mgr Lucyna KRYLA
Biuro Hydrograficzne Marynarki Wojennej

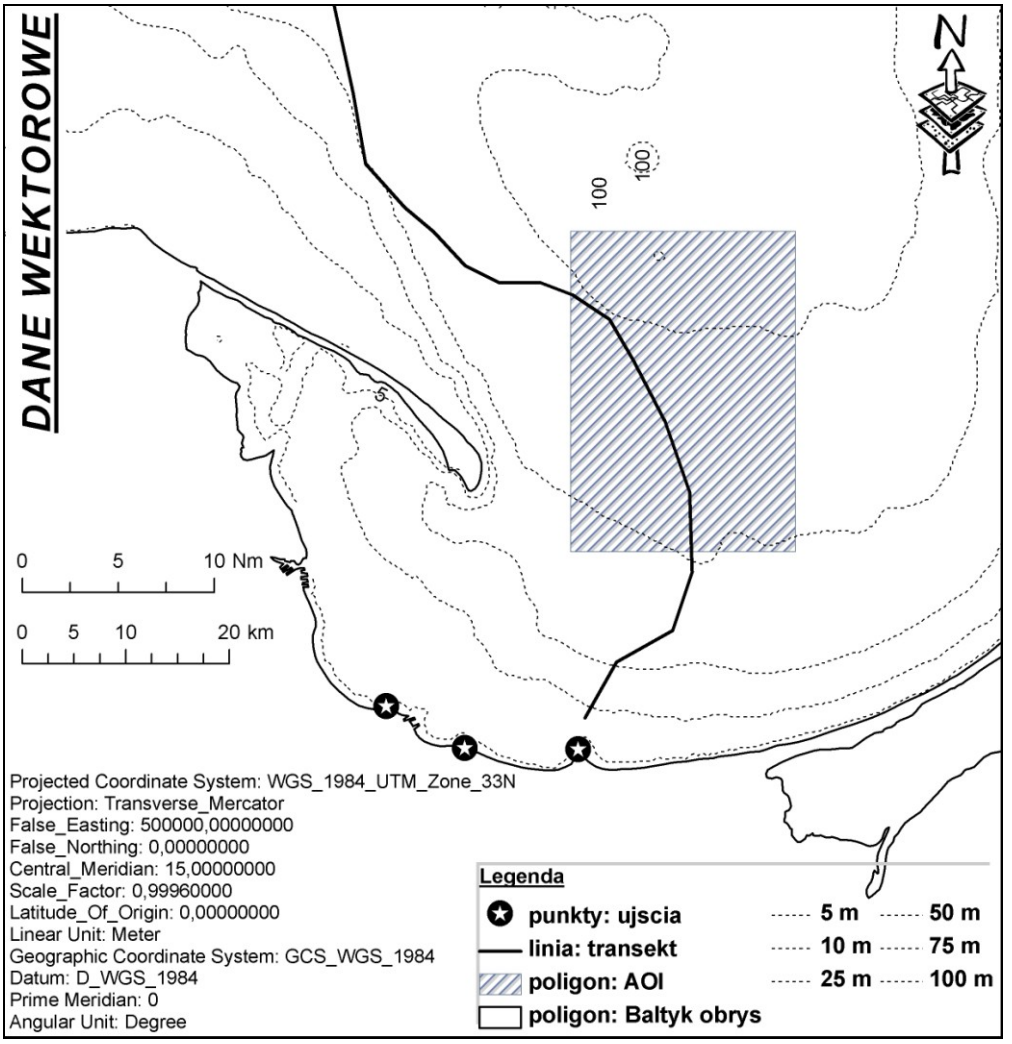
GEOGRAFICZNE SYSTEMY INFORMACYJNE – ZASTOSOWANIE FUNKCJI ANALIZY WEKTOROWEJ

Wstęp

Przez szereg wieków postać papierowa map była jedyną formą przedstawiania informacji geograficznej. Rozwój nowych technologii w przeciągu ostatnich 30 lat całkowicie zrewolucjonizował jednak proces ich powstawania, doprowadzając do wypracowania numerycznej reprezentacji w postaci map elektronicznych. Taka nowa forma danych przestrzennych daje wiele możliwości analizy zarówno przy wykorzystaniu funkcji analizy danych wektorowych, jak i rastrowych.

Model danych wektorowych zbudowany jest w oparciu o dane dyskretne: punkty, linie i poligony (Rys.1). Model danych rastrowych skonstruowany został w taki sposób, by najpełniej opisywać powierzchnie o strukturze ciągłej i zbudowany jest w oparciu o komórki zwane pikselami (Rys.2). W przypadku map nawigacyjnych przewaga modelu wektorowego nad rastrowym jest niezaprzeczalna i wynika głównie z większych możliwości precyzyjnego odzwierciedlenia rzeczywistości poprzez zapewnienie lepszej dokładności danych i pełniejszej informacji opisowej (nadawanie atrybutów) w procesie kodowania.

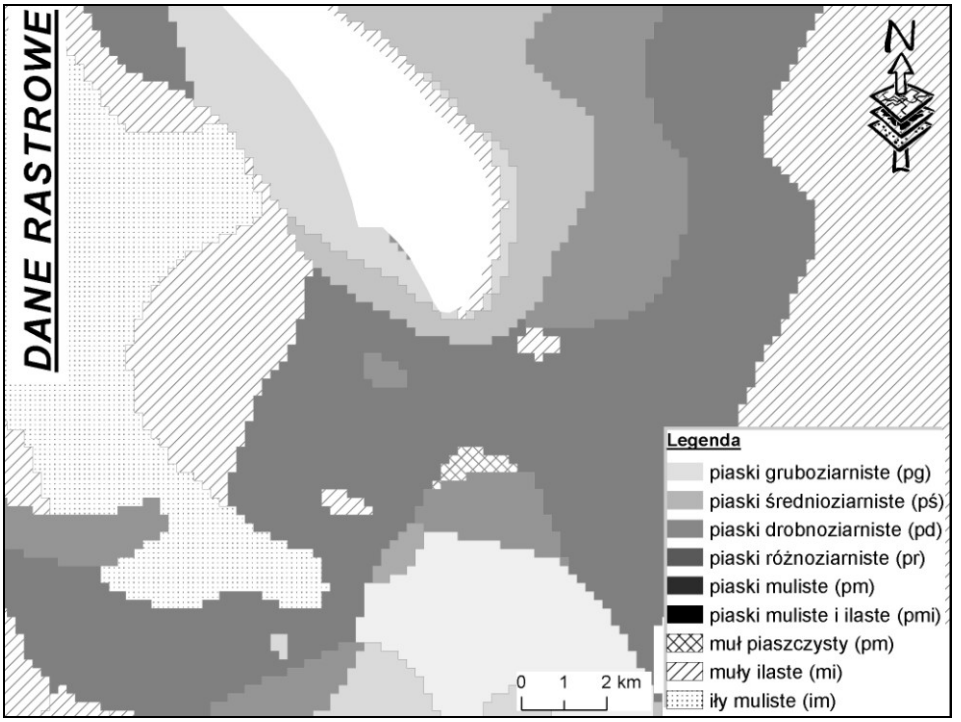
Niniejszy artykuł ma na celu przedstawienie szeregu funkcji analizy danych wektorowych w aspekcie ich potencjalnego wykorzystania w produkcji map tematycznych i specjalnych na potrzeby MW. Dla pełniejszego przedstawienia zagadnień zostaną one pokazane na wybranych przykładach. Cała analiza danych wykonana została w programie ArcGIS firmy ESRI. Zbiory danych pochodzą z zestawów ćwiczeń *Zastosowanie GIS w oceanografii*. Schematy zadań opracowane zostały przez dr. hab. Jacka URBAŃSKIEGO, uzupełnione i zmienione przez mgr Lucynę KRYLA.



