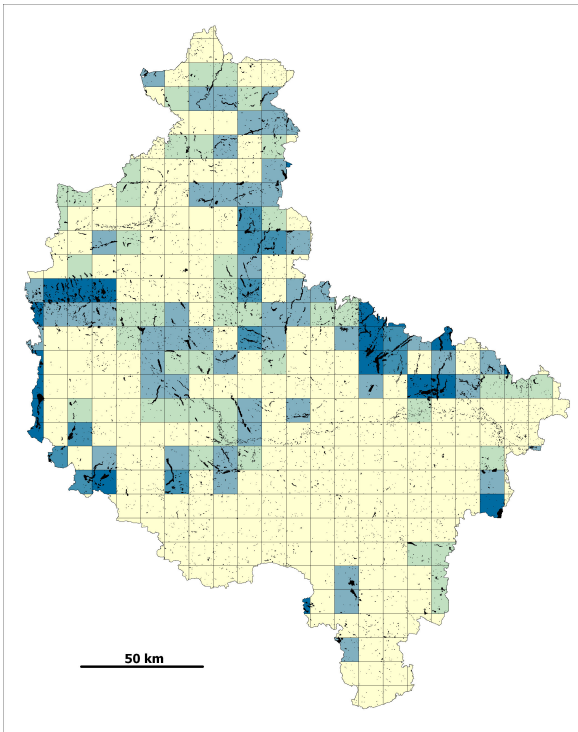
Za wskaźnik jeziorności przyjęto stosunek powierzchni zajmowanej przez zbiorniki wodne w poszczególnych oczkach siatki do powierzchni oczka siatki pól podstawowych. Ten prosty wskaźnik obrazuje jaką część pola jednostkowego zajmuje powierzchnia wodna. Wyniki analizy zostały zebrane w poniższej tabeli (tab. 4.6.).

**Tabela 4.6. Uogólnione wartości wskaźnika jeziorności rozpatrywanego w siatce 10 x 10 km**  
**Źródło: opracowanie własne**

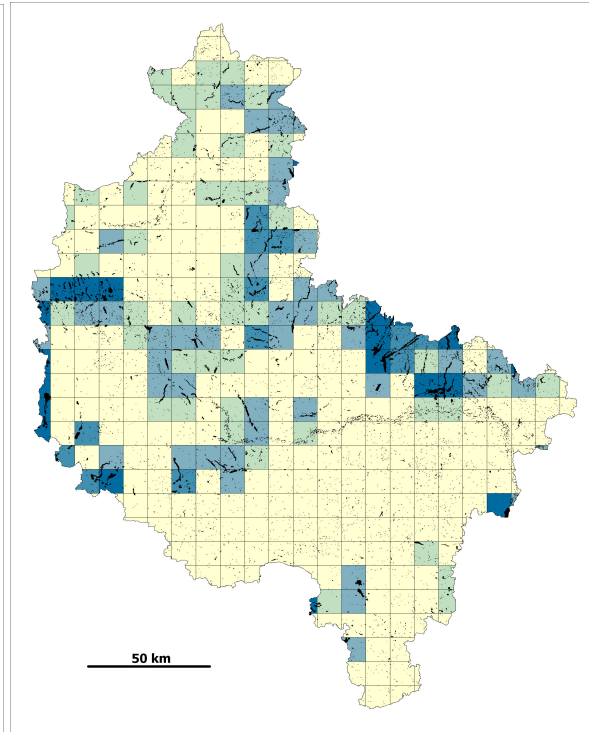
<b>Wskaźnik jeziorności</b>	<b>MHP</b>	<b>MPHP</b>	<b>VML2</b>
Wartość maksymalna	0,1784	0,1885	0,1788
Wartość średnia	0,014342	0,0142457	0,0123064
Odchylenie standardowe	0,024018	0,0243613	0,0221782
Suma wartości wskaźnika	5,1201	5,0857	4,9334
Powierzchnia zbiorników wodnych [km <sup>2</sup> ]	393,218	385,382	373,564

Zbiornicze wyniki analiz wskazują na największy odsetek powierzchni zbiorników wodnych w bazie MHP, a najmniejszy na VML2. Odzwierciedla się to w wartościach średnich i sumach wskaźnika jeziorności (tab. 4.6., ryc. 4.14.). Należy jednak zauważyć, że różnice między bazami są niewielkie (nieznacznie przekraczają 5%). Również zróżnicowanie przestrzenne wskaźnika jeziorności zobrazowane na ryc. nie wskazuje na znaczne lokalne różnice w rozkładzie tego wskaźnika dla poszczególnych baz.

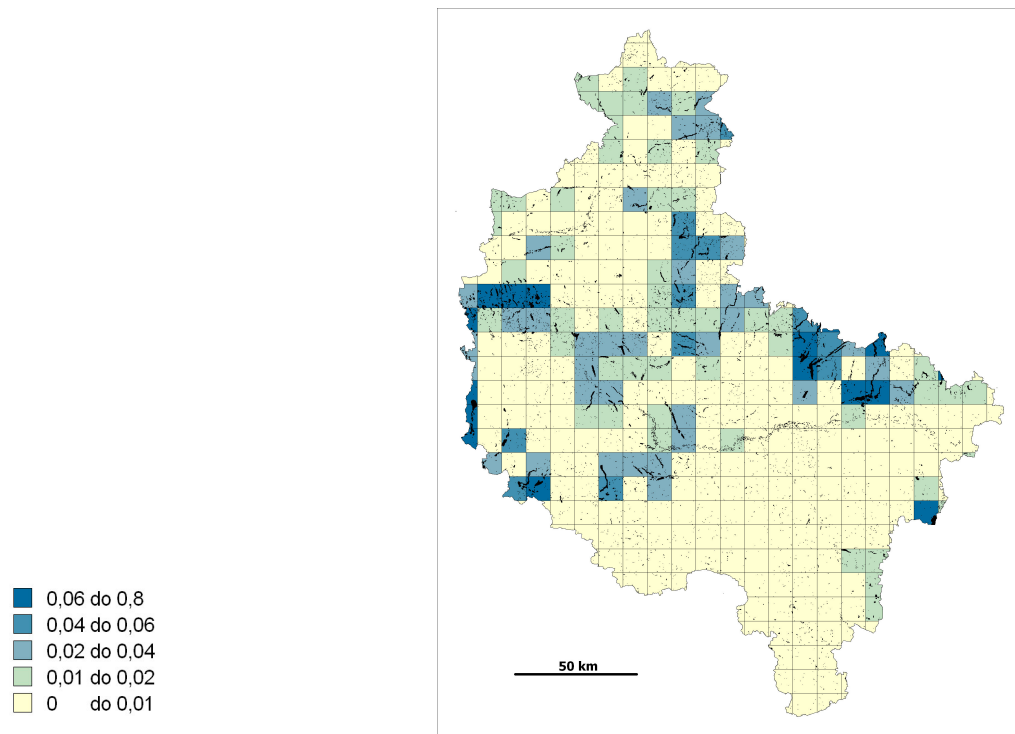
MHP



MPHP

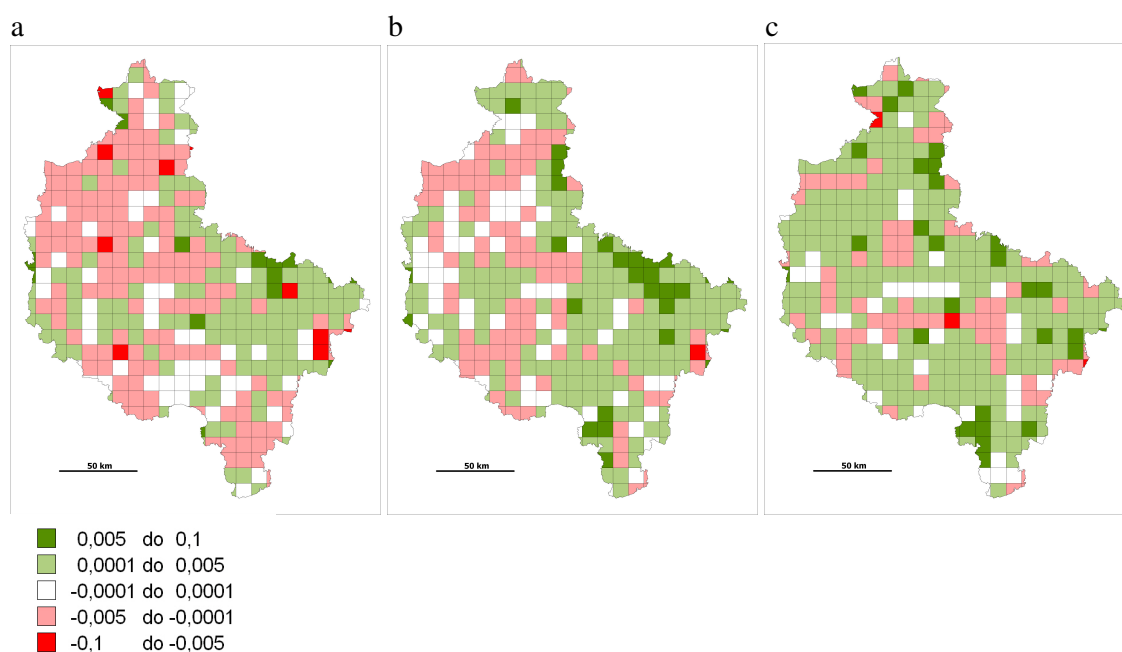


VML2



Ryc. 4.7. Wskaźnik jeziorności (jednostki niemianowane) w poszczególnych bazach danych  
Źródło: opracowanie własne

W celu uwidocznienia lokalnych różnic w jakości dostarczanych danych dotyczących zbiorników wodnych obliczono różnice między wskaźnikami jeziorności poszczególnych baz danych. Wyniki zilustrowano na ryc. 4.8. Widać na niej wyraźnie, że różnice powierzchni przekraczające 0,5% (różnica wskaźnika  $> 0,005$ ), zaznaczone na rycinie kolorem ciemnozielonym oraz czerwonym, dotyczą jedynie niewielkich obszarów województwa. Można więc stwierdzić homogeniczność tych baz pod względem prezentacji obszarów wód powierzchniowych. Należy jednak pamiętać, że ten wskaźnik nie przesądza o jakości danych wynikającej z lokalizacji i geometrii poszczególnych obiektów. Badania tych parametrów zostaną zaprezentowane w następnej części.



**Ryc. 4.8. Matryca różnic wskaźnika jeziorności w poszczególnych bazach danych w siatce 10 x 10 km: a) MPHP-MHP, b) MPHP-VML2, c) MHP-VML2**  
*Źródło: opracowanie własne*

#### *Analiza geometrii i lokalizacji zbiorników wodnych*

Do analizy kształtu i dokładności lokalizacyjnej zbiorników wodnych zastosowano cztery wskaźniki:

- wskaźnik rozwinięcia linii brzegowej,
- różnice powierzchni zbiorników wodnych,
- wskaźnik błędu lokalizacji,
- wskaźnik błędu batymetrii.

Do określenia pierwszych dwóch wskaźników wykorzystano zbiór jezior zawierających nazwy, wyłonionych z wybranych baz w następujący sposób:

Etap 1 z baz danych wyłoniono zbiorniki nazwane

Etap 2 wybrano jeziora o powierzchni równej lub większej 10 ha

Etap 3 wykluczono jeziora, które ze względów technicznych zostały niepoprawnie scalone (dotyczy bazy VML2)

W wyniku powyższej procedury wytypowano 267 zbiorników wodnych z terenu woj. wielkopolskiego, które spełniają trzy powyższe kryteria w każdej z trzech testowanych baz. Do określenia wskaźnika błędu lokalizacji oraz batymetrii wybrano dziesięć jezior, dla których autor wykonał w latach 2002-2003 szczegółowe pomiary batymetryczne.

Wskaźnik rozwinięcia linii brzegowej (RLB) określono jako stosunek długości linii brzegowej jeziora do długości okręgu o powierzchni równej powierzchni jeziora. Wynika z tej definicji, że wartość wskaźnika będzie liczbą niemianowaną i może przyjmować wartości większe lub równe 1, przy czym dla  $RLB=1$  jezioro w zarysie ma kształt okręgu.

Dla wytypowanych 267 jezior obliczono wskaźnik RLB, a wyniki pogrupowano w czterech klasach wartości wskaźnika (tab. 4.7.).

Zauważane prawidłowości w tym zestawieniu dotyczą m.in. następujących zagadnień:

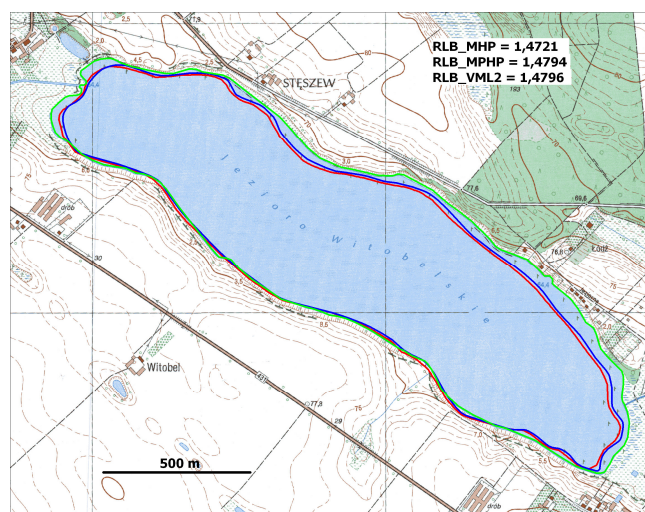
- w pierwszej klasie obiektów (o najbardziej regularnych kształtach,  $RLB \leq 1,5$ ) przeważają zbiorniki mniejsze,
- w trzeciej klasie obiektów (przedział  $RLB (2;3>)$ ) przeciętne zbiorniki są dwukrotnie większe niż w klasie 2
- obliczone w tab. 4.7. różnice wskaźnika RLB dla różnych baz jednoznacznie wskazują na najwyższe wskaźniki RLB dla bazy MHP

**Tabela 4.7. Wskaźniki rozwinięcia linii brzegowej (RLB) dla wybranych jezior woj. wielkopolskiego w bazach MHP, MPHP i VML2**

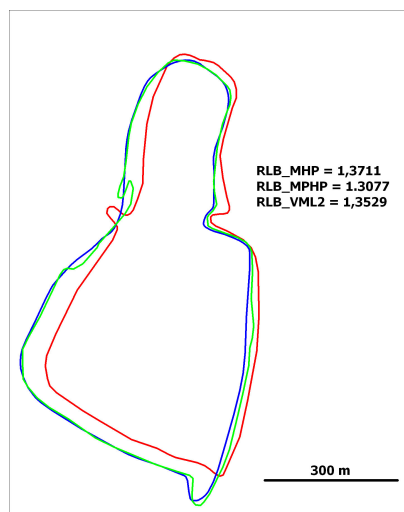
*Źródło: opracowanie własne*

Klasa RLB	Pow. jezior [ha] wg MHP	Wartości średnie RLB			Różnice RLB		
		dla MHP (ilość obiektów)	dla MPHP (ilość obiektów)	dla VML2 (ilość obiektów)	MPHP-MHP	MPHP-VML2	MHP-VML2
		średnia suma odch. std	średnia suma odch. std	średnia suma odch. std	średnia suma odch. std	średnia suma odch. std	średnia suma odch. std
<1;1,5>	2 894,64	1,29441 (98)	1,28277 (99)	1,29063 (104)	-0,0116 -1,1405 0,0408	-0,0079 -0,7707 0,0878	0,0038 0,3698 0,0918
(1,5;2>	6 168,88	1,68839 (100)	1,68758 (104)	1,67886 (98)	-0,0008 -0,0805 0,0789	0,0087 0,8723 0,0687	0,0095 0,9528 0,0757
(2;3>	6 899,18	2,37723 (54)	2,33088 (50)	2,33675 (51)	-0,0464 -2,5031 0,1591	-0,0059 -0,3172 0,1104	0,0405 2,1859 0,1670
>3	2 600,86	3,38447 (15)	3,27867 (14)	3,31835 (14)	-0,1058 -1,5870 0,1066	-0,0397 -0,5952 0,1010	0,0661 0,9918 0,1323
<b>RAZEM</b>	<b>18 563,56</b>	<b>1,77838 (267)</b>	<b>1,75849 (267)</b>	<b>1,76153 (267)</b>	<b>-0,0200 -5,3111 0,0970</b>	<b>-0,0030 -0,8108 0,0882</b>	<b>0,0169 4,5003 0,1104</b>

a)



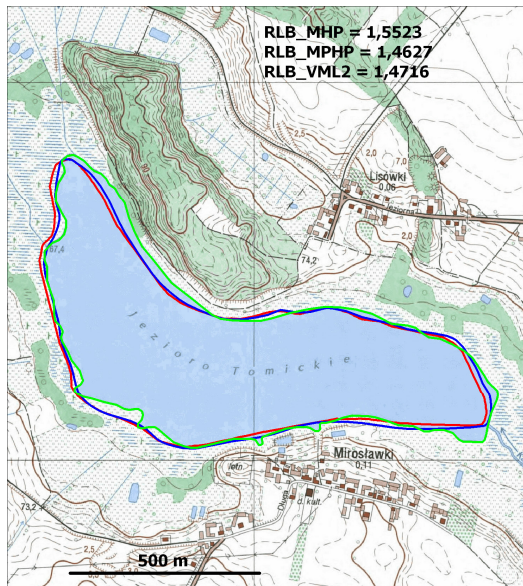
b)



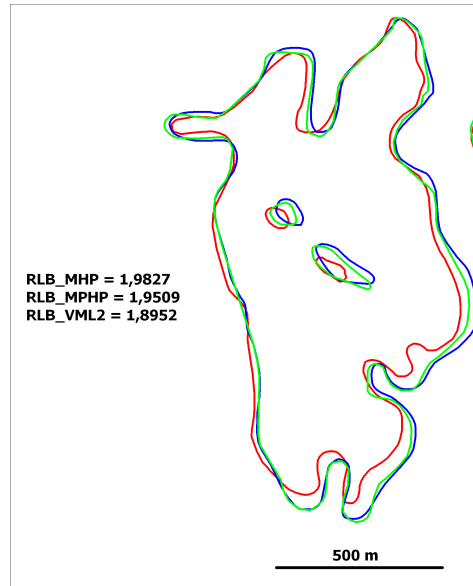
**Ryc. 4.9. Przykładowe jeziora o wskaźniku RLB <1;1,5> a) Witobelskie, b) Boszkowskie**

*Źródło: opracowanie własne*

a)

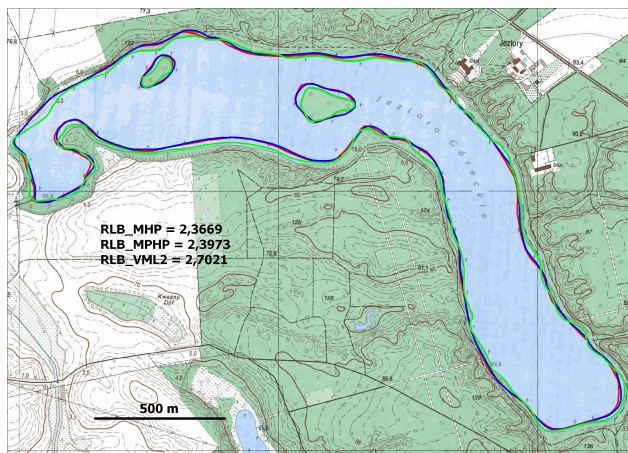


b)

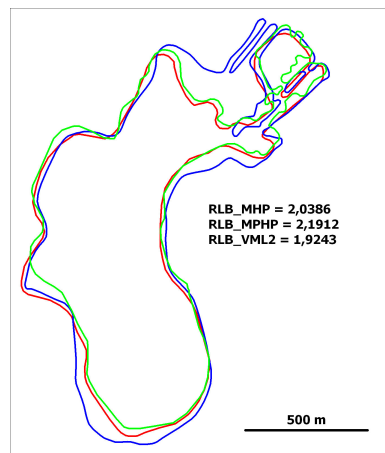


Ryc. 4.10. Przykładowe jeziora o wskaźniku RLB (1,5;2>): a) Tomickie, b) Bielskie  
Źródło: opracowanie własne

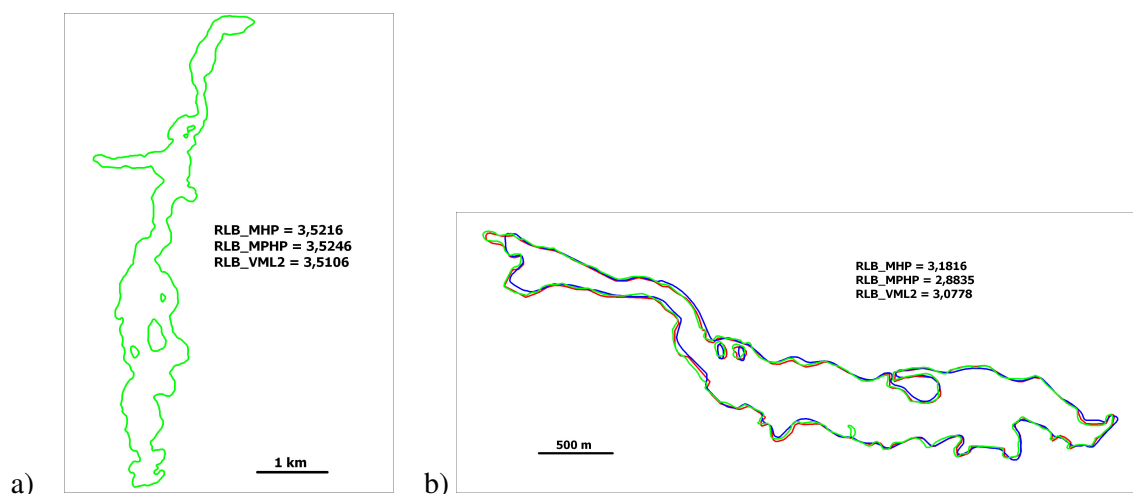
a)



b)



Ryc. 4.11. Przykładowe jeziora o wskaźniku RLB (2;3>): a) Góreckie, b) Mąkolno  
Źródło: opracowanie własne



**Ryc. 4.12. Przykładowe jeziora o wskaźniku RLB > 3: a) Lednica, b) Kowalskie**  
**Źródło: opracowanie własne**

Podsumowując, badania jakości danych przestrzennych przyniosły szereg spostrzeżeń dotyczących zarówno ograniczeń metody, jak i identyfikacji ogólnych cech jakości badanych baz danych. W szczególności:

1. Przeprowadzona dla wybranych danych z baz MHP, VML2 i MPHP próba empirycznego wyznaczenia miar jakości wskazała na problemy oceny, z których najważniejsze to:
  - brak jednorodnego modelu pojęciowego w różnych bazach, co utrudniało porównywalność wyników;
  - konieczność kompleksowego spojrzenia na miary jakości (np. na miarę dokładności geometrycznej składają się miary: długości, powierzchni, kształtu i lokalizacji obiektów);
  - jakość danych może być określana najczęściej w sposób względny poprzez odniesienie do danych przyjętych jako wzorcowe (np. mapa topograficzna, ortofotomapa, dane z pomiaru terenowego lub baza referencyjna).
2. Badania empiryczne jakości danych w zakresie kompletności, dokładności geometrycznej i lokalizacyjnej na przykładzie wód powierzchniowych dla baz MHP, VML2 i MPHP wykazały, że:
  - analiza obiektów wód powierzchniowych dla obszaru województwa wielkopolskiego udokumentowała wysoką jakość MHP oraz szereg istotnych braków i niedokładności baz MPHP i VML2;
  - wskaźniki kształtu obiektów jednoznacznie potwierdzają źródła pozyskiwania danych (mapa topograficzna w skali 1:10 000 lub 1:25 000 w przypadku MHP oraz mapa 1:50 000 dla VML2 i prawdopodobnie dla MPHP).

## **5. WYZNACZANIE I WYKORZYSTANIE POTENCJAŁU INFORMACYJNEGO DANYCH PRZESTRZENNYCH**

W tym rozdziale zaprezentowano koncepcję potencjału informacyjnego danych, a w szczególności czynniki mające główny wpływ na jego wartość (rozdział 5.1.). Uwaga była skupiona wokół atrybutów obiektów przestrzennych, które stanowiły nie tylko miarę dokładności tematycznej, ale były głównym czynnikiem wpływającym na pojemność informacyjną bazy danych, a także na analityczne i wizualizacyjne funkcje bazy. Dlatego też autor niniejszej rozprawy zaproponował podejście do klasyfikacji poziomów ujęcia atrybutów (rozdział 5.2.). W dalszej części rozdziału (5.3.) wskazano przykładowe procedury wyznaczania wskaźnika potencjału informacyjnego danych przestrzennych. Na koniec, w rozdziale 5.4. zaprezentowane zostało ujęcie dynamiczne potencjału informacyjnego, uwzględniające pełnienie przez bazę funkcji akumulacyjnej, analitycznej i wizualizacyjnej.

### **5.1. Koncepcja potencjału informacyjnego danych przestrzennych**

Jak określono we wstępie rozprawy potencjał informacyjny danych przestrzennych to miara zasobności modeli obiektów przestrzennych (encji) w informację o cechach tych obiektów, istotnych z punktu widzenia indywidualnego użytkownika danych. Zakłada się, że wartość potencjału może być liczbą bezwzględną, wartością względną lub cechą opisową.

W określaniu potencjału informacyjnego urzędowych baz danych przestrzennych przyjmuje się założenie, że elementy jakości danych (w rozumieniu normy ISO 19113) takie, jak kompletność, spójność logiczna, dokładność położenia i dokładność czasowa są zgodne ze specyfikacjami produktu lub rozpoznane w wyniku przeprowadzenia procedury oceny jakości danych. Specyfikację produktu stanowią ściśle zdefiniowane procedury tworzenia baz danych wydane w postaci dokumentu (instrukcji lub wytycznych technicznych). Badanie jakości danych jest czynnością uzupełniającą tę wiedzę o elementy istotne dla użytkownika danych z punktu widzenia jego oczekiwań, np. stopień kompletności danej klasy obiektów względem własnego wzorca. Badanie może obejmować wybrane klasy obiektów lub cały zbiór danych. Może się ono ponadto odbywać na całym obszarze zainteresowań lub tylko w polach testowych. Dobór metod i narzędzi oraz zakresu merytorycznego i przestrzennego oceny jakości danych zależy od indywidualnych potrzeb użytkownika danych. Przykładowe

procedury tak rozumianej oceny zaprezentowano w rozdziale 4. Przedmiotowe bazy danych mogą posiadać opracowane metadane, i wtedy one są źródłem informacji o jakości danych.

Istotą potencjału informacyjnego danych przestrzennych jest szczegółowe określenie miary jakości w zakresie dokładności tematycznej danych. Przy tym „dokładność tematyczna danych” nie jest ściśle rozumiana, jak w normie ISO 19113 jako poprawność danych atrybutowych i ich dokładność. Dane atrybutowe należy traktować jako czynnik decydujący o możliwości: 1) przechowywania danych na warstwie tematycznej (funkcja akumulacyjna), 2) realizacji analiz geoinformacyjnych (funkcja analityczna) i 3) prezentacji kartograficznej danych (funkcja wizualizacyjna). Jeśli dane atrybutowe są rzetelne i przechowują informację cenną z punktu widzenia użytkownika to zrozumiałe jest, że im więcej atrybutów przechowuje baza, tym jest bardziej wartościowa jako źródło danych. Bogactwo atrybutów o odpowiednio wysokim poziomie dokładności przekłada się bezpośrednio na możliwości analityczne w oparciu o taką warstwę, a także na wzrost funkcjonalności warstwy w zakresie wizualizacji kartograficznej.

Podstawową jednostką będącą przedmiotem określania potencjału musi być zawsze pojedyncza warstwa tematyczna. Aby uzyskać porównywalną miarę syntetyczną potencjału informacyjnego dla kolekcji grup warstw tematycznych różnych baz danych należy:

1. Wyselekcjonować w ramach grupy (lub grup) warstw tematycznych warstwy istotne dla celu badania i określić ich znaczenie w prezentowaniu określonego zjawiska (w postaci współczynnika wagowego);
2. Uśrednić wartości potencjałów informacyjnych poszczególnych warstw z uwzględnieniem współczynników wagowych warstw (średnie ważone).

Konieczność przyjęcia warstwy jako podstawowej jednostki określania potencjału informacyjnego wynika z faktu, że najczęściej przy przedstawianiu określonego zjawiska (lub klasy obiektów) w bazach danych, informacja jest umieszczana na kilku warstwach. Warstwy te mogą posiadać różny zakres informacji atrybutowej, stąd niemożliwe byłoby jednoznaczne określenie zawartości tematycznej. Przykładem może tu być informacja o zbiornikach wodnych w bazach MHP, MPHP i VML2 (por. rozdz. 4). Większe jeziora posiadają z reguły bogaty zestaw danych atrybutowych podczas, gdy pozostałe zbiorniki są ubogie w dane.

Jeżeli porównywane są dane z różnych źródeł należy pamiętać o konieczności spełnienia warunku zgodności semantycznej między warstwami (grupami warstw) z różnych baz danych. Zastosowanie sumy iloczynów potencjałów cząstkowych z ich współczynnikami wagowymi pozwoli na uzyskanie porównywalnej wartości dla określonej klasy obiektów dla różnych baz danych przestrzennych. W przypadku braku zgodności zakresów znaczeniowych

danych należy w każdej bazie danych wyselekcjonować grupy obiektów zgodnych semantycznie lub odstąpić od porównywania potencjałów.

Tak rozumiana koncepcja potencjału informacyjnego baz danych przestrzennych odnosi się wprost do zwaloryzowania jakości bazy danych dla indywidualnych działań użytkownika, a więc do konkretnych jego potrzeb wobec bazy danych. Miara potencjału informacyjnego powinna być sumą miar ilościowych jakości według ISO: kompletności, spójności logicznej oraz dokładności czasowej i przestrzennej, a także dokładności tematycznej, która jest kluczowa przy określaniu potencjału. Dokładność tematyczna jest przy tym rozumiana szerzej niż w normie ISO 19113 – jako cecha warstwy tematycznej, która w decydujący sposób wpływa na ilość i jakość przechowywanych danych oraz na możliwości analityczne i wizualizacyjne warstwy. Warunkiem porównywalności miar potencjału informacyjnego jest ściśle określenie semantycznego zakresu poszczególnych fizycznych warstw geoinformacyjnych oraz nadanie im właściwego współczynnika.

## **5.2. Poziomy ujęcia atrybutów**

W podsumowaniu poprzedniego rozdziału określono, że zasadniczym czynnikiem wielkości potencjału informacyjnego danych przestrzennych jest poziom dokładności tematycznej rozumianej jako zdolność do pełnienia przez bazę danych funkcji akumulacyjnej, analitycznej i wizualizacyjnej. Określenie poziomu dokładności tematycznej będzie polegało na uporządkowanej analizie danych atrybutowych zapisanych w strukturze bazy danych. Istotne z punktu widzenia potencjału informacyjnego będą zarówno format przechowywania danych atrybutowych, jak i precyzja wyrażania wartości atrybutu. Do sklasyfikowania tych cech wykorzystano pojęcie skal pomiarowych. Odniesienie tego pojęcia do cyfrowych baz danych, a nie wyłącznie do produktów kartograficznych wymagało nowego podejścia. Odniesienie skal pomiarowych do baz danych przestrzennych było sygnalizowane do tej pory w nielicznych publikacjach (Żyszkowska 2000, Kraak, Ormelling 2006, Bielecka 2006).

Skale pomiarowe są traktowane przez różnych autorów z odmienną szczegółowością: od trzech poziomów (np. Korycka-Skorupa 2002a) aż po kilkanaście i więcej wydzieleni (np. Forrest 1999). Wynika to z różnego podejścia do tego pojęcia:

- 1) traktowania skali pomiarowej jako kartograficznego ujęcia danych nieprzestrzennych;
- 2) charakteryzowania atrybutów pod kątem ich funkcji w tworzeniu wizualizacji;
- 3) koncepcji związanej z fizycznym tworzeniem bazy danych i wielofunkcyjnym wykorzystaniem atrybutów bazy danych.

Autor niniejszej rozprawy skłania się do zastosowania ostatniego z wymienionych kryteriów i przyjmuje się za Żyszkowską (2000) określenie „poziomy ujęcia atrybutów”. Poziomy w ujęciu niniejszych badań są zgrupowane zgodnie z podziałem na dane jakościowe i ilościowe, a obejmują pięć poziomów jakościowych: logiczny, nominalny, klasyfikacyjny, porządkowy i hierarchiczny oraz pięć poziomów ilościowych: absolutny, skalarny, interwałowy oraz wskaźnikowy skalarny i wskaźnikowy interwałowy (tab. 5.1.). Podstawą uporządkowania poziomów pomiarowych nie jest dokładność pomiaru (skale „mocniejsze” i „słabsze”, Frączek 1981), lecz następujące cechy:

**Tab. 5.1. Propozycja klasyfikacji poziomów ujęcia atrybutów**  
**Źródło: opracowanie własne**

Poziomy ujęcia atrybutów			Charakterystyka	Dopuszczalne wartości	Przykład
Symbol					
<b>JAKOŚCIOWE</b>	LOGICZNY	<b>JL</b>	Wskazanie obszarów (miejsz) występowania zjawiska/obiektu w przestrzeni geograficznej	1 (Tak, Yes, True), 0 (Nie, No, False)	Ciek okresowo zanikający - wartość Y lub N
	NOMINALNY	<b>JN</b>	Wyróżnienie indywidualnych, niepowtarzalnych cech obiektu (dotyczy głównie nazw geograficznych)	Nazwa własna lub unikalny ID	Nazwa jeziora lub nr katalogowy jeziora
	KLASYFIKACYJNY	<b>JK</b>	Grupowanie obiektów w niehierarchiczne klasy w oparciu o jedną lub wiele cech	Nazwa lub kod klasy obiektów	Nazwa formy pokrycia terenu lub kod Corine Land Cover
	PORZĄDKOWY	<b>JP</b>	Grupowanie obiektów w klasy, pomiędzy którymi zachodzi relacja porządku	Nazwa lub kod klasy obiektów	Klasa bonitacyjna gleb lub liczbowy kod klasy bonitacyjnej gleb
	HIERARCHICZNY	<b>JH</b>	Grupowanie obiektów w klasy, pomiędzy którymi zachodzi relacja hierarchii	Unikalny ID hierarchiczny	TERYT, ID hydrograficzny cieków
<b>ILOŚCIOWE</b>	ABSOLUTNY	<b>IA</b>	Określenie liczby obiektów w sztucznej lub naturalnej jednostce przestrzennej	Count (liczba) – w postaci liczby całkowitej	Liczba drzew pomnikowych w podziale na gminy
	SKALARNY	<b>IS</b>	Ustawienie obiektów w relacji porządku w oparciu o wartość atrybutu liczbowego	Dowolna wartość liczbową	Obwód drzewa
	INTERWAŁOWY	<b>II</b>	Grupowanie obiektów w oparciu o wartości atrybutu liczbowego do ustalonych przedziałów liczbowych i ustawienie ich w relacji porządku	Klasa wartości liczbowych	Klasy szerokości koryta cieków
	WSKAŹNIKOWY SKALARNY	<b>IWS</b>	Ustawienie obiektów w relacji porządku w oparciu o obliczone wskaźniki	Dowolna wartość liczbową wskaźnika	Wartości wskaźnika rozwinięcia linii brzegowej
	WSKAŹNIKOWY INTERWAŁOWY	<b>IWI</b>	Grupowanie obiektów w oparciu o obliczone wskaźniki do ustalonych przedziałów liczbowych i ustawienie ich w relacji porządku	Klasa wartości liczbowych wskaźnika	Klasy wskaźnika rozwinięcia linii brzegowej

- wzrost stopnia porównywalności i możliwości interpretacyjnych danych jakościowych (od wskazania obszarów występowania zjawiska oraz identyfikacji niepowtarzalnych obiektów, poprzez klasyfikację do rankingu oraz tworzenie klas hierarchicznych – por. ryc. 5.1.) oraz danych ilościowych (od liczby obiektów, poprzez porównanie

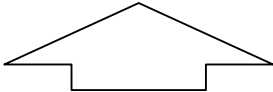
- indywidualnych wartości i klasyfikację z użyciem przedziałów liczbowych do porównywania syntetycznych wskaźników wartości – por. ryc. 5.2.);
- wzrost stopnia zależności między cechami jakościowymi obiektów (proste wskazanie klasy – ranking – zależność hierarchiczna);
  - wzrost stopnia przetworzenia informacji ilościowej (od liczebności obiektów do wskaźników liczbowych).

Z powyższych powodów autor niniejszej rozprawy zdecydował się na przesunięcie w grupie jakościowej poziomu hierarchicznego na ostatnią pozycję.