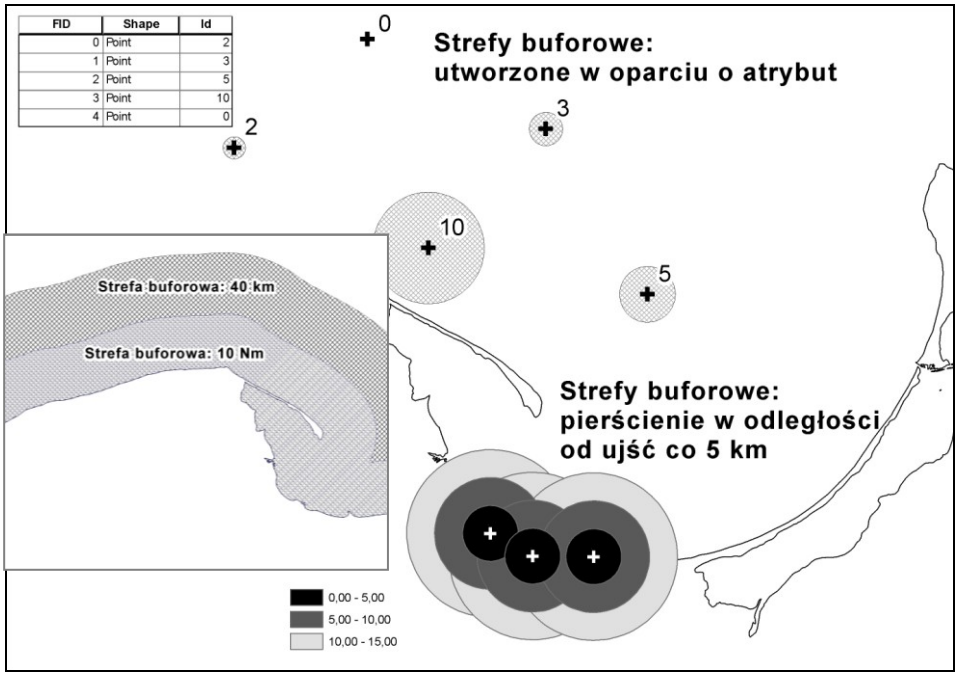
Rys. 1. Model danych wektorowych.

1. Buforowanie i funkcja wycinania (CLIP).

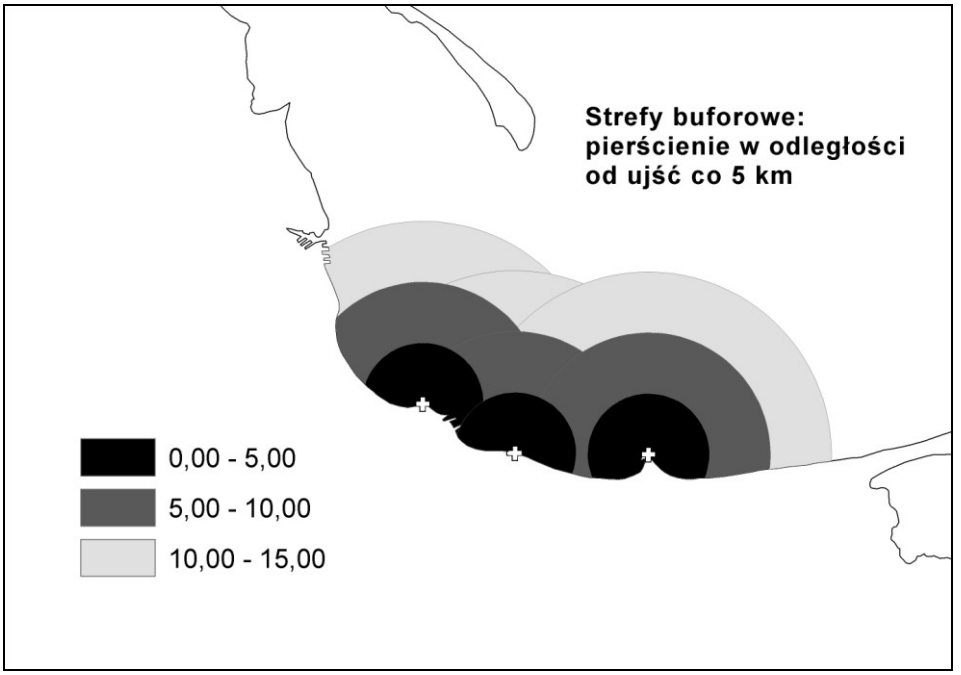
Załóżmy, że chcielibyśmy utworzyć obszar zabroniony dla żeglugi w odległości 2 mil morskich od rejonu bądź punktu niebezpiecznego lub wyznaczyć statystyki, np. średnią głębokość, liczbę wraków, w danej odległości od punktu, w danym rejonie lub na danej trasie. Kluczem do rozwiązana tego problemu byłoby utworzenie tzw. stref buforowych. Bufory mogą być tworzone zarówno dla punktów, jak i linii, czy poligonów, mogą być skalowane w zależności od wartości atrybutów poszczególnych obiektów, tworzone wewnątrz lub na zewnątrz poligonów.



Rys. 2. Model danych rastrowych.