Próba nazwania procesów generowania atrybutów jakościowych została zawarta na ryc. 5.1. Zależność funkcjonalna cech jakościowych wzrasta od dołu ryciny ku górze. Poziom logiczny odpowiada procesowi detekcji obiektów o określonej cesze. Poziom nominalny jest związany z identyfikacją obiektów poprzez nadanie im nazw lub identyfikatorów. Grupowanie obiektów w oparciu o wybraną cechę stanowi przejście do poziomu klasyfikacyjnego. Przypisanie klasom obiektów miar względnych pozwalających na uszeregowanie klas w odniesieniu do „siły” danego atrybutu można określić mianem waloryzacji i jest to przejście do poziomu porządkowego. Natomiast poziom hierarchiczny wiąże się najczęściej z zastosowaniem wieloatrybutowej klasyfikacji obiektów, w którym występują atrybuty nadrzędne i podrzędne. W typowych zjawiskach o charakterze hierarchicznym (podział administracyjny, sieć hydrograficzna i zlewnie, użytkowanie terenu) dla określenia hierarchii obiektu stosuje się jednolity system identyfikatorów hierarchicznych. Należy zaznaczyć, że opisany schemat (ryc. 5.1.) porządkuje jedynie procesy i metody generowania atrybutów jakościowych. Przykładowe reprezentacje obiektów zaprezentowane na tej rycinie nie muszą, a w niektórych przypadkach nawet nie mogą dotyczyć tych samych obiektów.

Kolejność poziomów ilościowych związana jest głównie ze wzrostem możliwości interpretacyjnych danych liczbowych (ryc. 5.2.). Poziom absolutny informuje jedynie o frekwencji obiektów w danej jednostce przestrzennej. Poziom skalarny charakteryzuje się już przypisaniem konkretnej wartości cechy do każdego obiektu. Zastosowanie poziomu interwałowego pozwala na grupowanie obiektów w klasy scharakteryzowane określonymi przedziałami wartości atrybutu. Najwyższy poziom interpretacji danych ilościowych pozwala na zastosowanie liczbowych wskaźników obliczonych z innych atrybutów ilościowych oraz ich zinterpretowanie z wykorzystaniem wartości wskaźnika w ujęciu skalarnym lub interwałowym.

Poziom ujęcia atrybutów jakościowych	Przykładowe reprezentacje obiektów				Proces generowania atrybutów jakościowych
<b>JH</b> hierarchiczny	<b>RZĄD I</b>	<b>RZĄD II</b>	<b>RZĄD III</b>	...	<b>HIERARCHIZACJA</b>
	ob.1	ob11 ob12 ob13	ob111 ob112 ob113 ob121 ob122 ob123 ob131 ob132 ob133	...	
	ob.2	ob21 ob22 ob23	ob211 ob212 ob213 ob221 ob222 ob223 ob231 ob232 ob233	...	
	ob.3	ob31 ob32 ob33	ob311 ob312 ob313 ob321 ob322 ob323 ob331 ob332 ob333	...	
	...	...	...	...	
<b>JP</b> porządkowy	<b>TYP A</b> bardzo dobry - ob.3 - ob.9	<b>TYP B</b> przeciętny - ob.1 - ob.2 - ob.8	<b>TYP C</b> słaby - ob.4 - ob.5 - ob.6 - ob.7	...	<b>WALORYZACJA</b>
<b>JK</b> klasyfikacyjny	<b>KLASA 1</b> (ob.1, ob.5, ob.6)	<b>KLASA 2</b> (ob.7)	<b>KLASA 3</b> (ob.3, ob.9)	...	<b>KLASYFIKACJA</b>
<b>JN</b> nominalny	ob.1 ob.2 ob.3 ob.4 ob.5 ob.6 ob.7 ob.8 ob.9 ...				<b>IDENTYFIKACJA</b>
<b>JL</b> logiczny	□ □ □ □ □ □ □ □ □ ...				<b>DETEKCJA</b>

  
**R Z E C Z Y W I S T O Ś Ć**

*Ryc. 5.1. Schemat generowania atrybutów jakościowych  
Źródło: opracowanie własne*

Poziom ujęcia atrybutów ilościowych	Przykładowe reprezentacje obiektów			Proces generowania atrybutów ilościowych
<b>II, IWI</b> interwałowy, wskaźnikowy interwałowy	<b>PRZEDZIAŁ 1</b> <0; 10> - ob.7 - ob.2	<b>PRZEDZIAŁ 2</b> (10; 20> - ob.4 - ob.1 - ob.9	<b>PRZEDZIAŁ 3</b> > 20 - ob.6 - ob.3 - ob.8 - ob.5	<b>KLASYFIKACJA</b>
<b>IS, IWS</b> skalarny, wskaźnikowy skalarny				<b>WARTOŚCIOWANIE</b>
<b>IA</b> absolutny				<b>ZLICZANIE</b>



*Ryc. 5.2. Schemat generowania atrybutów ilościowych*

*Źródło: opracowanie własne*

### 5.3. Pomiar potencjału informacyjnego danych przestrzennych - przykłady

Poniżej przedstawiono propozycję wyznaczania potencjału informacyjnego danych przestrzennych w ujęciu zaprezentowanym w rozdziale 5.1. W oparciu o informacje ogólne o bazach danych (rozdział 3) oraz przeprowadzone badania jakości danych (rozdział 4) wyznaczono względne wskaźniki jakości danych zgodnie z normą ISO 19113 i 19114 w zakresie ilościowych wskaźników jakości: kompletności danych, spójności logicznej, dokładności położenia i dokładności czasowej. Przy określaniu dokładności tematycznej wykorzystano poziomy ujęcia atrybutów zaproponowane w rozdziale 5.2. W określaniu potencjału wykorzystano przykłady poddane szczegółowym badaniom jakości danych w rozdziale 4, mianowicie warstwy rzek oraz jezior z trzech baz danych: Mapy Hydrograficznej Polski (MHP), Mapy Podziału Hydrograficznego Polski (MPHP) oraz Mapy Wektorowej Poziomu Drugiego (VML2).

Punktem wyjścia do określenia potencjału informacyjnego warstwy tematycznej było określenie oczekiwanych cech danych, a więc:

1. Pożądanego modelu pojęciowego warstwy;
2. Przyjętych metod określania jakości danych w zakresie: kompletności, spójności logicznej, dokładności położenia i dokładności czasowej;
3. Oczekiwanego zestawu danych atrybutowych.

Wyniki badania poszczególnych elementów składowych potencjału informacyjnego zamieszczono w tabeli zbiorczej. Na każdym etapie określania potencjału zastosowano wskaźniki, gdzie wartość „1” określała stuprocentowe spełnianie danego kryterium. Dzięki temu każdy wynik cząstkowy: kompletność, spójność logiczną, dokładność położenia, czasową i tematyczną można było przedstawić jako wskaźnik. Potencjał informacyjny warstwy tematycznej (PIWT) wyliczano ze wzoru:

$$\text{PIWT} = w1 * K + w2 * SL + w3 * DP + w4 * DC + w5 * DT$$

gdzie:

K – wskaźnik kompletności obiektów;

SL – kompleksowy wskaźnik spójności logicznej;

DP – wskaźnik dokładności przestrzennej;

DC – wskaźnik dokładności czasowej;

DT – kompleksowy wskaźnik dokładności tematycznej;

w1...w5 – wagi poszczególnych wskaźników.

Wagi poszczególnych wskaźników są dobierane eksperymentalnie przez użytkownika danych. To podkreśla jeszcze bardziej podstawową cechę określania potencjału – jego subiektywizm oparty na potrzebach użytkownika. Suma wszystkich wag powinna wynosić 1.

Ostatecznym wynikiem pomiaru potencjału informacyjnego dla grupy warstw tematycznych będzie uśredniona wartość PIWT wszystkich warstw rozpatrywanej bazy danych, tworzących tę grupę warstw. Każdej z warstw należy przypisać wagę w zależności od roli, jaką pełni w grupie warstw. Wagi należy dobierać indywidualnie, kierując się takimi cechami, jak: ilość obiektów odwzorowywanych na warstwie względem innych warstw, ważność tych obiektów, bogactwo informacji atrybutowej oraz zależności semantyczne między warstwami. Wartość potencjału informacyjnego grupy warstw tematycznych (PIGWT) ustala się ze wzoru:

$$\text{PIGWT} = w1 * \text{PIWT1} + w2 * \text{PIWT2} + \dots + wn * \text{PIWTn}$$

gdzie:

w1, w2, ... wn – wagi poszczególnych warstw tematycznych;

PIWT1, PIWT2, ..., PIWTn – wartości potencjału informacyjnego poszczególnych warstw tematycznych.

### 5.3.1. Pomiar potencjału informacyjnego dla wód płynących

*Pożądany model pojęciowy:* dane dotyczące rzek powinny być zgromadzone w warstwach tematycznych obejmujących: 1) rzeki nazwane (obiekty liniowe) – dla rzek o szerokości poniżej 30m; 2) rzeki nazwane (obiekty powierzchniowe) – dla rzek o szerokości 30m i więcej; 3) cieki pozostałe (obiekty liniowe) – cieki bez nazwy, istotne rowy o charakterze melioracyjnym; 4) kanały (obiekty liniowe) – dla obiektów o szerokości poniżej 30m; 5) kanały (obiekty powierzchniowe) – dla obiektów o szerokości 30m i więcej. Obiekty powinny być utworzone zgodnie z regułami topologicznymi (połączenia na węzłach; zapis łańcuchów zgodny z nurtem; obecność łączników sieciowych w postaci łańcuchów na odcinkach powierzchniowych – jeziora przepływowe i rzeki przedstawione jako obiekty powierzchniowe).

#### *Metoda określania jakości danych*

Kompletność – jest określana jako niedomiar względem bazy zawierającej najliczniejszą populację obiektów (w przypadku rzek miarą jest łączna długość cieków).

#### Spójność logiczna:

- w zakresie spójności pojęciowej: przyjmuje się za wzorcowe następujące kategorie obiektów: obiekty liniowe – cieki główne, cieki pozostałe, rowy melioracyjne, kanały; obiekty powierzchniowe – rzeki o szerokości powyżej 30m;
- spójność dziedziny: przyjęto poprawną definicję;
- spójność formatu: przyjęto poprawną definicję;
- spójność topologiczna: założono obligatoryjność:
  - \* połączeń obiektów w węzłach,
  - \* zgodności kierunku digitalizacji z nurtem,
  - \* obecności połączeń liniowych w obrębie obiektów powierzchniowych rzek.

Dokładność położenia: określana w sześciu polach testowych o rozmiarach arkusza mapy topograficznej 1:10000 w układzie 1992; materiałem referencyjnym jest sieć hydrograficzna widoczna na analogowej mapie topograficznej 1:10000 w układzie współrzędnych 1992.

Dokładność czasowa - przyjęto poprawną definicję.

*Oczekiwany zestaw danych atrybutowych (dokładność tematyczna):*

- 1) identyfikator hydrograficzny (JN);
- 2) nazwa cieku (JN);
- 3) numer nadrzędnego identyfikatora hydrograficznego (recypienta) (JN);

- 4) rząd cieków (JH);
- 5) szerokość cieków (II);
- 6) charakter przepływu (stały lub okresowy), a dla kanałów funkcja (JK);

(w nawiasach podano przyporządkowanie atrybutu do poziomu ujęcia atrybutów zgodnie z tab. 5.1. ).

Zestawienie wyników cząstkowych wartości potencjału informacyjnego warstw prezentuje tab. 5.2.

**Tabela 5.2. Zestawienie cząstkowych wartości potencjału informacyjnego dla warstw rzek baz danych MHP, MPHP i VML2**  
**Źródło: opracowanie własne**

Elementy jakości danych		MHP			MPHP		VML2		
KOMPLETNOŚĆ (K)		574,08 / 574,08			523,60 / 574,08		477,75 / 574,08		
		<b>1,0</b>			<b>0,91</b>		<b>0,83</b>		
SPÓJNOŚĆ LOGICZNA (SL)	POJĘCIOWA	1,0			0,8		1,0		
	- cieków główne	0,2			0,2		0,2		
	- cieków pozostałe	0,2			0,2		0,2		
	- rowy meliorac.	0,2			0,0		0,2		
	- kanały	0,2			0,2		0,2		
	- obszary wodne	0,2			0,2		0,2		
	DZIEDZINY	1,0			1,0		1,0		
	FORMATU	1,0			1,0		1,0		
	TOPOLOGICZNA	1,0			1,0		0,33		
	- połączeń w węzłach	0,33			0,33		0,33		
	- popr. kier. digit.	0,33			0,33		0,0		
- łączniki liniowe	0,33			0,33		0,0			
ŁĄCZNIE		4/4 = <b>1,0</b>			3,8/4 = <b>0,95</b>		3,33/4 = <b>0,83</b>		
DOKŁADNOŚĆ POŁOŻENIA (DP)		<b>0,93</b>			<b>0,67</b>		<b>0,69</b>		
DOKŁADNOŚĆ CZASOWA (DC)		<b>1,0</b>			<b>1,0</b>		<b>1,0</b>		
DOKŁ. TEMATYCZNA (DT)	DANE ATRYBUTOWE	5/6=	1/6=	5/6=	5/6=	2/6=	3/6=	3/6=	3/6=
		0,83	0,17	0,83	0,83	0,33	0,50	0,50	0,50
	- ID hyd.	1/6	0	1/6	1/6	0	0	0	0
	- nazwa cieków	1/6	0	1/6	1/6	0	1/6	1/6	1/6
	- Id hyd recipienta	1/6	0	1/6	1/6	0	0	0	0
	- rząd cieków	0	0	0	1/6	0	0	0	0
	- szer. cieków	1/6	0	1/6	0	1/6	1/6	1/6	1/6
- char. przepływu	1/6	1	1/6	1/6	1/6	1/6	1/6	1/6	
Warstwy tematyczne		Cieki_n, Pow_wod	Cieki_bn	Kanal	Rzeki_r, S_rzeki	Rzekin	Rivera, Riverl	Ditchl	Canala, Canall
		MHP			MPHP		VML2		

Przy określaniu ostatecznej wartości wskaźnika potencjał informacyjny warstwy tematycznej (PIWT) uwzględniono powyższe miary (tab. 5.2.) z następującymi wagami:

$$PIWT = 0,2 * K + 0,15 * SL + 0,2 * DP + 0,05 * DC + 0,4 * DT$$

Zatem za najważniejszą miarę wpływającą na potencjał informacyjny uznano zasobność danej warstwy w atrybuty oczekiwane z punktu widzenia zastosowania bazy (40%). Obliczenie potencjału informacyjnego dla warstw tematycznych rzek zestawiono w tab. 5.3.

**Tabela 5.3. Potencjał informacyjny warstw tematycznych rzek dla baz danych MHP, MPHP i VML2**  
Źródło: opracowanie własne

Baza	warstwa	0,2 * K	0,15 * SL	0,2 * DP	0,05 * DC	0,4 * DT	PIWT
<b>MHP</b>	Cieki_n*					0,33	<b>0,92</b>
	Cieki_bn	0,20	0,15	0,19	0,05	0,07	<b>0,66</b>
	Kanal					0,33	<b>0,92</b>
<b>MPHP</b>	Rzeki_r*	0,18	0,14	0,13	0,05	0,33	<b>0,83</b>
	Rzekin					0,13	<b>0,63</b>
<b>VML2</b>	Rivera*					0,20	<b>0,68</b>
	Ditchl,	0,17	0,12	0,14	0,05	0,20	<b>0,68</b>
	Canala*					0,20	<b>0,68</b>

\*Zarówno warstwa liniowa, jak i powierzchniowa

Aby określić potencjał informacyjny grupy warstw „rzeki”, należy odnieść się do zakresów semantycznych poszczególnych warstw (por. rozdział 4, ryc. 4.1.).

W zależności od miejsca danej warstwy tym schemacie, przypisano poszczególnym warstwom wskaźniki wagowe wynikające z ich roli w prezentacji rzek i ich atrybutów:

**MHP:** cieki\_n=0,7, kanal=0,1, cieki\_bn=0,2

**MPHP:** rzeki\_r=0,8, rzekin=0,2

**VML2:** Rivera, riverl=0,85, canala, canal=0,07, ditcha, ditchl=0,08

Potencjał informacyjny grupy warstw tematycznych (PIGWT) przedstawia się zatem następująco:

$$PIGWT\_MHP = 0,7*0,92 + 0,1*0,66 + 0,2*0,92 = 0,644 + 0,066 + 0,184 = \mathbf{0,894}$$

$$PIGWT\_MPHP = 0,8*0,83 + 0,2*0,63 = 0,664 + 0,126 = \mathbf{0,790}$$

$$PIGWT\_VML2 = 0,85*0,68 + 0,07*0,68 + 0,08*0,68 = 0,578 + 0,0476 + 0,0544 = \mathbf{0,680}$$

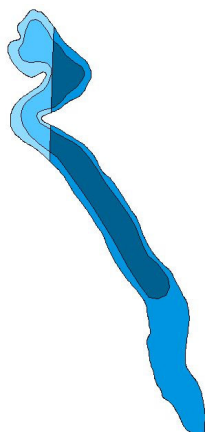
Podsumowując, warstwy rzek według założeń użytkownika spełnią jego oczekiwania w niemal 90% w przypadku bazy MHP, niecałych 80% w przypadku MPHP, oraz w 68% dla bazy VML2.

### 5.3.2. Pomiar potencjału informacyjnego dla wód stojących

*Pożądany model pojęciowy:* dane dotyczące jezior powinny być zgromadzone w warstwach tematycznych obejmujących: 1) zbiorniki wodne o powierzchni powyżej 1 250 m<sup>2</sup>; 2) zbiorniki retencyjne o powierzchni powyżej 1 250 m<sup>2</sup>; 3) stawy hodowlane o powierzchni powyżej 1250 m<sup>2</sup>.

*Metoda określania jakości danych*

Kompletność – jest określana jako niedomiar względem bazy zawierającej najliczniejszą populację obiektów (w przypadku jezior miarą jest łączna powierzchnia wszystkich obiektów stojących wód powierzchniowych o powierzchni powyżej 1 250 m<sup>2</sup>). W przypadku VML2 zaistniał problem ustalania kompletności obiektów ze względu na specyficzny zapis obiektów jezior (ryc. 5.3.). Każdy obszar między hydroizobatami stanowi odrębny obiekt, ponadto zaznacza się podział arkuszowy bazy.



Ryc. 5.3. Brak integralności obiektu jeziora w VML2 na przykładzie Jeziora Łódzko-Dymaczewskiego

#### Spójność logiczna:

- w zakresie spójności pojęciowej: przyjmuje się za wzorcowe następujące kategorie obiektów: zbiorniki wodne, zbiorniki retencyjne i stawy hodowlane;
- spójność dziedziny: przyjęto poprawną definicję;
- spójność formatu: przyjęto poprawną definicję;
- spójność topologiczna: założono obligatoryjność:
  - \* połączeń obiektów w węzłach,
  - \* zgodności kierunku digitalizacji z nurtem,
  - \* obecności połączeń liniowych w obrębie obiektów powierzchniowych rzek.

Dokładność położenia: określana w sześciu polach testowych o rozmiarach arkusza mapy topograficznej 1:10 000 w układzie współrzędnych 1992; materiałem referencyjnym jest sieć hydrograficzna widoczna na analogowej mapie topograficznej 1:10 000 w układzie współrzędnych 1992.

Dokładność czasowa - przyjęto poprawną definicję.

*Oczekiwany zestaw danych atrybutowych (dokładność tematyczna):*

- 1) identyfikator katalogowy jeziora (JN);
- 2) nazwa katalogowa jeziora (JN);
- 3) rzędna zwierciadła wody (IS);
- 4) szerokość (IS);
- 5) długość (IS);
- 6) powierzchnia (IS);
- 7) głębokość maksymalna (IS);
- 8) głębokość średnia (IS);
- 9) długość linii brzegowej (IS);
- 10) pojemność (IS);
- 11) funkcja (JK).

Zestawienie wyników cząstkowych wartości potencjału informacyjnego warstw prezentuje tab. 5.4.

Przy określaniu ostatecznej wartości wskaźnika potencjał informacyjny warstwy tematycznej (PIWT) uwzględniono powyższe miary z następującymi wagami:

$$\text{PIWT} = 0,2 * K + 0,15 * SL + 0,2 * DP + 0,05 * DC + 0,4 * DT$$

Zatem za najważniejszą miarę wpływającą na potencjał informacyjny uznano zasobność danej warstwy w atrybuty oczekiwane z punktu widzenia zastosowania bazy (40%). Obliczenie potencjału informacyjnego dla warstw tematycznych jezior zestawiono w tab. 5.5.

**Tabela 5.4. Zestawienie cząstkowych wartości potencjału informacyjnego dla warstw jezior baz danych MHP, MPHP i VML2**

*Źródło: opracowanie własne*

Elementy jakości danych		MHP	MPHP		VML2		
KOMPLETNOŚĆ (K)		393,218 / 393,218	385,382 / 393,218		373,564 / 393,218		
		<b>1,0</b>	<b>0,98</b>		<b>0,95</b>		
SPÓJNOŚĆ LOGICZNA (SL)	POJĘCIOWA	1,0	0,66		1,0		
	- zbiorniki wodne	0,33	0,33		0,33		
	- zbiorniki retenc.	0,33 <sup>*)</sup>	0,165 <sup>**)</sup>		0,33		
	- stawy hodowl.	0,33 <sup>*)</sup>	0,165 <sup>**)</sup>		0,33		
	DZIEDZINY	1,0	1,0		1,0		
	FORMATU	1,0	1,0		1,0		
	TOPOLOGICZNA	0,8	0,8		0,2		
	- integraln. wewn.	0,8	0,8		0,0		
	- wyspy	0,0	0,0		0,2		
ŁĄCZNIE		3,8/4 = <b>0,95</b>	3,46/4 = <b>0,87</b>		3,2/4 = <b>0,80</b>		
DOKŁADNOŚĆ POŁOŻENIA (DP)		<b>1,0</b>	<b>0,98</b>		<b>0,97</b>		
- wskaźnik RLB		1,0	0,99		0,99		
- powierzchnia		1,0	0,96		0,95		
DOKŁADNOŚĆ CZASOWA (DC)		<b>1,0</b>	<b>1,0</b>		<b>1,0</b>		
DOKŁ. TEMATYCZNA (DT)	DANE ATRYBUTOWE	0,5	0,2	0,0	0,9	0,35	0,45
	- ID katalogowe	0,1	0,1	0,0	0,0	0,0	0,0
	- nazwa katalog.	0,1	0,1	0,0	0,1	0,1	0,1
	- rzędna zw. wody	0,1	0,0	0,0	0,1	0,1	0,1
	- szerokość	0,0	0,0	0,0	0,1	0,0	0,0
	- długość	0,0	0,0	0,0	0,1	0,0	0,0
	- powierzchnia	0,05 <sup>***)</sup>	0,0	0,0	0,05	0,05	0,05
	- głęb. maksym.	0,05	0,0	0,0	0,1	0,0	0,0
	- głęb. średnia	0,0	0,0	0,0	0,1	0,0	0,0
	- dł. linii brzeg.	0,0	0,0	0,0	0,05	0,0	0,0
	- pojemność	0,0	0,0	0,0	0,1	0,0	0,1
- funkcja	0,1	0,0	0,0	0,1 <sup>****)</sup>	0,1 <sup>****)</sup>	0,1	
Warstwy tematyczne	Zb_wod		jeziora	jez_n	lakea	fishfirma	reservra
		MHP	MPHP		VML2		

<sup>\*)</sup> identyfikowane przez określenie funkcji (atrybut)

<sup>\*\*)</sup> obecność obiektów, ale brak identyfikacji funkcji

<sup>\*\*\*)</sup> informacja występuje na odrębnej warstwie

<sup>\*\*\*\*)</sup> brak określenia funkcji, ale jest rozwarstwienie obiektów na zbiorniki wodne i stawy hodowlane

**Tabela 5.5. Potencjał informacyjny warstw tematycznych jezior dla baz danych MHP, MPHP i VML2**

**Źródło: opracowanie własne**

Baza	warstwa	0,2 * K	0,15 * SL	0,2 * DP	0,05 * DC	0,4 * DT	PIWT
MHP	zb_w	0,20	0,14	0,2	0,05	0,20	<b>0,79</b>
MPHP	jeziora	0,20	0,13	0,2	0,05	0,08	<b>0,66</b>
	jez_n					0,00	<b>0,58</b>
VML2	lakea	0,19	0,12	0,19	0,05	0,36	<b>0,91</b>
	fishfrma					0,14	<b>0,69</b>
	reservra					0,18	<b>0,73</b>

W zależności od miejsca danej warstwy w reprezentowaniu jezior, przypisano poszczególnym warstwom wskaźniki wagowe wynikające z ich roli w prezentacji jezior i ich atrybutów:

**MHP:** zb\_w=1,0

**MPHP:** jeziora=0,8, jez\_n=0,2

**VML2:** lakea=0,85, fishfrma=0,07, reservra=0,08

Potencjał informacyjny grupy warstw tematycznych (PIGWT) wynosi:

$$\text{PIGWT\_MHP} = 1,0 * 0,79 = \mathbf{0,790}$$

$$\text{PIGWT\_MPHP} = 0,8 * 0,66 + 0,2 * 0,58 = 0,528 + 0,116 = \mathbf{0,644}$$

$$\text{PIGWT\_VML2} = 0,85 * 0,91 + 0,07 * 0,69 + 0,08 * 0,73 = 0,7735 + 0,0483 + 0,0584 = \mathbf{0,8802}$$

Można podsumować, że warstwy jezior według założeń użytkownika spełnią jego oczekiwania w niemal 80% w przypadku bazy MHP, około 65% w przypadku MPHP, oraz w 88% przy wykorzystaniu bazy VML2.

**Tak uzyskane miary potencjału informacyjnego należałoby traktować jako wskaźnik stopnia dopasowania danej warstwy (PIWT) lub grupy warstw tematycznych (PIGWT) do oczekiwanej przez użytkownika struktury tematycznej oraz innych parametrów jakości danych. Ze względu na indywidualność procedury otrzymane wyniki mogą być bezpośrednio porównywalne wyłącznie w ramach badanej próby.**

#### 5.4. Dynamiczny potencjał informacyjny

Nowoczesne metody geoinformacyjnego przetwarzania danych oraz postępujący proces doskonalenia baz danych sprzyjają pragmatycznemu podejściu do wykorzystania potencjału informacyjnego baz danych w badaniach środowiska przyrodniczego. W tym rozdziale zaprezentowane zostanie dynamiczne ujęcie zagadnienia potencjału informacyjnego. Jako materiał ilustracyjny posłużą najnowsze dane z zasobu geodezyjno-kartograficznego.

Badania potencjału informacyjnego baz danych przestrzennych w ujęciu prezentowanym w poprzednim rozdziale są podyktowane potrzebą optymalizacji wyboru źródła danych w sytuacji ich redundancji oraz niepewności (Goodchild 2000). Konieczne jest wtedy wyznaczanie wskaźników jakości danych (rozdział 4) i uwzględnianie ich przy obliczaniu potencjału informacyjnego warstw tematycznych (PIWT) łącznie z miarą dokładności tematycznej (DT; rozdział 5.3.).

Najnowsze trendy rozwojowe zasobu geodezyjno-kartograficznego pozwalają spojrzeć na problematykę potencjału informacyjnego w sposób odmienny od zaprezentowanego w poprzednim rozdziale. Przy użytkowaniu danych przestrzennych, w wybranych klasach obiektów możliwe jest już pominięcie żmudnego etapu oceny jakości danych lub przynajmniej ograniczenie go do wybiórczej kontroli. Sytuacja taka wynika w szczególności z postępu dwóch procesów: sukcesywnego przetwarzania danych geodezyjnych do postaci cyfrowej oraz tworzenia spójnej topologicznie bazy topograficznych danych referencyjnych. Modelowanie wyżej wymienionych danych jest oparte na solidnych podstawach formalnych (Instrukcja techniczna K-1 i Wytyczne techniczne TBD), rzetelnych danych źródłowych (dane z pomiarów geodezyjnych, rejestry państwowe, ortofotomapa, weryfikacja terenowa) oraz zaawansowanych metodach kontroli zasobu (Bielecka 2010).

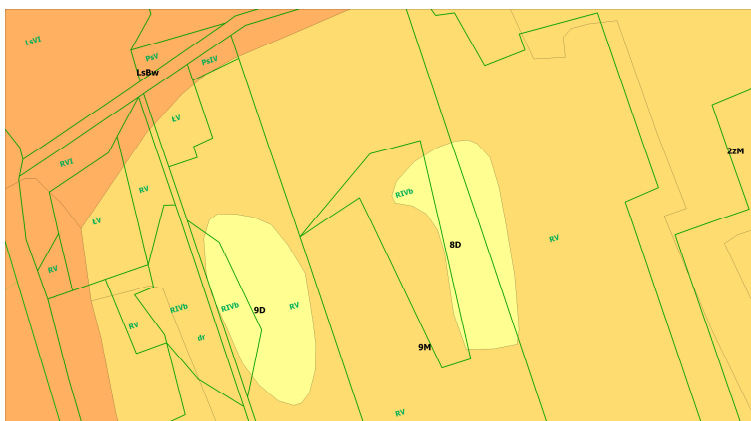
Wysoką jakość danych przestrzennych o których mowa powyżej potwierdza wstępna ocena jakości danych, przeprowadzona metodą zaprezentowaną w rozdziale 4 (tab. 5.6.). Dane w większości pochodzą z bazy danych obiektów topograficznych opracowywanych w oparciu o wytyczne techniczne TBD w części na podstawie cyfrowych danych ewidencji gruntów i budynków. Ponadto w zestawieniu tym znalazły się dane tematyczne oparte na pomiarach i obliczeniach geodezyjnych oraz urzędowej taksacji (klasyfikacji): dane Leśnej Mapy Numerycznej z bazą SILP i klasoużytki z Ewidencji Gruntów i Budynków. Uzupełnieniem danych hydrograficznych są zasięgi zlewni elementarnych z Mapy Podziału

Hydrograficznego Polski, która jest bazą referencyjną w hydrografii. Informacja o glebach została przyjęta z opracowań przygotowanych w oparciu o szczegółowe badania terenowe.

Warto podkreślić, że badany zestaw danych charakteryzuje się dużą wzajemną zgodnością topologii, co stwarza warunki porównywalności tych baz oraz możliwości wspólnego występowania ich elementów w analizach przestrzennych i wizualizacjach. Istotne konflikty topologiczne dotyczą w zasadzie tylko niezgodności przestrzennych między wydzieleniami mapy glebowo-rolniczej i klasami bonitacyjnymi gleb zawartymi na warstwie klasoużytków (ryc. 5.4.).

**Tabela 5.6. Wstępna ocena jakości dla wybranych grup tematycznych współczesnych baz danych przestrzennych**  
**Źródło: opracowanie własne**

GRUPA TEMATYCZNA	BAZA ŹRÓDŁOWA	WARSTWA TEMATYCZNA	SYMBOL WARSTWY	OCENA JAKOŚCI	
Wody powierzchniowe	GBDOT	Rzeki i kanały	SWRK	3/3/2	<b>8</b>
		Rowy melioracyjne	SWML	3/3/2	<b>8</b>
		Obszary wód płynących i stojących	PKWO	3/3/2	<b>8</b>
		Mokradła	OIMO	3/3/2	<b>8</b>
	MPHP	Zlewnie elementarne	zlew_el	3/2/3	<b>8</b>
Gleby	MGR	Typy gleb	gleby	3/3/3	<b>9</b>
	EGiB	Klasy bonitacyjne gleb	GESKLU	3/3/3	<b>9</b>
Obszary leśne	LMN +SILP	Pododdziały leśne + dane SILP	wydz_pol	3/3/3	<b>9</b>
		Lasy prywatne	syt_pol	3/3/2	<b>8</b>
Obszary zabudowane	GBDOT	Budynki	BBBD	3/3/3	<b>9</b>
		Punkty adresowe	ARAD	2/3/3	<b>8</b>
Obiekty komunikacyjne	GBDOT	Drogi	SKJZ	3/3/2	<b>8</b>
		Linie kolejowe	SKKL	3/3/3	<b>9</b>
		Przystanki	OIKM	3/3/2	<b>8</b>
		Obszary komunikacyjne	KUKO	3/3/3	<b>9</b>
Jednostki administracyjne	GBDOT	Gminy	ADGM	3/3/3	<b>9</b>
		Obłęby geodezyjne	ADOE	3/3/3	<b>9</b>
		Miejscowości	ADMS	3/3/3	<b>9</b>



**Ryc. 5.4. Brak zgodności topologicznej między warstwą klasoużytków EGiB (zielone linie) a wydzieleniami kompleksów glebowo-rolniczych MGR (wieloboki z wypełnieniem)**  
**Źródło: opracowanie własne**

Należy jednak pamiętać, że proces modernizacji danych przestrzennych nie jest zakończony, a powstałe we wcześniejszym okresie bazy danych (MHP, MSP, VML2 i inne) nadal stanowią bogate źródło informacji. W związku z tym trzeba traktować zaproponowaną w rozdziale 5.3. metodę obliczania potencjału informacyjnego jako nadal przydatną w użytkowaniu tych baz danych.

Poziom potencjału informacyjnego baz danych przestrzennych jest w głównej mierze uwarunkowany jakością atrybutów, co zaprezentowano w rozdziale 5.2. i 5.3. Atrybuty wg tych rozważań pełnią funkcję akumulacyjną, analityczną i wizualizacyjną. Ocena potencjału, zaprezentowana na przykładach w rozdziale 5.3. jest jednak statycznym podejściem do danych – wskazuje, na ile spełniają one oczekiwania użytkownika przy założeniu niezmienności danych oraz autonomicznego traktowania każdej warstwy. W rzeczywistości obecnie dane przestrzenne wykorzystuje się w systemach geoinformacyjnych w sposób aktywny, pozwalający na ich przetwarzanie, często z wykorzystaniem wielu warstw tematycznych. Potencjał informacyjny powinien uwzględniać zatem możliwość pełnienia przez dane funkcji akumulacji danych, ich analizy i wizualizacji w ujęciu dynamicznym.

**Funkcja akumulacyjna** dynamicznego potencjału informacyjnego jest opisana przez sumę atrybutów warstwy tematycznej (ryc. 5.5.), rozpatrywanych w poziomach ujęcia atrybutów (tab. 5.7.). Sama funkcja akumulacyjna informuje jedynie opisowo o ilości i rodzaju występujących atrybutów i sama w sobie nie stanowi wystarczającej informacji o potencjale informacyjnym. Jest jednak podstawą ustalenia funkcji analitycznej i wizualizacyjnej.

X UWAGI	INFORM DOD	FUNKCJA OG	FUNKCJA SZ	HAZWA	L KONDIYGHA
302103_4.0001.596/4;3_BUD		g	Gw		1
302103_5.0005.2_BUD		g	Gp		1
302103_5.0005.3_BUD		s	Sm		1
302103_5.0005.4_BUD		g	Gp		2
302103_5.0005.6_BUD		s	Sm		2

A

rtt	rtt description	tuc	tuc description	use	use description	wd1	wtc	wtc description	wd2	mcc	mcc description	rpc	rpc description
16	Limited Access	4	Road	4	National	0	1	All Weather	0	0	Unknown	1	Passing
16	Limited Access	4	Road	4	National	0	1	All Weather	0	0	Unknown	1	Passing
16	Limited Access	4	Road	4	National	0	1	All Weather	0	0	Unknown	1	Passing
0	Unknown	4	Road	114	Communal	0	1	All Weather	0	0	Unknown	1	Passing

B

ADR LES	ADR ADM	TSL	RODZ POW	GP IGL UDZ	GP LISC UD	GP WIEK	GP WYS	GP ZWARCIE
09-22-1-01-44 -a -00	30-21-125-0007	BMSW	D-STAN	10	0	43	20	UM
09-22-1-01-44 -b -00	30-21-125-0007	OL	D-STAN	0	10	49	20	UM
09-22-1-01-44 -f -00	30-21-125-0007	BMSW	D-STAN	10	0	43	19	UM
09-22-1-01-45 -a -00	30-21-125-0007	LMW	D-STAN	0	10	48	20	PRZ
09-22-1-01-45 -b -00	30-21-125-0007	LMW	D-STAN	10	0	49	20	UM
09-22-1-01-45 -c -00	30-21-125-0007	LMW	D-STAN	0	10	49	19	UM

C

Ryc. 5.5. Przykłady akumulacji wybranych atrybutów w źródłowych bazach danych

A – warstwa budynków BDOT, B - warstwa dróg VML2,

C – pododdziały LMN z przyłączoną bazą SILP

Źródło: opracowanie własne

Tabela 5.7. Przykład opisu funkcji akumulacyjnej dla dynamicznego potencjału informacyjnego dla wód powierzchniowych (zgodnie z tab. 5.1.)