Rys. 3. Tworzenie stref buforowych.



Rys. 4. Zastosowanie funkcji wycinania.

Na rysunku 3 pokazane są możliwości ich tworzenia:

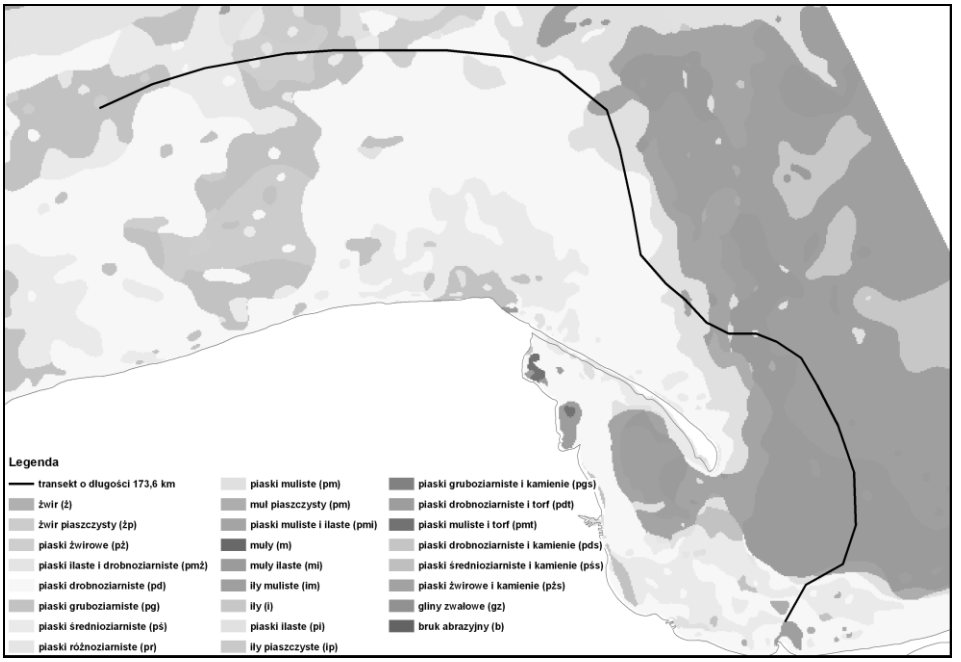
- bufor w odległości 10 mil morskich od brzegu,
- bufor w odległości 40 km od brzegu,
- strefy buforowe w oparciu o wartość atrybutu z tablicy,
- strefy buforowe w postaci szeregu pierścieni utworzone w odległościach 0-5 km, 5-10 km, 10-15 km od ujść rzecznych.

Przy tworzeniu stref buforowych dla punktów w ujściach zauważyć należy, iż strefy te znajdują się zarówno na lądzie, jak i na morzu. Pojawia się tutaj konieczność usunięcia lądowej części poligonów, co jest możliwe poprzez zastosowanie funkcji wycinania (CLIP). Aby móc wykorzystać tą funkcję, należy posiadać obszar – poligon, którym dokona się wyodrębnienia obiektów wewnątrz niego, którymi mogą to być dowolne dane typu wektorowego. Zastosowanie tej funkcji przedstawia rysunek 4 – jako CLIP wykorzystano *shapefile* obszar Morza Bałtyckiego.

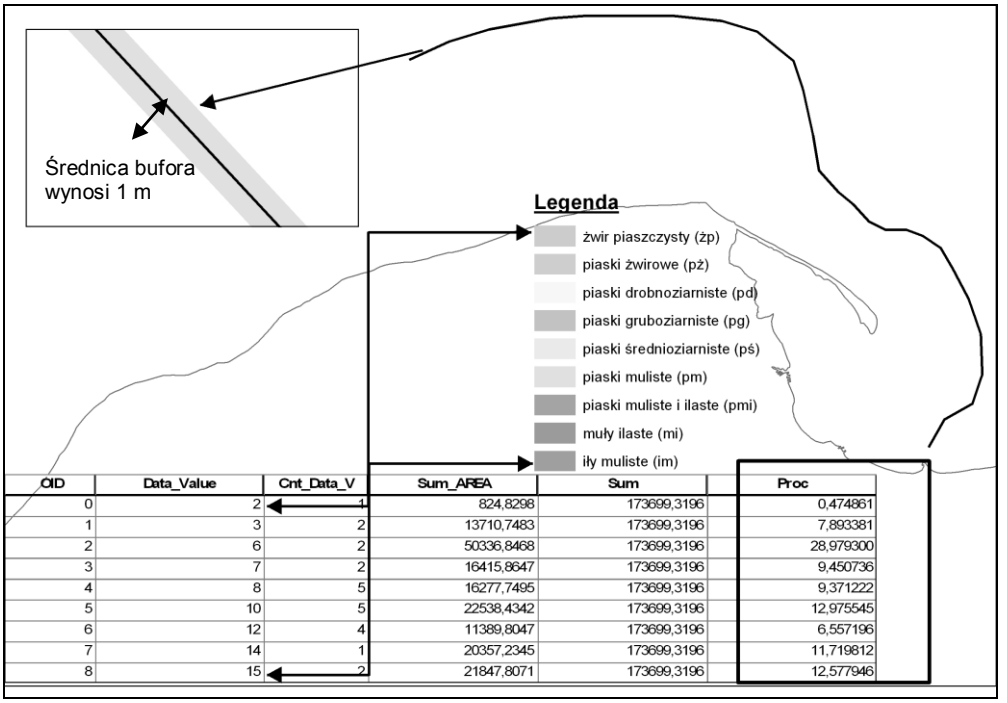
2. Buforowanie, funkcja łączenia (UNION), praca na tablicach atrybutowych

Nieocenioną zaletą funkcji analitycznych GIS jest możliwość ich łączenia w celu otrzymania żądanej informacji. Ciekawy przykład stanowi zadanie obliczenia procentowego udziału osadów poszczególnych typów na zadanym transekcie. Do wykonania tego zadania potrzebna jest mapa rastrowa lub wektorowa osadów w rejonie zainteresowania oraz sam przebieg trasy (Rys.5).