Źródło: opracowanie własne

Nazwa warstwy	Symbol warstwy	Nazwa atrybutu	Obligatoryjność (O) lub fakultatywność (F) warstwy	Uwagi i/lub słownik wartości atrybutu	Poziom ujęcia atrybutu
Rzeki i kanały	SWRK	RODZAJ	O	Rzeka (Rz), kanał (Kn), strumień lub potok (St)	JK
		PRZEBIEG	O	Cieki główne (1) lub ramię boczne (2)	JL
		OKRESOWOSC	O	Ciek stały (0) lub okresowy (1)	JL
		ID-HYDRO*)	F	Zgodny z MPHP	JH
		NAZWA*)	F	Nazwa geograficzna	JN
		INFORM_DOD	F	Alternatywna nazwa	JN
Rowy melioracyjne	SWML	RODZAJ	O	Rów melioracyjny zwykły (Rm) lub zbiorczy (Rb)	JL
		SZEROKOSC	F	Szerokość w metrach	IS
		OKRESOWOSC	O	Rów stały (0) lub okresowy (1)	JL
		INFORM_DOD	F	Opis rowu	JN
Obszary wód płynących i stojących	PKWO	RODZAJ	O	Wody morskie (Pm), powierzchniowe płynące (Pp) lub powierzchniowe stojące (Ps)	JK
		ID-HYDRO*)	F	Zgodny z MPHP	JH
		NAZWA*)	F	Nazwa geograficzna	JN
Mokradła	OIMO	RODZAJ	O	Teren podmokły (Mo) lub bagno (Ms)	JL
Zlewnie elementarne	zlew_el	ID_HYD	O	Identyfikator hydrograficzny	JH
		POZIOM	O	Poziom podziału	JH
		NAZ_ZLEWNI	O	Nazwa zlewni	JN

**Funkcja analityczna** dynamicznego potencjału informacyjnego opisuje możliwości przetwarzania danych atrybutowych do celów analitycznych. Przetwarzanie danych ograniczono do części atrybutowej danych z dwóch powodów:

1. Koncepcja potencjału informacyjnego danych przestrzennych została w udokumentowany sposób sprowadzona w niniejszej pracy do części atrybutowej warstw tematycznych
2. Wszelkie zmiany w obrębie atrybutów mają charakter ubogacający lub zubażający zawartość informacyjną o obiektach przestrzennych bez zmiany liczby, kształtów i lokalizacji tych obiektów. Zmiany geometrii obiektów przestrzennych natomiast zmieniają założenia dotyczące modelowania danych przestrzennych i tworzą nową jakość danych, nieporównywalną ze źródłem.

Zakres operacji geoinformacyjnych na danych atrybutowych odpowiadających powyższym założeniom przedstawiono w tab. 5.8 oraz tab. 5.9.

**Tabela 5.8. Zestawienie operacji na atrybutach właściwych do określania funkcji analitycznej dynamicznego potencjału informacyjnego**  
*Źródło: opracowanie własne na podstawie Kaczmarek 2010*

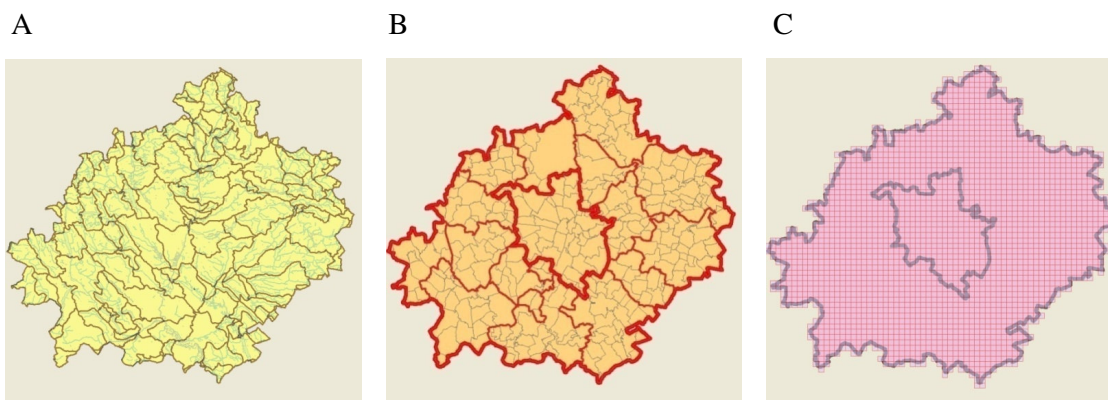
Nazwa operacji na danych atrybutowych		skrót	
Generowanie danych atrybutowych	Przekształcanie atrybutów	arytmetyczne	<b>PAa</b>
		tekstowe	<b>PA<sub>t</sub></b>
		słownikowe	<b>PA<sub>s</sub></b>
	Obliczenia atrybutów z geometrii obiektów		<b>GA</b>
	Przyłączanie atrybutów ze źródła zewnętrznego	z tabeli (po identyfikatorze)	<b>AZ<sub>t</sub></b>
		z innej warstwy (po identyfikatorze)	<b>AZ<sub>w</sub></b>
Obliczenia statystyk atrybutów	dla całego zbioru obiektów warstwy	<b>ST<sub>w</sub></b>	
Tworzenie słowników atrybutowych		<b>SA</b>	
Klasyfikacja obiektów	Jednoatrybutowa	<b>KO<sub>j</sub></b>	
	wieloastrybutowa	<b>KO<sub>w</sub></b>	

**Tabela 5.9. Przykład opisu funkcji analitycznej dla dynamicznego potencjału informacyjnego dla wód powierzchniowych (zgodnie z tab. 5.8.)**

*Źródło: opracowanie własne*

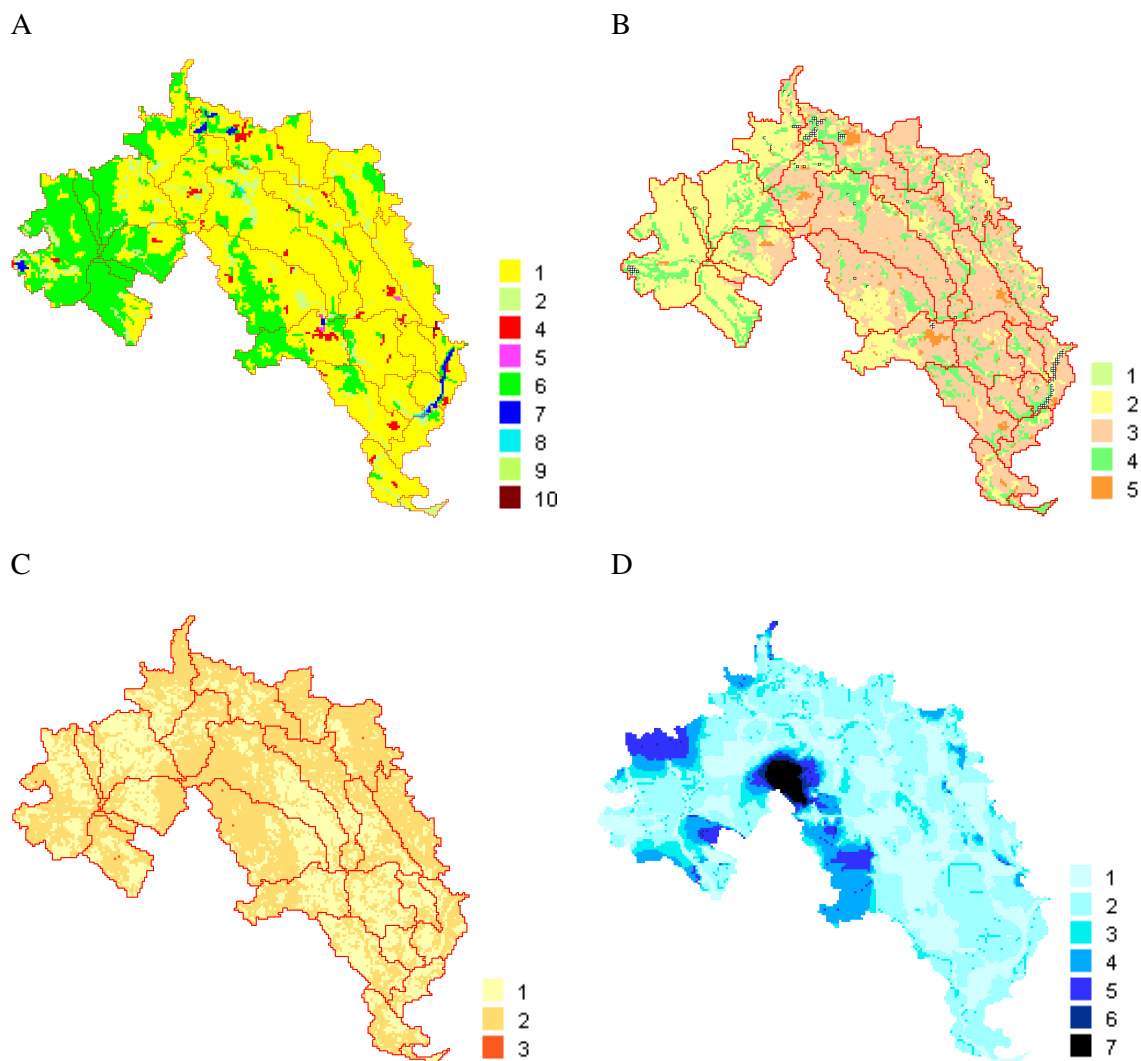
WARSTWA	ATRYBUT	POZ.	PAa	PA <sub>t</sub>	PA <sub>s</sub>	GA	AZ <sub>t</sub>	AZ <sub>w</sub>	ST <sub>w</sub>	SA	KO <sub>j</sub>	KO <sub>w</sub>
SWRK	RODZAJ	JK				+			+	+	+	+
	PRZEBIEG	JL							+	+	+	
	OKRESOWOSC	JL							+	+	+	
	ID-HYDRO*)	JH					+	+			+	
	NAZWA*)	JN		+								
	INFORM_DOD	JN		+								
SWML	RODZAJ	JL				+			+	+	+	+
	SZEROKOSC	IS							+		+	
	OKRESOWOSC	JL							+	+	+	
	INFORM_DOD	JN										
PKWO	RODZAJ	JK				+			+	+	+	
	ID-HYDRO*)	JH										
	NAZWA*)	JN										
OIMO	RODZAJ	JL				+			+	+	+	
zlew_el	ID_HYD	JH				+	+	+			+	
	POZIOM	JH									+	
	NAZ_ZLEWNI	JN		+								

Warto zauważyć, że generowanie danych atrybutowych (tab. 5.8.) może wiązać się ze zmianą jednostki przestrzennej. Wynik obliczania atrybutów z geometrii obiektów lub obliczania statystyk atrybutów jest odnoszony wtedy do innego układu przestrzennego (ryc. 5.6.), a dane są generowane poprzez ich sumowanie, uśrednianie, wybór wartości reprezentatywnej lub wskaźniki (ryc. 5.7.). Taka zmiana układu odniesienia jest metodą na uspójnianie wybranych danych pochodzących z niespójnych topologicznie baz danych (ryc. 5.8.).



**Ryc. 5.6. Przykład jednostek przestrzennych generowania danych atrybutowych: A – naturalne (zlewnie elementarne), B – sztuczne nieregularne (jednostki administracyjne), C – sztuczne regularne (siatka kwadratów)**

*Źródło: opracowanie własne*



**Ryc. 5.7. Przykładowe cechy środowiska przyrodniczego wygenerowane w regularnej siatce kwadratów z bazy CLC (A), MHP (B), DTED2 (C) i MHgP (D)**

**A. Formy użytkowania terenu (CLC): 1 – grunty orne, 2 – łąki i pastwiska, 4 – zabudowa rozproszona, 5 - obszary przemysłowe i komunikacyjne, 6 – lasy, 7 – obszary wodne, 8 – obszary podmokłe, 9 – inne formy roślinności, 10 – obszary eksploatacji odkrywkowej;**

**B. Przepuszczalność gruntu (MHP): 1 – łatwa, 2 – średnia, 3 – słaba, 4 – zmienna, 5 – zróżnicowana;**

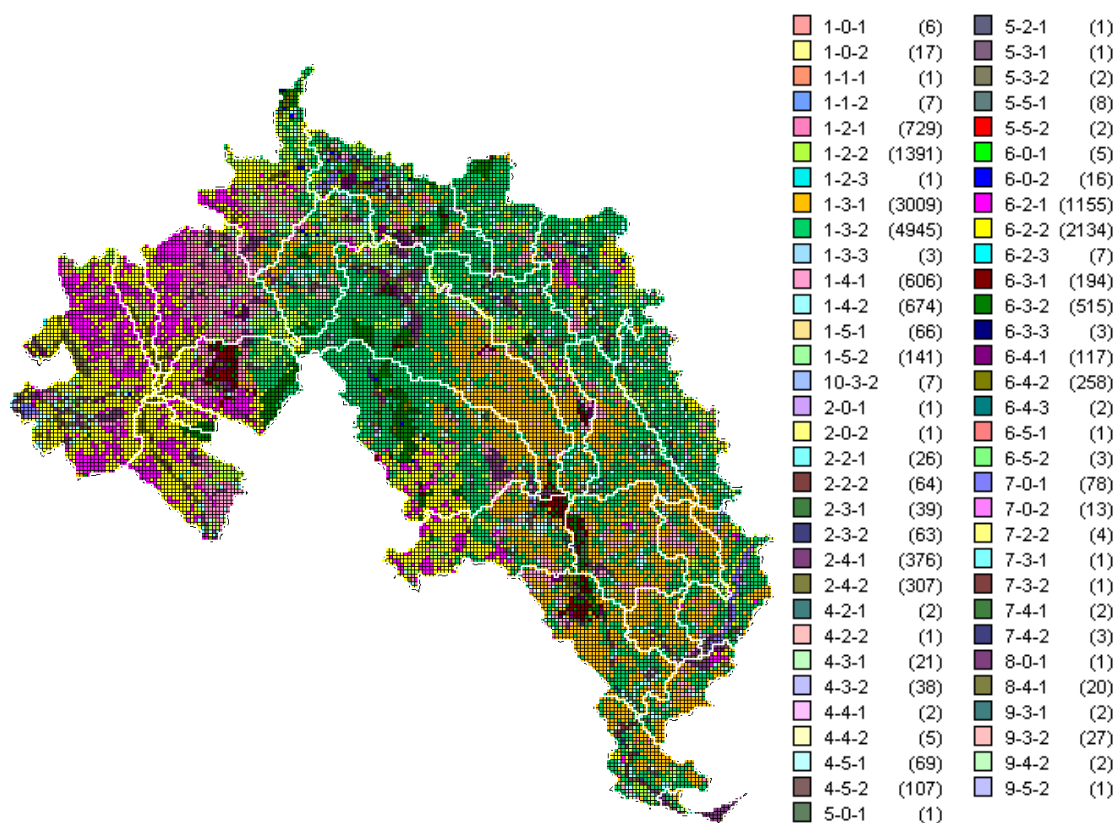
**C. Nachylenia stoków (DTED2): 1 – poniżej 0,5%, 2 – 0,5-5%, 3 – powyżej 5%;**

**D. Głębokość I poziomu wodonośnego (MHgP): 1 – 1 m lub mniej, 2 – 1-2 m, 3 – 2-3 m, 4 – 3-5 m, 5 – 5-10 m, 6 – 10-15 m, 7 – więcej niż 15 m poniżej poziomu terenu.**

*Źródło: opracowanie własne*

**Funkcja wizualizacyjna** dynamicznego potencjału informacyjnego obejmuje zakres możliwości przetwarzania kartograficznego przy realizacji kartograficznych metod badań z wykorzystaniem technologii geoinformacyjnych. Określenie tej funkcji jest ściśle uzależnione od celu badań oraz wizualizacyjnych preferencji operatora w procesie kompilowania kartograficznego (Dent 1990, Medyńska-Gulij 2007).

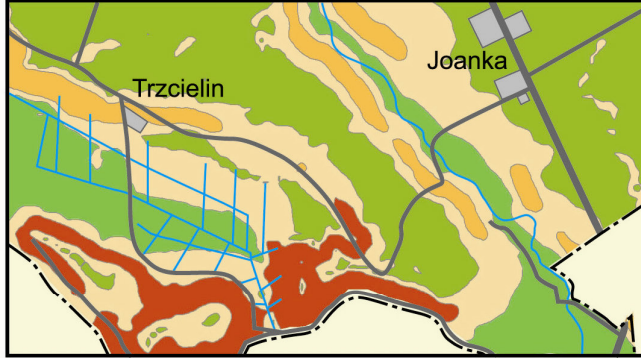
Wykorzystanie funkcji wizualizacyjnej ma szczególne znaczenie w kartograficznej metodzie badań i może przebiegać kilkietapowo, np. od etapu klasyfikacji terenu (ryc. 5.9.), poprzez etap oceny (ryc. 5.10) do etapu syntezy na drodze nakładania ocen cząstkowych (ryc. 5.11.). Najczęściej funkcja wizualizacyjna wymaga wcześniejszego przetworzenia danych atrybutowych (funkcja analityczna, tab. 5.8.).



**Ryc. 5.8. Klasyfikacja obszaru badań na podstawie przetworzonych do siatki kwadratów danych z baz CLC, MHP i DTED2 (por. ryc 5.8. A, B i C). W nawiasach legendy podano liczebność**

**Źródło: opracowanie własne**

### Ukształtowanie powierzchni terenu



skala 1:25 000

0 500m

#### Obszary płaskie:

- den dolinnych
- wyniesione ponad dna dolin

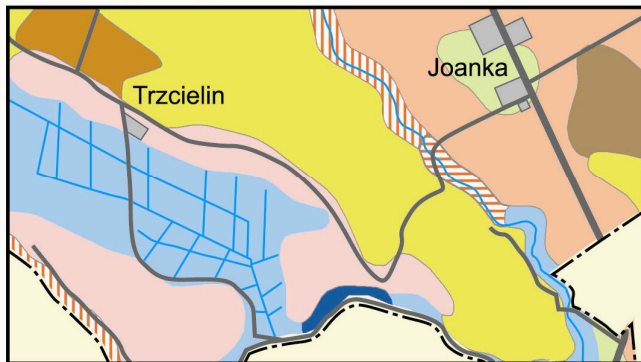
#### Obszary krawędziowe:

- połogie
- strome

#### Obszary zróżnicowane morfometrycznie:

- z przewagą stromych wzniesień

### Budowa geologiczna

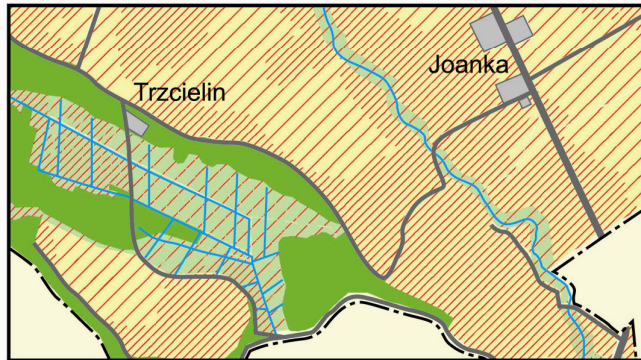


skala 1:25 000

0 500m

- gliny zwałowe
- ▨ namuły piaszczyste i piaski rzeczne
- ▨ namuły torfiaste
- piaski i żwiry ozów
- piaski i żwiry wodnolodowcowe
- piaski lodowcowe
- piaski lodowcowe na glinach zwałowych
- piaski pyłowate, zwietrzelinowe (eluwialne) na glinach
- torfy
- torfy na gytach

### Użytkowanie terenu



skala 1:25 000

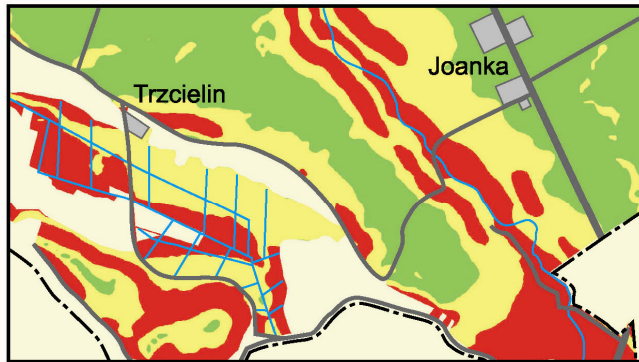
0 500m

- wody powierzchniowe
- lasy
- obszary użytkowane rolniczo:
- grunty orne klas bonitacyjnych:
- ▨ - średnich
- ▨ - niskich
- użytki zielone klas bonitacyjnych:
- ▨ - średnich
- ▨ - niskich
- obszary zabudowane lub w zabudowie

**Ryc. 5.9. Fragmenty mapy tematycznych prezentujące ukształtowanie powierzchni terenu, budowę geologiczną i użytkowanie terenu**  
**Źródło: Bródka, Medyńska-Gulij, Kaczmarek 2010**

Ocena przydatności terenu do zainwestowania  
ze względu na ukształtowanie powierzchni terenu

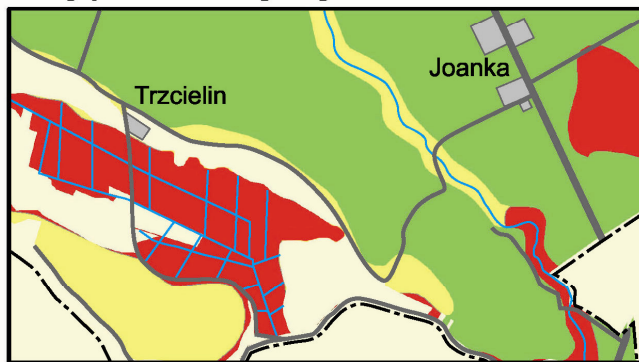
skala 1:25 000  
0 500m



Tereny o warunkach:  
■ korzystnych  
■ korzystnych z ograniczeniami  
■ niekorzystnych  
 - - - granice gminy  
 == drogi  
 — ciek  
 ■ zabudowa

Ocena przydatności terenu do zainwestowania  
ze względu na warunki geologiczne

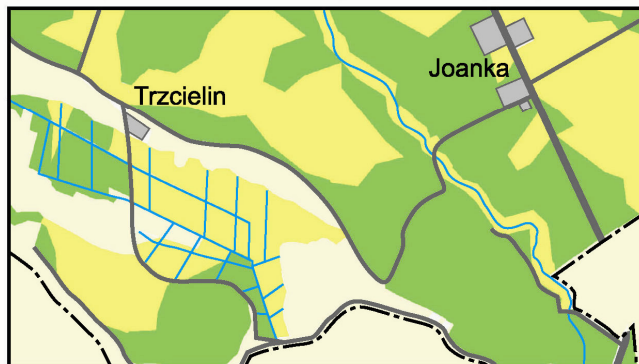
skala 1:25 000  
0 500m



Tereny o warunkach:  
■ korzystnych  
■ korzystnych z ograniczeniami  
■ niekorzystnych  
 - - - granice gminy  
 == drogi  
 — ciek  
 ■ zabudowa

Ocena przydatności terenu do zainwestowania ze względu na  
sposób użytkowania i zdolności do pełnienia funkcji przyrodniczych

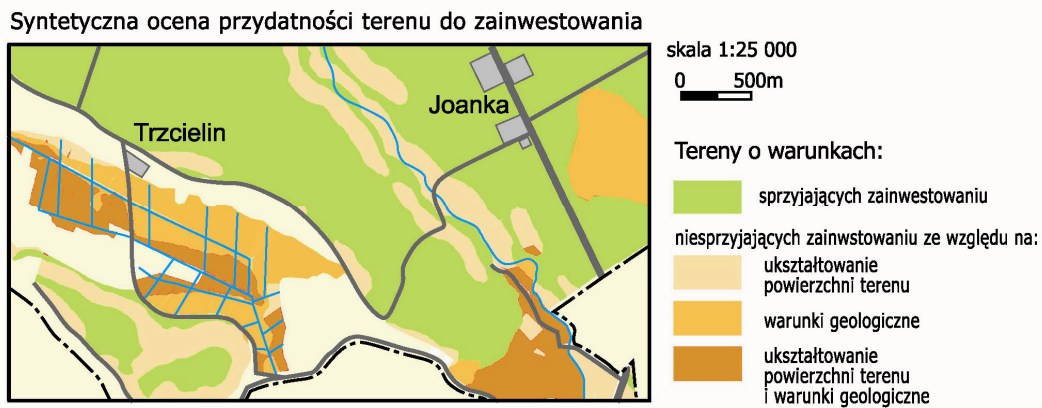
skala 1:25 000  
0 500m



Tereny o warunkach:  
■ korzystnych  
■ korzystnych z ograniczeniami  
 - - - granice gminy  
 == drogi  
 — ciek  
 ■ zabudowa

Ryc. 5.10. Fragmenty mapy tematycznych prezentujące oceny przydatności terenów do zainwestowania

Źródło: Bródka, Medyńska-Gulij, Kaczmarek 2010



*Ryc. 5.11. Fragmenty mapy dotyczącej syntetycznej oceny przydatności terenów do zainwestowania  
Źródło: Bródka, Medyńska-Gulij, Kaczmarek 2010*

## 5.5. Dualizm pojęcia potencjału informacyjnego danych przestrzennych

Zaprezentowana w rozdziale 5 koncepcja potencjału informacyjnego danych przestrzennych charakteryzuje się dualizmem podejścia do wyznaczania jego wartości. Takie podejście wynika z potrzeb, jakie stają przed badaczem zamierzającym wykorzystać istniejące dane przestrzenne w swoich badaniach. Z jednej strony jest to konieczność określenia stopnia ufności do danych, wyznaczenia poziomu ich precyzji, a z drugiej ustalenie potencjalnych możliwości przetworzenia danych źródłowych dla osiągnięcia celów badawczych.

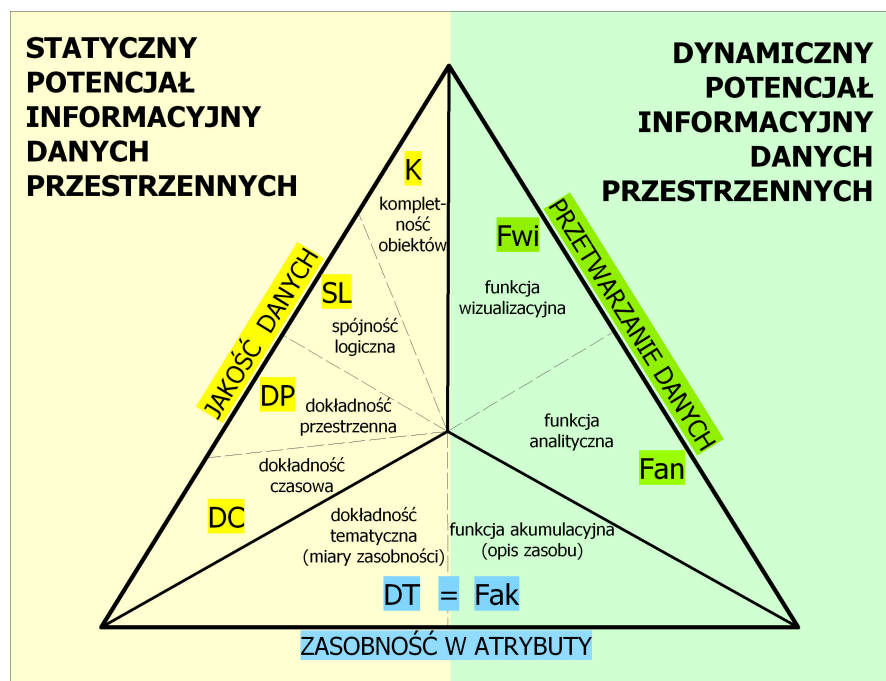
Należy zatem rozdzielić pojęcia statycznego i dynamicznego potencjału informacyjnego danych przestrzennych oraz określić ich przeznaczenie w ocenie zasobu baz danych przestrzennych. Pojęcie statycznego i dynamicznego potencjału informacyjnego przyjęto za Saliszczewem (1998), który określił, że statyczny potencjał mapy jest sumą zawartych w niej wiadomości, a dynamiczny potencjał mapy wyznacza możliwości rozszerzenia wiedzy w wyniku wnioskowania indukcyjnego i dedukcyjnego.

Na statyczny potencjał informacyjny składają się (ryc. 5.12.):

- elementy jakości danych wg PN ISO 19113: kompletność, spójność logiczna, dokładność przestrzenna i dokładność czasowa oznaczane w zakresie istotnym dla celu użytkownika danych;

- dokładność tematyczna rozumiana jako stopień spełniania oczekiwań użytkownika w zakresie liczby i jakości danych atrybutowych rozpatrywanych w aspekcie poziomów ujęcia atrybutów.

Celem wyznaczania statycznego potencjału informacyjnego jest określenie syntetycznej miary jakości danych i zasobności w dane atrybutowe. Potencjał statyczny przyjmują postać liczbowych miar bezwzględnych lub względnych.



Ryc. 5.12. Składowe statycznego i dynamicznego potencjału informacyjnego danych przestrzennych  
 Źródło: opracowanie własne

Na dynamiczny potencjał informacyjny składają się (ryc. 5.12.):

- funkcja akumulacyjna danych, stanowiąca opis atrybutów w podziale na poziomy ujęcia atrybutów;
- funkcja analityczna, która opisuje zakres przetwarzania analitycznego danych atrybutowych;
- funkcja wizualizacyjna, określająca ramy kompilowania kartograficznego danych w obrębie zainteresowań użytkownika.

Celem wyznaczania dynamicznego potencjału informacyjnego jest określenie potencjalnych możliwości realizacji analiz przestrzennych oraz kompilowania kartograficznego przedmiotowych danych w zakresie przyjętego celu badawczego i w oparciu o dostępne dane atrybutowe. Potencjał dynamiczny jest określany jako miara opisowa w postaci zestawień tabelarycznych (por. tab. 5.7. i tab. 5.9.).

## 6. Podsumowanie

Przeprowadzone badania potencjału informacyjnego danych przestrzennych miały na celu wykazanie potrzeby wyznaczenia tego jego wartości przy realizacji kartograficznych badań środowiska przyrodniczego. Rozprawa zwraca uwagę Czytelnika na bogactwo krajowych baz danych przestrzennych i ich aplikacyjność w badaniach środowiska przyrodniczego, co jest szczególnie istotne obecnie, w okresie ciągłego przyrostu danych w ramach budowania krajowej infrastruktury informacji przestrzennej.

Przedstawione w rozprawie badania wskazały na ważny aspekt wykorzystania krajowych baz danych przestrzennych w badaniach przyrodniczych. Zaproponowano procedurę wyznaczenia potencjału informacyjnego, która może być podstawą rekomendacji krajowych baz danych przestrzennych do wykorzystania w badaniach naukowych. Procedura ta prowadzi do określenia swoistych metadanych uwzględniających oczekiwania użytkownika względem bazy danych z punktu widzenia jego celów badawczych. W literaturze problemu wiele miejsca poświęca się wyznaczaniu jakości danych przestrzennych, jednakże jest to podejście z punktu widzenia producenta lub administratora danych. Schematy wyznaczenia metadanych również odnoszą się do technologii wytworzenia danych i stawiają nacisk na zobiektywizowanie oceny. Natomiast potencjał informacyjny w ujęciu niniejszej rozprawy stanowi subiektywne podejście do oceny, zgodne z dążeniami użytkownika danych przestrzennych.

Szczególnie istotne w podjętych badaniach jest zjawisko redundancji informacji zawartej w różnych bazach danych. Koncepcja potencjału informacyjnego baz danych przestrzennych pozwala na optymalizację wyboru danych do badań z punktu widzenia celu. Wyboru można dokonać z poziomu uproszczonej oceny baz danych przestrzennych lub poprzez porównanie potencjałów informacyjnych baz danych w zakresie badanego komponentu środowiska przyrodniczego.

W trakcie postępowania badawczego zauważono jednak, że wyznaczenie potencjału informacyjnego musi odnosić się do pojedynczej warstwy tematycznej. Wynika to z konieczności zastosowania ujednoczonych miar składowych oceny, a więc kompletności, spójności logicznej, dokładności położenia, czasowej, oraz tematycznej w ramach całego badanego zbioru. Dlatego zbiór badanych danych musi być homogeniczny. Określanie potencjału informacyjnego grupy warstw tematycznych może być zrealizowane po zdefiniowaniu liczbowych wartości wag, jakie przypisuje się poszczególnym warstwom

tematycznym. Porównanie potencjału informacyjnego warstw pochodzących z różnych baz danych, a przedstawiających ten sam element komponentu środowiska przyrodniczego może być zrealizowane po zbadaniu relacji semantycznych między tymi warstwami. W świetle przeprowadzonych badań nie wydaje się aby było możliwe i celowe określanie potencjału informacyjnego całych baz danych, ponieważ z oczywistych względów uzyskane rezultaty nie będą porównywalne.

Potencjał informacyjny danych przestrzennych może być realizowany w oparciu o kilkietapową procedurę, która obejmuje: określenie pożądanego modelu pojęciowego danych, wyznaczenie jakości danych w zakresie kompletności, spójności logicznej, dokładności położenia i dokładności czasowej, zdefiniowanie oczekiwanego zestawu danych atrybutowych oraz wyznaczenie dokładności tematycznej danych. Dokładność tematyczna jest kluczowym elementem potencjału informacyjnego, ponieważ atrybuty w decydujący sposób wpływają na możliwości analityczne i wizualizacyjne warstwy tematycznej. Atrybuty powinny być rozpatrywane w odniesieniu do koncepcji skal pomiarowych w kartografii. Dlatego opracowano klasyfikację poziomów ujęcia atrybutów jakościowych i ilościowych wraz ze schematami generowania atrybutów. Jest to szczególnie istotne osiągnięcie pracy, ponieważ skale pomiarowe dla danych cyfrowych powinny uwzględniać możliwość przetwarzania danych atrybutowych.

Istota przedmiotowych baz danych w zastosowaniach badawczych kryje się właśnie w możliwościach geoinformacyjnego przetwarzania. Dlatego w końcowej części pracy zwrócono uwagę na możliwość dynamicznego ujęcia potencjału informacyjnego danych przestrzennych. Polega ono na określeniu potencjalnych możliwości realizacji analiz przestrzennych oraz kompilowania kartograficznego w zakresie przyjętego celu badawczego w oparciu o istniejące źródła danych atrybutowych. Dynamiczny potencjał informacyjny jest miarą pełnienia przez dane przestrzenne funkcji akumulacyjnej, analitycznej i wizualizacyjnej.

W podsumowaniu rozprawy autor rozdziela pojęcia statycznego i dynamicznego potencjału informacyjnego danych przestrzennych i przypisuje im role. Koncepcja statycznego i dynamicznego potencjału informacyjnego została przejęta z pojęć wprowadzonych przez Saliszczewa (1998) w odniesieniu do map analogowych. Statyczny potencjał jest miarą spełniania przez dane oczekiwań ze strony naukowca badającego środowisko przyrodnicze. Natomiast potencjał dynamiczny opisuje możliwości pełnienia

przez dane źródłowe funkcji akumulacyjnej, analitycznej i wizualizacyjnej przy założeniu wykorzystania metod przetwarzania geoinformacyjnego i kompilowania kartograficznego.

Przeprowadzone w rozprawie postępowanie badawcze pozwoliło na sformułowanie wniosków o charakterze ogólnym i szczegółowym.

Wnioski ogólne odnoszą się do celu głównego rozprawy:

1. Wyznaczanie potencjału informacyjnego krajowych baz danych przestrzennych jest kluczem do wykorzystywania tych danych w badaniach środowiska przyrodniczego. Jest to narzędzie badawcze, które z jednej strony pozwala na wybór właściwych danych źródłowych do procedury badawczej w sytuacji redundancji danych i niepewności dotyczącej ich jakości (potencjał statyczny), a z drugiej strony umożliwia zdefiniowanie zakresu realizacji przez te dane funkcji akumulacyjnej, analitycznej i wizualizacyjnej w pracach badawczych (potencjał dynamiczny).
2. Możliwe jest przeprowadzanie procedur wyznaczania potencjału informacyjnego krajowych danych przestrzennych zarówno w ujęciu statycznym jako miar liczbowych jakości, jak i dynamicznym w postaci opisu funkcjonalności danych.
3. Statyczny potencjał informacyjny danych przestrzennych określa poziom zasobności bazy w dane atrybutowe istotne z punktu widzenia użytkownika. W mniejszym stopniu wartość potencjału statycznego uwzględnia miary jakości wg normy ISO 19113.
4. Dynamiczny potencjał informacyjny danych przestrzennych odnosi się do możliwości przetworzenia danych atrybutowych w systemach geoinformacyjnych w celach analitycznych lub wizualizacyjnych.

Wnioski szczegółowe rozprawy odnoszą się do realizacji celów cząstkowych:

1. Analiza zawartości baz danych przestrzennych wskazała na dużą redundancję danych, co rodzi potrzebę określenia procedur wyboru danych optymalnych do konkretnych zastosowań.
2. Najlepsze dane, o najwyższej dokładności i najszerszym zastosowaniu w analizach lokalnych powstają w ostatnim czasie (BDOT, EGIB, LMN), a wcześniejsze bazy (najczęściej tematyczne) nadają się bardziej do opracowań regionalnych.