Pierwszym problemem pojawiającym się w trakcie realizacji zadania jest: jak przy wykorzystaniu dostępnych funkcji policzyć udział procentowy osadów wzdłuż linii? Jednym z podejść jest utworzenie niewielkiego bufora wokół linii – tu zastosowano promień 0,5 m, a następnie zastosowanie funkcji nakładania – łączenia (UNION). W rezultacie da ona jedną warstwę typu poligon, w której atrybuty będą zachowane z obu warstw wejściowych – w tym przypadku, w tablicy atrybutowej otrzymamy informację, czy w danym miejscu osadów przebiegał metrowy bufor transektu, a jeżeli tak, to jaki osad znajdował się na danym odcinku transektu. Dalszym krokiem będzie wyodrębnienie tylko tych poligonów, które posiadają pozytywny identyfikator dla przebiegu trasy (eksport do nowej warstwy) oraz policzenie powierzchni poszczególnych poligonów (ciągły bufor wokół transektu poprzecinany został granicami nałożonych, zmieniających się osadów). Ponieważ każdy osad posiada inny atrybut w tablicy, a ilość rekordów dla poszczególnych osadów jest większa od 1 (1 – 5), należy przeprowadzić tzw. sumaryzację. Funkcja sumaryzacji działa w ten sposób, że traktuje jedną z kolumn jako definicję klasy (identyfikator osadu), licząc dla tej klasy statystyki z innych, wybranych kolumn (suma z pola powierzchni w tym przypadku). Dane statystyczne umieszczone zostają w tablicy *.dbf (Rys. 6). Obliczenie udziału procentowego to podzielenie otrzymanych podsum dla poszczególnych typów osadów przez sumę całkowitą powierzchni. Ostatecznie należy połączyć kod osadu z nazwą osadu. Na zadanym transekcie najwięcej znajduje się piasków drobnoziarnistych (ok. 29%), najmniej żwiru piaszczystego (ok. 0,5%).



Rys. 5. Usytuowanie transektu na tle mapy wektorowej osadów.

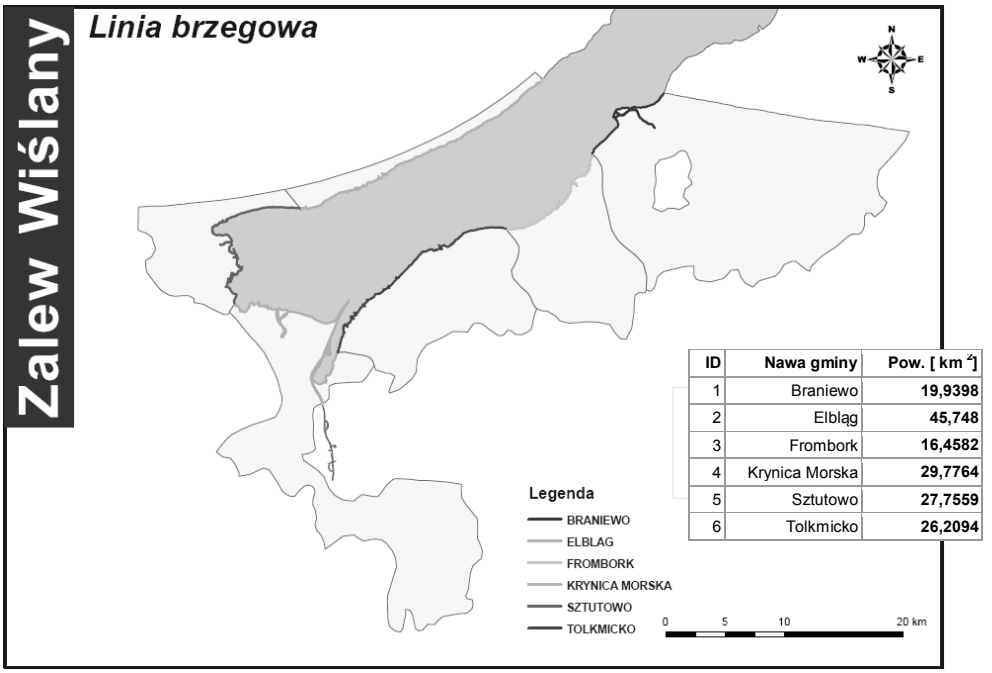


OID	Data_Value	Cnt_Data_V	Sum_AREA	Sum	Proc
0	2	1	824,8298	173699,3196	0,474861
1	3	2	13710,7483	173699,3196	7,893381
2	6	2	50336,8468	173699,3196	28,979300
3	7	2	16415,8647	173699,3196	9,450736
4	8	5	16277,7495	173699,3196	9,371222
5	10	5	22538,4342	173699,3196	12,975545
6	12	4	11389,8047	173699,3196	6,557196
7	14	1	20357,2345	173699,3196	11,719812
8	15	2	21847,8071	173699,3196	12,577946

Rys. 6. Wykaz osadów występujących na transekcie wraz z ich procentowym udziałem.

3. Funkcja przecinania (INTERSECT), topologia warstw, praca na tablicach atrybutowych

Oprócz dwóch wcześniej wymienionych funkcji nakładania (CLIP, UNION) istnieje jeszcze jedna, nosząca nazwę przecinania (INTERSECT). W efekcie zadziałania tej funkcji na dwóch (bądź większej ilości) warstwach wektorowych uzyskuje się warstwę, w której zachowane są jedynie te elementy, które wystąpiły we wszystkich danych wejściowych – przecięły się. Przykładem jej zastosowania jest wyliczenie przybliżonej długości linii brzegowej w rejonie poszczególnych gmin Zalewu Wiślanego mając do dyspozycji jedynie obrys zalewu i obrys gmin. Niestety, po zadziałaniu funkcji INTERSECT może okazać się, że granice warstw: obrys zalewu i obrys gmin, nie zachowują przestrzennych relacji – nie pokrywają się. Jest to definiowane jako błąd topologii danych. Rozwiązaniem tymczasowym, nie ingerującym w topologię warstw, będzie wprowadzenie pewnej tolerancji niedokładności danych, w obrębie której uznamy, że dane ulegają nałożeniu (*cluster tolerance* = 3 metry).



ID	Nawa gminy	Pow. [km ^{2}]
1	Braniewo	19,9398
2	Elbląg	45,748
3	Frombork	16,4582
4	Krynica Morska	29,7764
5	Sztutowo	27,7559
6	Tolkmicko	26,2094

Rys. 7. Przykład zastosowania funkcji INTERSECT.

Należy mieć jednak na uwadze, iż takie rozwiązanie wprowadza dodatkowy błąd do obliczeń. Jako typ danych wyjściowych w INTERSECT należy wybrać linię. Ułatwieniem w rozwiązaniu tego zagadnienia jest fakt, że funkcja przecinania zachowa atrybuty warstw nakładanych, a więc także nazwy gmin. Ostatnim etapem zadania będzie sumaryzacja po klasie: nazwa gminy, z sumy długości linii wyliczonych w uprzednio dodanej kolumnie (Rys. 7).