3. Testowanie jakości danych wykazało znaczne różnice w jakości danych, niezauważalne na poziomie ogólnej oceny jakości danych i dostarczyło rzetelnych miar jakości do procedury wyznaczania potencjału informacyjnego.
4. Badanie jakości danych przestrzennych może być realizowane na dużych zbiorach danych z wykorzystaniem prostych narzędzi geoinformacyjnych i statystycznych.
5. Kluczowe zagadnienie potencjału informacyjnego danych przestrzennych obejmuje określenie dokładności tematycznej.
6. Zaproponowana w pracy klasyfikacja poziomów ujęcia atrybutów (tab. 5.1.) nawiązuje do klasycznego zagadnienia skal pomiarowych w kartografii, a jednocześnie odnosi się do danych cyfrowych.
7. W pracy zdefiniowano zależności między poziomami ujęcia atrybutów, zarówno dla poziomów jakościowych jak i ilościowych oraz ujęto je w schematy generowania atrybutów (ryc. 5.1. i 5.2.)
8. Badanie potencjału musi się odnosić do pojedynczej warstwy tematycznej ze względu na niejednorodny zakres semantyczny warstw oraz różną zawartość atrybutową.
9. Wyznaczanie potencjału informacyjnego warstw wymaga określenia zależności semantycznych pomiędzy badanymi warstwami, ustalenia pożądanego modelu pojęciowego oraz zestawu atrybutów, a także metod wyznaczania jakości danych.
10. Aby wyznaczyć potencjał informacyjny grupy warstw tematycznych należy określić wagę poszczególnych warstw w tej grupie. Niecelowe jest wyznaczanie sumarycznego potencjału informacyjnego całych baz danych ze względu na ich różne zakresy semantyczne.

W obecnej sytuacji danych przestrzennych w Polsce uzasadnione wydaje się stosowanie kompleksowej analizy baz danych poprzez wyznaczanie wskaźników statycznego potencjału informacyjnego danych przestrzennych warstwy (PIWT) lub grupy warstw tematycznych (PIGWT) oraz opisu potencjału dynamicznego. Notowany w ostatnich latach wzrost poziomu jakości baz danych oraz tendencja do tworzenia ogólnonarodowych infrastruktur danych przestrzennych (Dyrektywa INSPIRE, Ustawa PIIP) spowoduje prawdopodobnie zainteresowanie się dynamicznym aspektem potencjału informacyjnego. Dlatego najważniejszym kierunkiem dalszych badań w tym zakresie powinny być próby skwantyfikowania tej miary.

## Literatura:

- Altheide P. 2008: Spatial Data Transfer Standard (SDTS). [In:] Shekhar S., Xiong H. (eds) 2008: Encyclopedia of GIS. SpringerScience+Business Media New York 2008: 1087-1095.
- Bac-Bronowicz J., Kołodziej A., Kowalski P., Olszewski R. 2007: Konwersja bazy danych VMap L2 pierwszej edycji do struktury użytkowej. Roczniki Geomatyki V, 2: 21-28.
- Barszczyńska M., Łasut E., Walczykiewicz T. 2008: Identyfikacja zasobów danych przestrzennych gospodarki wodnej w kontekście potrzeb planowania. Roczniki Geomatyki VI, 6: 7-14.
- Bartkowski T. 1977: Metody badań geografii fizycznej. PWN Warszawa-Poznań.
- Berlant A. M. 1978: Kartograficzeskij metod issliedowanija. Izdat. Moskow. Uniw., Moskwa: ss. 254
- Bernhardsen T. 2002: Geographic Information Systems – an introduction. John Wiley & Sons, New York 2002: pp. 422.
- Bertin 1983: Semiology of graphics: diagrams, networks, maps. The University of Wisconsin Press: ss. 415.
- Berus T., Kołodziej A., Olszewski R. 2007: Koncepcja harmonizacji baz danych tematycznych SOZO i HYDRO z referencyjną bazą VMap L2 drugiej edycji. Roczniki Geomatyki V, 2: 29-38.
- Bielecka E. 2006: Systemy informacji geograficznej. Teoria i zastosowania. Wyd. PJWSTK Warszawa: ss. 229.
- Bielecka E. 2010: Zasady oceny jakości danych przestrzennych oraz ich zastosowanie do oceny jakości danych gromadzonych w TBD. [W:] Modelowanie danych przestrzennych. Roczniki Geomatyki VIII, 4: 53-66.
- Bielecka E., Ciołkosz A. 2009: Baza danych o pokryciu terenu w Polsce CLC-2006. Pol. Przegl. Kartogr., 41: 227-236.
- Board Ch. 1967: Maps as models. [In:] R. J. Chorley, P. Haggett (eds), Models In geography. London, Methuen 1967: 671-725.
- Borowicz D., Krupski J. 2009: Wykorzystanie Bazy Danych Ogólnogeograficznych do opracowania map tematycznych. [W:] W. Żyszkowska, W. Spallek (red.), Bazy danych w kartografii. Główne problemy współczesnej kartografii, Wrocław: 48-72.
- Brodlie K., Dykes J., Gillings M., Haklay M., Kitchin R., Kraak M.-J. 2001: Geography in VR: context. [In:] P. Fisher, D. Unwin: Virtual Reality in Geography. Taylor & Francis, London and New York 2002: 7-16.
- Bródka S. 2007: Zakres informacji o środowisku przyrodniczym wykorzystywany na potrzeby oceny potencjału rekreacyjnego. [W:] B. Medyńska-Gulij, L. Kaczmarek (red.), Informacja geograficzna w kształtowaniu i ochronie środowiska przyrodniczego. Bogucki Wyd. Naukowe, Poznań: 113-130.
- Bródka S. red. 2010: Praktyczne aspekty ocen środowiska przyrodniczego. Bogucki Wyd. Naukowe, Poznań: ss. 330.
- Bródka S., Medyńska-Gulij B., Kaczmarek L. 2010: Projekty oceny środowiska przyrodniczego w technologii GIS. [W:] S. Bródka (red.), Praktyczne aspekty ocen środowiska przyrodniczego. Bogucki Wyd. Naukowe, Poznań: 247-330.
- Chang K.-t. 2004: Introduction to Geographic Information Systems. 2-nd ed., McGraw-Hill, New York 2004: pp. 400.
- Ciołkosz A., Bielecka E. 2005: Pokrycie terenu w Polsce. Bazy danych CORINE Land Cover. Inspekcja Ochrony Środowiska, Biblioteka Monitoringu Środowiska, Warszawa: ss. 76.
- Clarke K. C. 1995: Analytical and computer cartography. 2<sup>nd</sup> ed., Prentice-Hall, Engelwood Cliffs, New Jersey.

- CORINE Land Cover Technical Guide (1994). <http://sia.eionet.europa.eu>
- Czerny A. 1994: Rozwój koncepcji modelowania kartograficznego. Pol. Przegl. Kartogr., t. 26, nr 4: 185-201.
- Czerny A., Kałamucki K., Ostrowski W., Żyszkowska W., 1999: Dorobek i stan polskiej kartografii geograficznej w ostatnim dziesięcioleciu Pol. Przegl. Kartogr., 31: 3-14.
- Delaney J. 1999: Geographic Information Systems, an introduction. Oxford Univ. Press, New York.
- Dent B. D. 1990: Cartography: Thematic map design. 2<sup>nd</sup> ed., Dubuque, IA: Wm. C. Brown: ss. 422.
- Dent B. D. (1999), Cartography: Thematic Map Design, wyd. piąte, McGraw-Hill, Boston.
- Digital Cartographic Data Standard Task Force (DCDSTF), 1988: The proposed standard form digital cartographic data. Amer. Cartographer, 15, 1.
- Digital Geographic Information Exchange Standard DIGEST edition 2.1. (2000)  
<http://www.dgiwg.org/digest/>
- Dykes, J., MacEachren A. M., Kraak M. J. 2005: Exploring geovisualization. Elseviers. Imprint: PERGAMON Amsterdam.
- Dyrektywa ustanawiająca infrastrukturę informacji przestrzennej we Wspólnocie Europejskiej nr 2007/2/WE z dnia 14 marca 2007 r.
- Dzikowska T., Szczepański J., Akińcza M. 2007: Możliwości wykorzystania mapy sozologicznej w planowaniu prac z zakresu geodezyjnego urządzania obszarów wiejskich. [W:] B. Medyńska-Gulij, L. Kaczmarek (red.), Informacja geograficzna w kształtowaniu i ochronie środowiska przyrodniczego. Bogucki Wyd. Naukowe, Poznań: 131-138.
- Eckes K. 2006: Modelowanie rzeczywistości geograficznej w systemach Informacji Przestrzennej. Roczniki Geomatyki IV, 2: 43-73.
- Fisher P. F. 1999: Models of uncertainty in spatial data. [W:] P. Longley, M. Goodchild, D. M. Maguire and D. Rhind (red.), Geographical information systems: principles, techniques, management and applications. John Wiley & Sons, New York.
- Frączek I. 1981: Zagadnienie skal pomiarowych w kartografii. Pol. Przegl. Kartogr., 13, 3-4: 126-135.
- Freitag U. 1971: Semiotik und Kartographie. Über die Anwendung kybernetischer Disziplinen in der theoretischen Kartographie. w: Kartographische Nachrichten, 21, 171-182.
- Głazewski A. 2006: Modele rzeczywistości geograficznej a modele danych przestrzennych. Pol. Przegl. Kartogr., 38, 3:217-225.
- Gogołek W. 2007: Dane przestrzenne w działalności polskiej służby geologicznej. Roczniki Geomatyki V, 7: 47-50.
- Gomasasca M. A. 2009: Basic of geomatics. SpringerScience+Business Media B. V.
- Goodchild M. F. 2008: Spatial data analysis. [In:] Shekhar S., Xiong H. (eds) 2008: Encyclopedia of GIS. SpringerScience+Business Media New York 2008: 200-203.
- Gotlib D., Iwaniak A., Olszewski R. 2007a: GIS obszary zastosowań. PWN Warszawa, ss. 250.
- Gotlib D., Iwaniak A., Olszewski R. 2007b: Harmonizacja baz danych referencyjnych – niezbędny krok ku budowie krajowej infrastruktury danych przestrzennych. Roczniki Geomatyki, V, 1: 35-38.
- Gotlib D., Olszewski R. 2005: Możliwość wymiany danych między SITop a bazami danych Map. [W:] Makowski A. (red.): System informacji topograficznej kraju. Ofic. Wyd. Polit. Warszawskiej: 191-195.
- Graf R. 2004: Zastosowanie „Mapy hydrograficznej Polski w skali 1:50 000” w systemie oceny podatności wód podziemnych na zanieczyszczenia. [W:] L. Kozacki, B. Medyńska-Gulij (red.),

- Kartografia tematyczna w kształtowaniu środowiska geograficznego. Mat. Ogólnop. Konf. Kartogr., 25: 219-228.
- Graf R. 2007: Przydatność wybranych baz danych tematycznych w badaniach ekohydrologicznych systemów rzecznych. [W:] B. Medyńska-Gulij, L. Kaczmarek (red.), Informacja geograficzna w kształtowaniu i ochronie środowiska przyrodniczego. Bogucki Wyd. Naukowe, Poznań: 139-150.
- Guptill S. C., Morrison J. L. 1995: Elements of spatial data quality. Elsevier Science: ss. 202.
- Hake G., Grünreich D., Meng L. 2002: Kartographie. De Guyter Lehrbuch, Berlin 2002, ss. 564.
- Harvey F. 2008: A primer of GIS: fundamental geographic and cartographic concepts. The Guilford Press, New York 2008: pp. 310.
- Herbich P. i in, 2004 – Mapa hydrogeologiczna Polski w skali 1:50 000, udostępnianie, weryfikacja, aktualizacja i rozwój. Instrukcja, MŚ-PIG, Warszawa.
- Instrukcja opracowania i aktualizacji Mapy geologiczno-gospodarczej Polski w skali 1:50 000. PIG 2002.
- Instrukcja opracowania komputerowej edycji Mapy hydrogeologicznej Polski w skali 1:50 000. Część I Opracowanie autorskie. PIG 1999: ss. 93.
- Instrukcja opracowania komputerowej edycji Mapy hydrogeologicznej Polski w skali 1:50 000. Część II Opracowanie komputerowe. PIG 1999: ss. 65.
- Instrukcja opracowania Mapy Geośrodowiskowej Polski w skali 1:50 000. PIG, Warszawa 2005: ss. 117.
- Instrukcja techniczna G-5 (2003)
- Instrukcja wydania i opracowania SMGP
- ISO/AVI 19138:2002 *Geographic information – Data quality measures*.
- Iwaniak A. 2007: Infrastruktura informacji geodezyjnej i kartograficznej. Roczniki Geomatyki V, 6: 63-69.
- Jezioro P., Kozak J. 2004: Wprowadzenie do systemów informacji geograficznej. Ćwiczenia. IGiGP Uniw. Jagiellońskiego, Kraków: ss.109.
- Kaczmarek L. 2007: Pozyskiwanie i przetwarzanie danych do monitoring środowiska przyrodniczego w systemach geoinformacyjnych. [W:] B. Medyńska-Gulij, L. Kaczmarek (red.), Informacja geograficzna w kształtowaniu i ochronie środowiska przyrodniczego. Bogucki Wyd. Naukowe, Poznań: 151-164.
- Kaczmarek L. 2010: Cyfrowe bazy danych w Polsce. [W:] S. Bródka (red.), Praktyczne aspekty ocen środowiska przyrodniczego. Bogucki Wyd. Naukowe, Poznań: 118-126.
- Kaczmarek L., Medyńska-Gulij B. 2006: Krajowy zasób danych przestrzennych i jego wykorzystanie w badaniach środowiska przyrodniczego. [W:] Kostrzewski A., Czerniawska J. (red.), Przemiany środowiska geograficznego Polski północno-zachodniej. Bogucki Wyd. Naukowe, Poznań: 203-210.
- Kaczmarek L., Medyńska-Gulij B. 2007a: Przetwarzanie baz danych źródłowych w celach ochrony środowiska przyrodniczego w planowaniu przestrzennym. [W:] M. Kistowski, B Korwel-Lejkowska (red), Waloryzacja środowiska przyrodniczego w planowaniu przestrzennym. Problemy Ekol. Krajobrazu, XIX, Gdańsk-Warszawa: 271-278.
- Kaczmarek L., Medyńska-Gulij B. 2007b: Źródła i metody pozyskiwania danych przestrzennych w badaniach środowiska przyrodniczego. Bogucki Wyd. Naukowe, Poznań: ss. 145.
- Kaczmarek L., Medyńska-Gulij B. 2008: Method of assessing spatial data for the purpose of monitoring the natural environment. [In:] A. Demirci et. al., 5th International Conference on Geographic Information Systems (ICGIS), Fatih Univ. Publications Istanbul: 441-448.

- Kaniecki A. 2004: Mapa hydrograficzna Polski w skali 1:50 000 – ujęcie historyczne. [W:] L. Kozacki, B. Medyńska-Gulij (red.), Kartografia tematyczna w kształtowaniu środowiska geograficznego. Mat. Ogólnop. Konf. Kartogr., 25: 26-37.
- Korycka-Skorupa J. 2002a: Od danych do mapy. Część I. Pol. Przegl. Kartograf., 34, 2: 91-102.
- Korycka-Skorupa J. 2002b: Od danych do mapy. Część II. Pol. Przegl. Kartograf., 34, 3: 175-188.
- Kraak M.-J., Ormeling F. 1998: Kartografia, wizualizacja danych przestrzennych. PWN, Warszawa: ss. 276.
- Kraak M.-J., Ormeling F. 2003: Cartography: visualization of geospatial data. 2<sup>nd</sup> ed., Pearson Education Limited, ss. 205.
- Kraak M.-J., Ormeling F. 2010: Cartography: visualization of geospatial data. 3<sup>rd</sup> ed., Pearson Education Limited.
- Kurzeja G., Szlarczyńska H. 2007: Mapa Wektorowa Poziomu Drugiego – zalety i możliwości wykorzystania. [W:] B. Medyńska-Gulij, L. Kaczmarek (red.), Informacja geograficzna w kształtowaniu i ochronie środowiska przyrodniczego. Bogucki Wyd. Naukowe, Poznań: 23-30.
- Laurini R., Thompson D. 1992: Fundamentals of spatial information systems. Academic Press, London.
- Lehmann E., Ogrissek R. 1988: Thematic Cartography. [W:] Basic cartography for students and technicians: 85-104.
- Lexikon der Kartographie und Geomatik, 2001, red. Bolmann J., Koch W.G., Spektrum Akademischer Verlag, Springer, Heidelberg-Berlin.
- Longley P. A., Goodchild M. F., Maguire D. J., Rhind D. W. 2006: GIS – teoria i praktyka. Wyd. Naukowe PWN, Warszawa, ss. 520.
- Lorek D. 2011: Potencjał informacyjny map topograficznych z lat 1822-33 (Urmesstischblätter) z terenu Wielkopolski. Rozprawa doktorska UAM Poznań, maszynopis.
- Łowicki D., Stępniewska M. 2007: Priorytety w gospodarce wodno-ściekowej województwa wielkopolskiego. Przykład zastosowania numerycznej mapy geośrodowiskowej Polski w zarządzaniu środowiskiem. [W:] B. Medyńska-Gulij, L. Kaczmarek (red.), Informacja geograficzna w kształtowaniu i ochronie środowiska przyrodniczego. Bogucki Wyd. Naukowe, Poznań: 165-174.
- MacEarchen A. M. 1994: Visualization in modern cartography: Setting the agenda. [W:] A.M. MacEarchen i D.R. Taylor (red.), Visualization in Modern Cartography. Oxford, Pergamon: 1-12.
- Magnuszewski 1999: GIS w geografii fizycznej. Wyd. Naukowe PWN, Warszawa: ss. 188.
- Makowski A., 2005a: Spór o mapę. Wprowadzenie do pojęcia mapy. [W:] Główne problemy współczesnej kartografii, Co zwię się koncepcją mapy, Wrocław: 65-72.
- Makowski A. red. 2005b: System informacji topograficznej kraju. Ofic. Wyd. Politechniki Warszawskiej, ss. 454.
- McMaster R. B., Shea K. S. 1992: Generalization in digital cartography. Assoc. of Amer. Geographers, Washington: ss. 134.
- Medyńska-Gulij B. 2007: Pragmatyczne podstawy kompilowania kartograficznego. Wyd. Naukowe UAM, Poznań, ss. 140.
- Medyńska-Gulij B., Kaczmarek L. (red.) 2007: Informacja geograficzna w kształtowaniu i ochronie środowiska przyrodniczego. Bogucki Wyd. Naukowe, Poznań: ss. 207.
- Medyńska-Gulij B. 2010: Wizualizacja kartograficzna w ocenach środowiska przyrodniczego. [W:] S. Bródka (red.) 2010: Praktyczne aspekty ocen środowiska przyrodniczego. Bogucki Wyd. Naukowe, Poznań: 228-246.