4. Wyszukiwanie obiektów według wartości atrybutowych i na podstawie relacji przestrzennych, transformacja układu współrzędnych, rejestracja danych rastrowych i wektorowych, digityzacja – tworzenie warstw wektorowych, łączenie zbiorów danych na podstawie relacji przestrzennych (JOIN), interpolacja, filtrowanie, rasteryzacja, maskowanie (algebra map).

Jednym ze sposobów pozyskiwania danych jest skanowanie postaci papierowej, a następnie digityzacja obiektów ze skanu. Tak otrzymane dane mogą być poddane dalszej transformacji z wykorzystaniem metod analizy danych wektorowych i/lub rastrowych w celu uzyskania szukanej informacji. Zobrazowanie procesu digityzacji przedstawiono na przykładzie tworzenia mapy batymetrycznej dla Jeziora Stężyckiego. Do wykonania tego zadania dostarczone zostały dane:

- obrysy jezior Pomorza – dane wektorowe *.shp, układ współrzędny danych GCS WGS84,
- plan batymetryczny Jeziora Stężyckiego – postać *.jpg (Rys. 8).

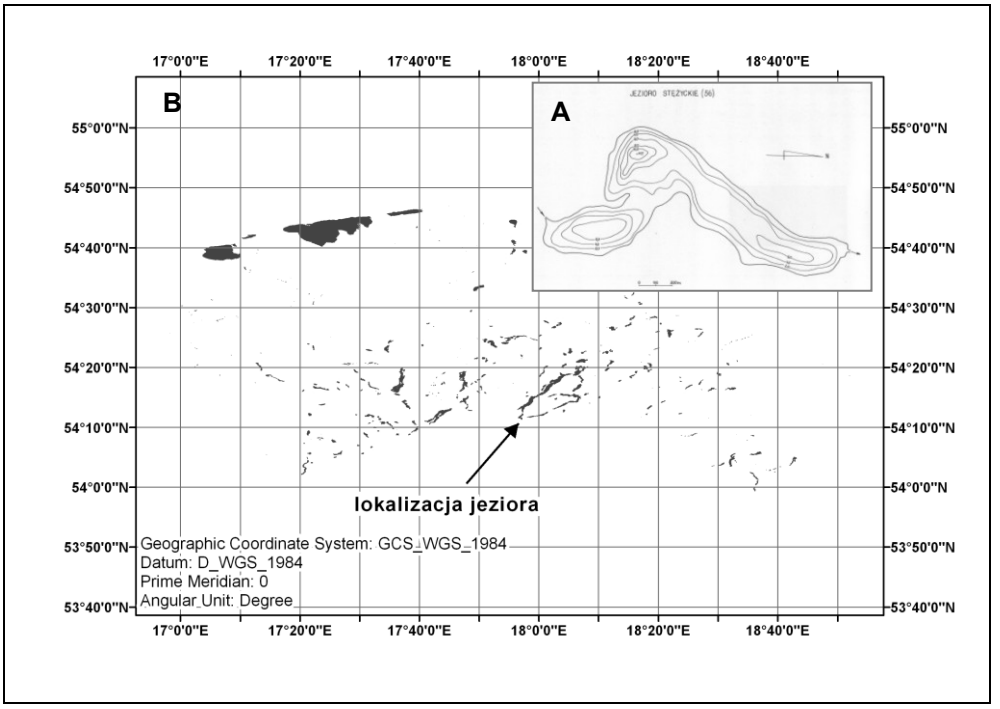
Produktem analizy ma być rastrowa mapa batymetryczną Jeziora Stężyckiego przedstawiona w układzie współrzędnych UTM 34N. Dodatkowym zadaniem jest obliczenie średniej głębokości w jeziorze, jego powierzchni i objętości wody.

Punkt 1. **Wyszukiwanie obiektów według wartości atrybutowych.**

Dane wektorowe do wykonania tego zadania zawierają informację nie tylko o wybranym jeziorze, ale także o innych jeziorach Pomorza. Pierwszym krokiem jest zatem znalezienie Jeziora Stężyckiego. Do danych wektorowych typu *shapefile* dołączona jest tzw. tablica atrybutowa, w obrębie której możliwe jest wyszukiwanie danych po nadanych wcześniej atrybutach. Tym razem należy wykorzystać informację o nazwie jeziora.

Punkt 2. **Eksport danych**

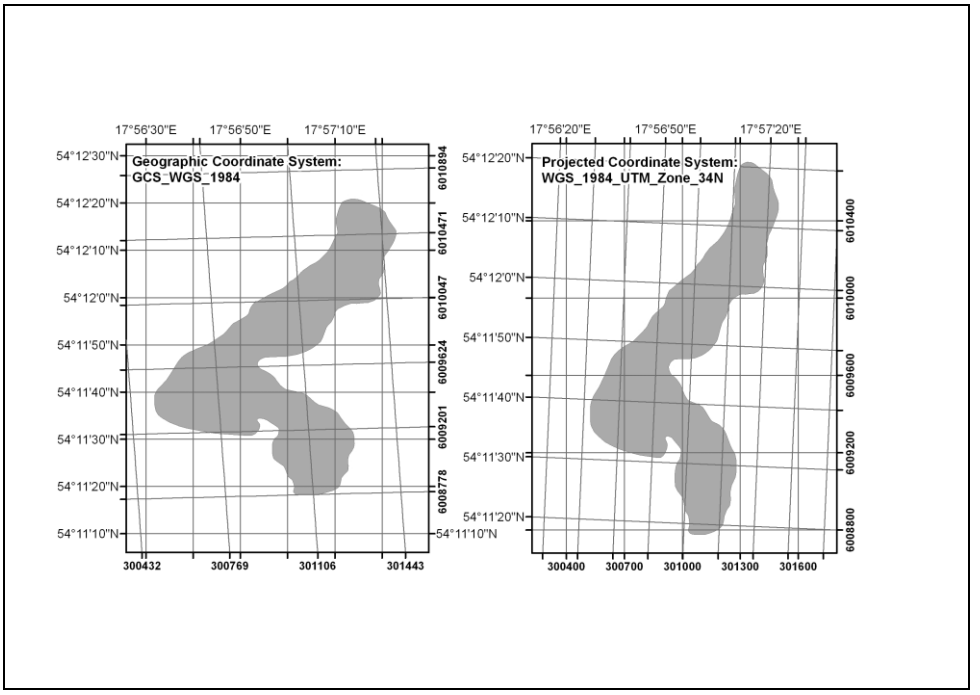
Poszukiwany obiekt/rekord po zaznaczeniu w tablicy atrybutowej można wyeksportować do nowej warstwy. Krok ten jest szczególnie wskazany w przypadku pracy na bardzo dużej liczbie danych (skraca czas dalszej analizy danych unikając transformacji niepotrzebnych danych).