- Muehrecke P.C., Muehrecke J. O., Kimerling A.J. 2001: *Map Use: Reading, Analysis and Interpretation*. 4rd ed., WI: JP Publications, Madison.
- Nałęcz T., 2007: Integracja danych przestrzennych o środowisku naturalnym – wyzwanie dla instytucji z branży ochrony środowiska? *Roczniki Geomatyki* 2007, t. V, z. 1, s. 81-90.
- OGC 1999:
- Olbricht G., Quick M., Schweikart J., 2002, *Desktop Mapping. Grundlagen und Praxis in Kartographie und GIS*”, Springer Verlag, Berlin.
- Olenderek H. 2000: *Historia Systemu Informacji Przestrzennej w lasach*. [W:] k. Okła (red.), *System informacji przestrzennej w Lasach Państwowych. Podręcznik użytkownika Leśnej Mapy Numerycznej*. Bogucki Wyd. Naukowe: 11-15.
- Olenderek H. 2010: *Możliwości wykorzystania najnowszych osiągnięć geomatyki w Lasach Państwowych*. *Roczniki Geomatyki*, VIII, 5: 87-94.
- Olszewski R. 2007: *Bazy danych tematycznych – rudymet czy konstytutywny komponent infrastruktury danych przestrzennych w Polsce?* [W:] B. Medyńska-Gulij, L. Kaczmarek (red.), *Informacja geograficzna w kształtowaniu i ochronie środowiska przyrodniczego*. Bogucki Wyd. Naukowe, Poznań: 31-42.
- Olszewski R., R. Graf, A. Macias, A. Kołodziej, T. Berus 2010, *Integracja baz danych tematycznych i referencyjnych w ramach tworzenia infrastruktury informacji przestrzennej w Polsce*, *Roczniki Geomatyki*, t 8, zeszyt 1(37) 89-98,
- Olszewski R., Stankiewicz M. 2004: *Mapa sozologiczna, geośrodowiskowa i hydrograficzna czy zintegrowana baza danych o środowisku geograficznym?* [W:] L. Kozacki, B. Medyńska-Gulij (red.), *Kartografia tematyczna w kształtowaniu środowiska geograficznego*. *Mat. Ogólnop. Konf. Kartogr.*, 25: 113-122.
- Ostrowski W. 2008: *Semiotyczne podstawy projektowania map topograficznych*. Uniw. Warszawski, Wyd. Geografii i Studiów Region., Warszawa 2008: ss. 210.
- O’Sullivan D., Unwin D. 2003: *Geographic information analysis*. J. Willey & Sons, Inc., New Jersey 2003: pp. 436.
- Pachelski W., Parzyński Z. 2007: *Aspekty metodyczne wykorzystania norm serii ISO 19100 do budowy georeferencyjnych składników krajowej infrastruktury danych przestrzennych*. *Roczniki Geomatyki* V, 3: 113-121.
- Paślawski J., J. Siwek 2010: *Użytkowanie map tematycznych*. [W:] J. Paślawski (red), *Wprowadzenie do kartografii i topografii*, Wyd. 2, Wyd. Nowa Era, Wrocław: 325-329.
- Peuquet D. J. 1984: *A conceptual Framework and comparison of spatial data models*. *Cartographica*, Vol. 21, no 4, 1984: 66-113.
- PN-EN ISO 19113:2005: *Informacja geograficzna. Podstawy opisu jakości*. (In english) Pol. Kom. Normalizacyjny, Warszawa 2009: ss 34.
- PN-EN ISO 19114:2005: *Informacja geograficzna – Procedury oceny jakości*. Pol. Kom. Normalizacyjny, Warszawa 2005: ss. 63.
- Przewoźniak 1979: *Struktura środowiska geograficznego Półwyspu Helskiego*. *Zesz. Nauk. BiNZ UG*, ser. Geogr. Fiz., 10.
- Ratajski L., 1973: *Metodyka kartografii społeczno-gospodarczej*. PPWK, Warszawa 2008: ss. 380.
- Ratajski L. 1978: *Unifikacja języka map tematycznych*. *Pol. Przegl. Kartogr.*, t. 10, nr 1: 22-26.
- Ratajski L. 1989: *Metodyka kartografii społeczno-gospodarczej*. PPWK Warszawa.
- Richling 1972: *Struktura krajobrazowa Krainy Wielkich Jezior Mazurskich*. *Prace i Studia IG UW*, 10 (Geogr. Fiz.), 4.

- Richling 1992: Kompleksowa geografia fizyczna. Wyd. Naukowe PWN, Warszawa: ss. 376.
- Richling 2007: Geograficzne badania środowiska przyrodniczego. Wyd. Naukowe PWN, Warszawa: ss. 323.
- Robinson A.H. Morrison J.L, Muehrecke P.C., Kimerling A.J., Guptil S.C. 1995: Elements of Cartography. 6 wyd.: Wiley, New York.
- Rozporządzenie MRRiB (2001)
- Rudowicz-Nawocka J. 2006: Przykłady wykorzystania danych z mapy glebowo-rolniczej z zastosowaniem Systemów Informacji Przestrzennej. Roczniki Geomatyki IV, 2: 135-138.
- Saliszczew K. A., 1955: O kartograficznym metodzie issledowania. Wiestn. Moskow. Uniw., Ser. Geogr., 10:161-170.
- Saliszczew K. A. 1967: Zadaci kartografii i aftomatizacja. Izw. Wysszych Zawied. Gieod. i Aerofot., 4: 7-10.
- Saliszczew K. A. 1998: Kartografia ogólna. Wyd. Naukowe PWN, Warszawa, ss.308.
- Sikorska-Maykowska M. 2004: Mapa geośrodowiskowa Polski w skali 1 : 50 000 - pomoc w zarządzaniu i ekorozwoju na szczeblu regionalnym. Przegl. Geol., 52, 10.
- Sikorska-Maykowska M. 2007: kartografia geologiczna i geośrodowiskowa w Państwowym Instytucie Geologicznym. [W:] B. Medyńska-Gulij, L. Kaczmarek (red.), Informacja geograficzna w kształtowaniu i ochronie środowiska przyrodniczego. Bogucki Wyd. Naukowe, Poznań: 53-62.
- Slocum T.A., McMaster R.B., Kessler F.C., Howard H.H. 2010: Thematic Cartography and Geographic Visualization, wyd. 3, Pearson Prentice Hall, Upper Saddle River.
- Sobczyński E. 2007: Systemy informacji geograficznej w wojsku. [W:] M. Kunz (red.), Systemy informacji geograficznej w praktyce (studium zastosowań). Wyd. Uniw. M. Kopernika, Toruń: 111-126.
- Spatial Data Transfer Standard. <http://mcmcweb.er.usgs.gov/sdts/>
- Spiess E., (2002), Map compilation., w: Basic Cartography for students and technicians, red. R. Anston, F. Ormeling, Reed Education and Professional Publishing.
- Standard Leśnej Mapy Numerycznej, 2005: ss. 133.
- Stankiewicz M. 2005: Współczesne rozumienie topografii. [W:] Makowski A. red. 2005: System informacji topograficznej kraju. Ofic. Wyd. Politechniki Warszawskiej: 49-56.
- Stones, Matthew 2003: Bazy danych i MySQL od podstaw. Wyd. Helion, Gliwice: ss. 560.
- System informacji o terenie. Podstawowa mapa kraju. Instrukcja techniczna K-1. Główny Geodeta Kraju, 1995: ss. 164.
- Tobler W. R. 1959: Automation and cartography. Geogr. Review, 49, 4: 526-534.
- Tobler W. R. 1961: Geographical filters and their inverses. Geogr. Analysis, 1, 3, [www.geog.ucsb.edu/~tobler](http://www.geog.ucsb.edu/~tobler)
- Tomlin C. D. 1990: GIS and cartographic modeling. Prentice Hall, New Jersey.
- Urbański 2008: GIS w badaniach przyrodniczych. Wyd. Uniw. Gdańskiego: ss. 252.
- US Federal Geographic Data Committee
- Ustawa o infrastrukturze informacji przestrzennej (Dz. U. nr 76, poz. 489 z dnia 4 marca 2010 r.)
- Wytyczne techniczne GIS-3. Mapa Hydrograficzna Polski, skala 1:50 000 w formie analogowej i numerycznej. GUGiK, Warszawa 2005: ss. 40.
- Wytyczne techniczne GIS-4. Mapa Sozologiczna Polski, skala 1:50 000 w formie analogowej i numerycznej. GUGiK, Warszawa 2005: ss. 58.

- Wytyczne Techniczne K-3.4. System Informacji o Terenie. Mapa Hydrograficzna Polski, skala 1:50 000 w formie analogowej i numerycznej. GUGiK 1997
- Wytyczne Techniczne K-3.6. System Informacji o Terenie. Mapa Sozologiczna Polski, skala 1:50 000 w formie analogowej i numerycznej. GUGiK 1997
- Wytyczne Techniczne. Baza danych topograficznych (TBD). GUGiK, Warszawa 2003
- Virrantaus K., D. Fairbairn, M-J. Kraak 2009: ICA Research Agenda on Cartography and GIScience, *Cartography and Geographic Information Science*, 36,2, 209-222.
- Yeung A. K. W., Brent Hall G. 2007: *Spatial database systems. Design, implementation and project management*. Springer: ss. 553.
- Zarządzenie nr 23 Dyrektora Generalnego Lasów Państwowych (1998).
- Zarządzenie nr 74 Dyrektora Generalnego Lasów Państwowych z dnia 23 sierpnia 2001r. w sprawie zdefiniowania standardu leśnej mapy numerycznej dla poziomu nadleśnictwa oraz wdrażania systemu informacji przestrzennej w nadleśnictwach.
- Zeiler T. A. 1999: *Modeling our World: the ESRI guide to geodatabase design*. ESRI Press, Redlands, pp. 199.
- Zhang J., Goodchild M. 2002: *Uncertainty in geographical information*. Taylor & Francis New York 2002: pp. 266.
- Zwoliński Z. 2009: *Rozwój myśli geoinformacyjnej*. [W:] GIS. Platforma integracyjna geografii, Z. Zwoliński (red.), Bogucki Wyd. Naukowe, Poznań: 9-22.
- Żynda S. 1978: *Podział środkowego Nadodrza na fizycznogeograficzne jednostki przestrzenne i ich ocena dla niektórych potrzeb planowania przestrzennego*. Wyd. UAM Poznań, ser. Geografia, 16.
- Żynda S. 2004: *Mapa sozologiczna Polski w skali 1:50 000 – ujęcie historyczne*. [W:] L. Kozacki, B. Medyńska-Gulij (red.), *Kartografia tematyczna w kształtowaniu środowiska geograficznego*. Mat. Ogólnop. Konf. Kartogr., 25: 26-37.
- Żyszkowska W. 2000: *Semiotyczne aspekty wizualizacji kartograficznej*. *Studia Geograficzne LXXIII*, Wyd. Uniwersytetu Wrocławskiego, Wrocław 2000: ss. 132.
- Żyszkowska W. 2003: *Analizy przestrzenne w systemach informacji geograficznej*. *Pol. Przegl. Kartograf.*, 35, 2: 100-113.

## Spis rycin:

- Ryc. 1.1. Procedura badawcza
- Ryc. 2.1. Cykl mapa-model
- Ryc. 2.2. Kartograficzna sieć komunikacyjna na tle trójstopniowego modelu kartograficznego (Hake i in. 2002)
- Ryc. 2.3. Schematyczne reprezentacje wirtualnej rzeczywistości a transformacja kartograficzna (Brodlić i in 2004)
- Ryc. 2.4. Przejście ze świata rzeczywistego do GIS poprzez tworzenie kolejnych modeli upraszczających rzeczywistość (Bernhardsen 2002)
- Ryc. 2.5. Model obiektów geometrycznych wg OGC (1999)
- Ryc. 2.6. Sposoby reprezentacji pola wg Longley i in. (2006)
- Ryc. 2.7. Schemat progów generalizacji według Ratajskiego (1973: 225)
- Ryc. 2.8. Kolejne etapy uproszczenia rzeczywistości wpływające na niepewność zobrazowań zjawisk i obiektów (Longley i in. 2006)
- Ryc. 2.9. Etapy oceny jakości danych geograficznych wg ISO 19114
- Ryc. 2.10. Problematyka badawcza rozprawy
- Ryc. 3.1. Podstawowe jednostki opracowania Bazy Danych Ogólnogeograficznych w skali 1:250000
- Ryc. 3.2. Stopień pokrycia Topograficzną Bazą Danych – pełna treść
- Ryc. 3.3. Stopień pokrycia wybranymi elementami Bazy Danych Obiektów Topograficznych
- Ryc. 3.4. Zakres przestrzenny opracowania Mapy Hydrograficznej Polski
- Ryc. 3.5. Zakres przestrzenny opracowania Mapy Sozologicznej Polski
- Ryc. 3.6. Dostępność bazy Szczegółowej Mapy Geologicznej Polski
- Ryc. 3.7. Stan opracowania Mapy Hydrogeologicznej Polski
- Ryc. 3.8. Stan opracowania Mapy Geologiczno-Gospodarczej Polski (kolor zielony z kratkowaniem) i Mapy Geośrodowiskowej Polski (kolor jasnozielony)
- Ryc. 3.9. Zestawienie waloryzujące najważniejsze cechy baz danych przestrzennych
- Ryc. 4.1. Znaczeniowy zakres poszczególnych warstw tematycznych płynących wód powierzchniowych trzech wybranych baz danych przestrzennych
- Ryc. 4.2. Gęstość sieci rzecznej w  $\text{km}/\text{km}^2$  w poszczególnych bazach danych
- Ryc. 4.3. Matryca różnic wskaźnika gęstości sieci hydrograficznej poszczególnych baz danych w siatce 10 x 10 km
- Ryc. 4.4. Matryca różnic wskaźnika gęstości sieci hydrograficznej bazy MHP i MPHP w siatce 10 x 10 km w zakresie warstw cieków zawierających nazwę
- Ryc. 4.5. Lokalizacja pól testowych do badań wskaźników kształtu i błędów lokalizacji cieków
- Ryc. 4.6. Trzy przykłady obrazujące dokładność lokalizacyjną i geometryczną MHP (linia zielona), MPHP (linia niebieska) oraz VML2 (linia czerwona) na tle mapy topograficznej w skali 1:10000
- Ryc. 4.7. Wskaźnik jeziorności (jednostki niemianowane) w poszczególnych bazach danych
- Ryc. 4.8. Matryca różnic wskaźnika jeziorności w poszczególnych bazach danych w siatce 10x10 km

- Ryc. 4.9. Przykładowe jeziora o wskaźniku RLB <1;1,5>
- Ryc. 4.10. Przykładowe jeziora o wskaźniku RLB (1,5;2>
- Ryc. 4.11. Przykładowe jeziora o wskaźniku RLB (2;3>
- Ryc. 4.12. Przykładowe jeziora o wskaźniku RLB > 3
- Ryc. 5.1. Schemat generowania atrybutów jakościowych
- Ryc. 5.2. Schemat generowania atrybutów ilościowych
- Ryc. 5.3. Brak integralności obiektu jeziora w VML2 na przykładzie Jeziora Łódzko-Dymaczewskiego
- Ryc. 5.4. Brak zgodności topologicznej między warstwą klasoużytków EGiB (kolor zielony) a wydzieleniami kompleksów glebowo-rolniczych MGR (wieloboki z wypełnieniem)
- Ryc. 5.5. Przykłady akumulacji wybranych atrybutów w źródłowych bazach danych
- Ryc. 5.6. Przykład jednostek przestrzennych generowania danych atrybutowych: A – naturalne (zlewnie elementarne), B – sztuczne nieregularne (jednostki administracyjne), C – sztuczne regularne (siatka kwadratów)
- Ryc. 5.7. Przykładowe cechy środowiska przyrodniczego wygenerowane w regularnej siatce kwadratów z bazy CLC (A), MHP (B), DTED2 (C) i MHgP (D)
- Ryc. 5.8. Klasyfikacja obszaru badań na podstawie przetworzonych do siatki kwadratów danych z baz CLC, MHP i DTED2
- Ryc. 5.9. Fragmenty mapy tematycznych prezentujące ukształtowanie powierzchni terenu, budowę geologiczną i użytkowanie terenu
- Ryc. 5.10. Fragmenty mapy tematycznych prezentujące oceny przydatności terenów do zainwestowania
- Ryc. 5.11. Fragmenty mapy dotyczącej syntetycznej oceny przydatności terenów do zainwestowania
- Ryc. 5.12. Składowe statycznego i dynamicznego potencjału informacyjnego danych przestrzennych

## Spis tabel

- Tabela 1.1. Zestawienie wybranych baz danych przestrzennych
- Tabela 2.1. Zestawienie skal pomiarowych (poziomów pomiarowych, poziomów ujęcia atrybutów) według wybranych autorów
- Tabela 2.2. Podtypy danych atrybutowych w wybranych aplikacjach bazodanowych oraz geoinformacyjnych
- Tabela 3.1. Podział środowiska przyrodniczego na geokomponenty
- Tabela 3.2. Przykładowe 30 grup warstw charakteryzujących wybrane elementy geokomponentów
- Tabela 4.1. System wstępnej oceny jakości cyfrowych baz danych
- Tabela 4.2. Zestawienie wybranych grup warstw baz danych przestrzennych z zastosowaniem systemu wstępnej oceny jakości opracowań numerycznych
- Tabela 4.3. Wartości wskaźnika gęstości sieci hydrograficznej dla poszczególnych warstw tematycznych

- Tabela 4.4. Wartości wskaźnika kształtu dla testowych odcinków cieków
- Tabela 4.5. Wskaźnik błędu lokalizacji cieków w poszczególnych bazach względem wzorca dla testowych odcinków cieków
- Tabela 4.6. Uogólnione wartości wskaźnika jeziorności rozpatrywanego w siatce 10 x 10 km
- Tabela 4.7. Wskaźniki rozwinięcia linii brzegowej (RLB) dla wybranych jezior woj. wielkopolskiego w bazach MHP, MPHP i VML2
- Tabela 5.1. Propozycja klasyfikacji poziomów ujęcia atrybutów
- Tabela 5.2. Zestawienie cząstkowych wartości potencjału informacyjnego dla warstw rzek baz danych MHP, MPHP i VML2
- Tabela 5.3. Potencjał informacyjny warstw tematycznych rzek dla baz danych MHP, MPHP i VML2
- Tabela 5.4. Zestawienie cząstkowych wartości potencjału informacyjnego dla warstw jezior baz danych MHP, MPHP i VML2
- Tabela 5.5. Potencjał informacyjny warstw tematycznych jezior dla baz danych MHP, MPHP i VML2
- Tabela 5.6. Wstępna ocena jakości dla wybranych grup tematycznych współczesnych baz danych przestrzennych
- Tabela 5.7. Przykład opisu funkcji akumulacyjnej dla dynamicznego potencjału informacyjnego dla wód powierzchniowych (zgodnie z tab. 5.1.)
- Tabela 5.8. Zestawienie operacji na atrybutach właściwych do określania funkcji analitycznej dynamicznego potencjału informacyjnego
- Tabela 5.9. Przykład opisu funkcji analitycznej dla dynamicznego potencjału informacyjnego dla wód powierzchniowych (zgodnie z tab. 5.8.)