Rys. 8. Schemat batymetryczny Jeziora Stęzyckiego – JPG (A) i obrysy jezior Pomorza – dane wektorowe (B).

Punkt 3. **Transformacja układu współrzędnych**

Mając na uwadze, że mapa końcowa wykonana ma być w UTM 34N, w przypadku tego zadania jest to optymalny moment na transformację układu współrzędnych. We większości programów GIS proces konwersji układów przebiega bez większego udziału użytkownika przy wykorzystaniu predefiniowanych geograficznych układów współrzędnych i odwzorowań kartograficznych (Rys. 9).



Rys. 9 Transformacja układu współrzędnych.

Punkt 4. **Rejestracja danych rastrowych w układzie współrzędnych**

Mając do dyspozycji obrys jeziora i jego plan batymetryczny można dokonać dowiązania grafiki do danych wektorowych – rejestracji w układzie współrzędnych. Jest to etap niezbędny przed wykonaniem kolejnego kroku, tj. przed digityzacją izobat. Istnieją różne metody transformacji obrazu w celu jego jak najlepszego dopasowania.

W tym przypadku zastosowano najprostszą postać równania dopasowującego współrzędne tzw. punktów kontrolnych, wskazujących punkty wzajemnie sobie odpowiadające na JPG i obrysie (strzałki na Rys. 10 A pokazują umiejscowienie punktów kontrolnych), wielomian pierwszego rzędu (transformacja polegająca na przesunięciu i obrocie skanu). Po wskazaniu punktów kontrolnych można przeprowadzić georeferencję lub rektyfikację zdjęcia. Pierwszy z tych procesów spowoduje utworzenie nowego pliku zawierającego informację

o ustawieniu zdjęcia w przestrzeni, nie zmieniając samego zdjęcia, drugi będzie skutkował przeliczeniem współrzędnych każdego piksela rejestrowanego rastra na nowe i utworzeniem GeoTIFF (Rys. 10 B, zastosowano georeferencję).

Punkt 5. **Digityzacja – tworzenie warstw wektorowych, edycja danych**

Posiadając dowiązany JPG można dokonać procesu digityzacji izobat. Do sczytanych linii należy w tablicy atrybutowej umieścić informację o wartości głębokości – dokonać edycji danych (Rys. 10 C).

Punkt 6. **Rejestracja danych wektorowych w układzie współrzędnych**

Niestety, może się okazać, że dokładność danych wektorowych utworzonych w powyższy sposób nie jest wystarczająca i należy ponownie dokonać dowiązania danych, tym razem bazując już na danych wektorowych. W tym przypadku spośród kilku typów transformacji wybrana została metoda elastycznego dopasowania (rubbersheeting), która polega na lokalnym dopasowaniu wskazanych punktów kontrolnych (proces tworzenia punktów kontrolnych przedstawiono na Rys. 10 D, lokalizację wszystkich punktów dociągania pokazuje Rys. 11 E).

Punkt 7. **Konwersja danych typu linia-punkt**

Kolejnym etapem pracy jest interpolacja danych. Z uwagi jednak na fakt, że standardowe metody interpolacji bazują na danych punktowych, należy przeprowadzić zamianę izobat na zbiory punktowe. Zastosowano tutaj skrypt konwertujący linie do punktów, napisany w języku Visual Basic (uruchamianie makro odbywa się bezpośrednio spod aplikacji ArcGIS, pod tą aplikacją jest także możliwe napisanie skryptu). Po zadziałaniu skryptu otrzymamy plik punktowy z zaznaczonych warstw *.shp typu polilinia (Rys. 11 F). W danych punktowych nie będą jednak zachowane atrybuty warstwy izobat – wartości głębokości. Kolejne dwa punkty zadania (8a i 8b) opisują różne scenariusze rozwiązania powyższego problemu.

Punkt 8a. **Wyszukiwanie obiektów na podstawie relacji przestrzennych**

Jeden ze sposobów przypisania atrybutów (dane głębokości) jednej warstwy danych (izobaty) do innej (punkty) bazuje na selekcji obiektów na podstawie wzajemnych relacji przestrzennych. Proces ten przebiega następująco: zaznaczamy izobaty o danej głębokości (np. izobaty o głębokości 2 m), przeprowadzamy selekcję danych punktowych na podstawie lokalizacji z wykorzystaniem zaznaczonej izobaty (polecenie: znajdź punkty które leżą na zaznaczonej izobacie 2 m), do zaznaczonych rekordów w tablicy punktów przypisujemy atrybut 2 (kolumna głębokości). W analogiczny sposób przeprowadza się wyliczenie wartości głębokości dla pozostałych punktów (zaznaczenie kolejnej izobaty, wyszukanie punktów leżących na niej, przypisanie wartości głębokości dla zaznaczonych punktów; Rys. 11 G).

Punkt 8b. **Łączenie zbiorów danych na podstawie relacji przestrzennych (JOIN)**

Szybszym sposobem przypisania atrybutów jednej warstwy danych do innej jest zastosowanie funkcji łączenia danych na podstawie wzajemnych relacji przestrzennych (JOIN). W efekcie zastosowania tej funkcji na danych punktowych otrzymuje się nową warstwę danych punktowych z dołączoną informacją do tablicy atrybutowej ze wskazanej drugiej warstwy (izobaty). Funkcja łączenia do każdego punktu przypisze pełną informację (wartość głębokości) o najbliżej leżącym obiekcie z drugiej warstwy (izobaty; Rys. 11 G).

Punkt 9. **Interpolacja (dane wektorowe: punkty -> dane rastrowe)**

Mając do dyspozycji zbiór punktów z wartościami głębokości można dokonać procesu interpolacji, tj. obliczenia wartości głębokości w dowolnym punkcie przestrzeni, w której zawarte są dane punktowe. Podstawowym założeniem wszystkich metod interpolacji jest stwierdzenie, że bliżej znajdujące się punkty mają większy wpływ na wartość wyznaczaną, niż te, które znajdują się dalej. W realizacji tego podpunktu wykorzystano funkcję tzw. krigingu (Rys. 11 H).

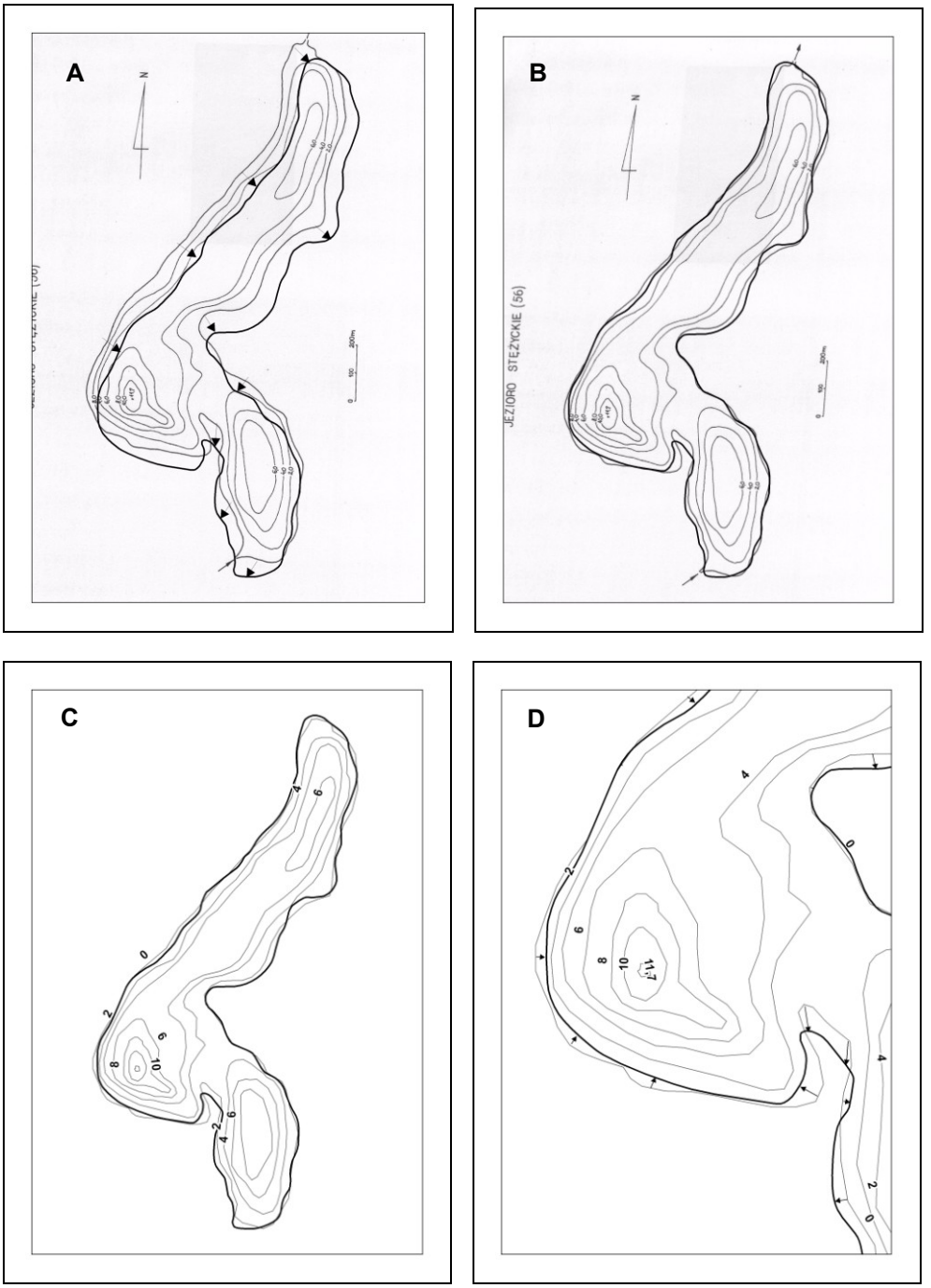
Punkt 10. **Filtrowanie (dane rastrowe)**

Dane rastrowe otrzymane po procesie interpolacji można poddać filtracji (np. filtr średniej z oknem 5 na 5), prowadzącej do wygładzenia powierzchni (Rys. 12 I).

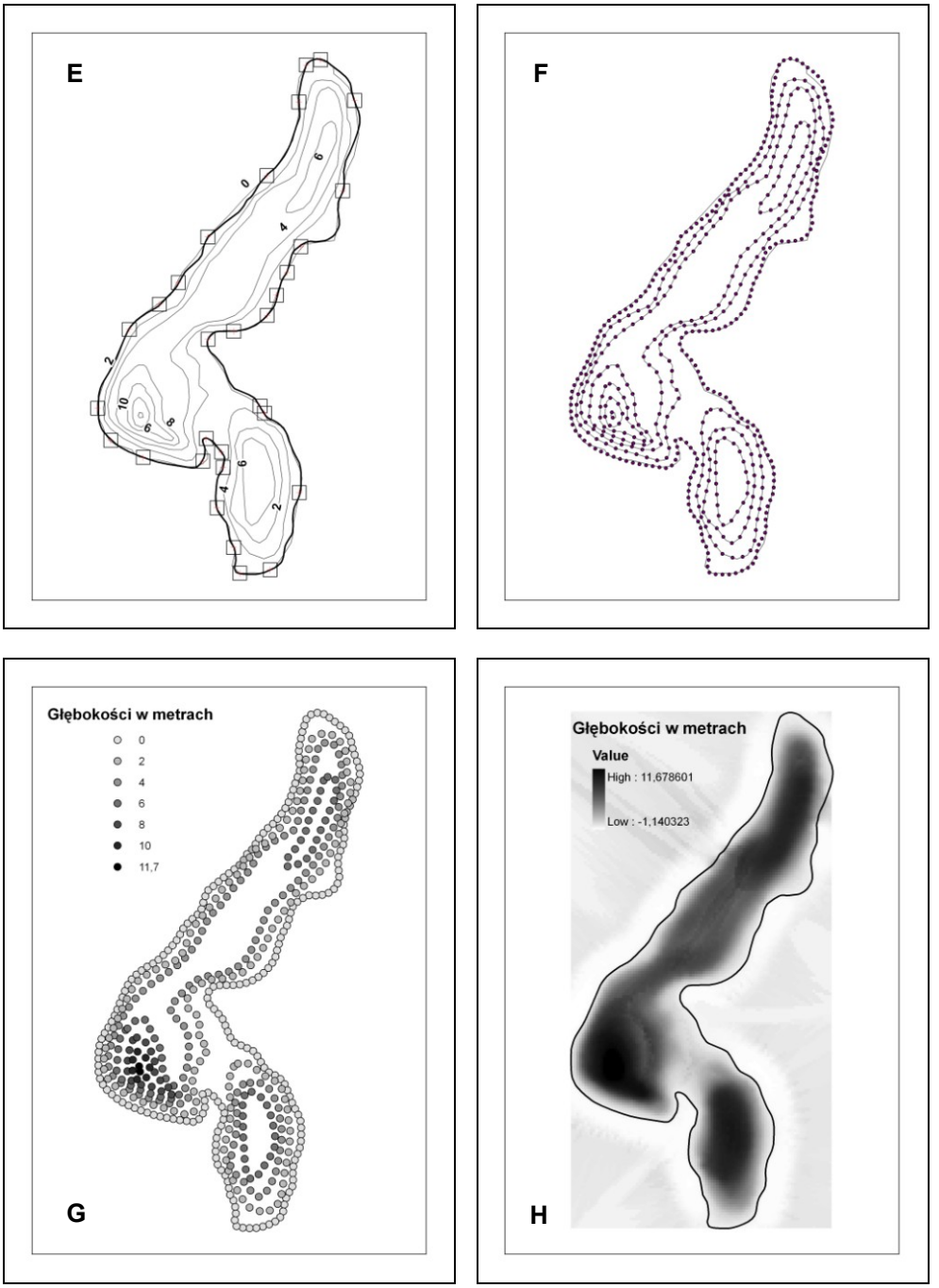
Punkt 11. **Rasteryzacja (dane wektorowe: poligon -> dane rastrowe), maskowanie (algebra map)**

Z uwagi na fakt, że wartości głębokości zostały wyznaczone zarówno na obszarze wody, jak i w części lądowej, należy dokonać usunięcia – wymaskowania obszaru wychodzącego poza obrys jeziora. Utworzenie maski, definiującej obszar zainteresowania oraz obszar podlegający usunięciu, w tym przypadku polega na rasteryzacji danych wektorowych typu poligon: obrys Jeziora Stężyckiego z wykorzystaniem kolumny zawierającej wartość 0 (rastrowa maska będzie posiadała wartość 0 w obszarze zajmowanym przez jezioro i wartość NoData na obszarze lądu; Rys. 12 J).

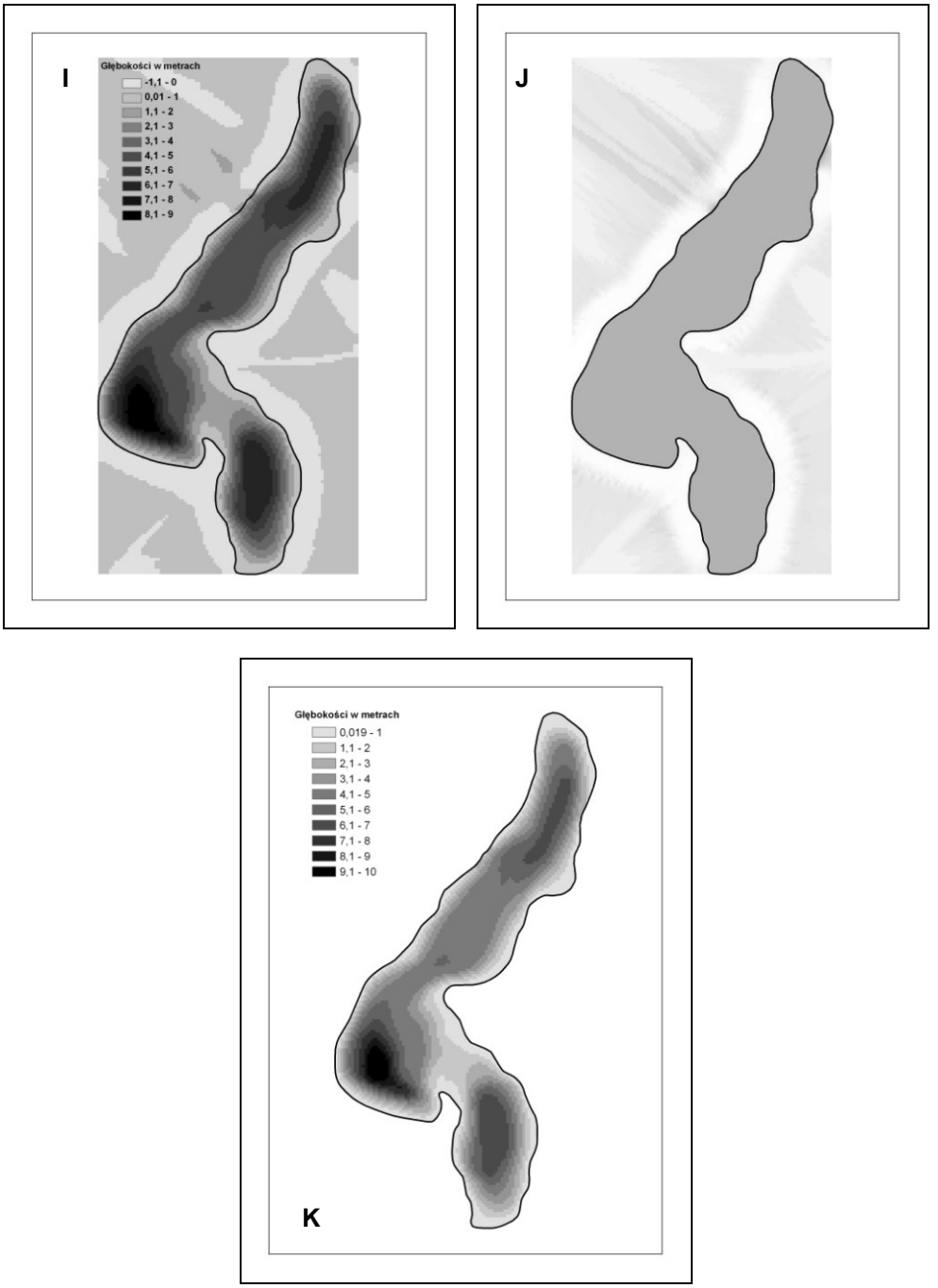
Po zastosowaniu algebry map, dodaniu do siebie mapy po filtracji i maski, otrzymana zostanie rastrowa mapa batymetryczna Jeziora Stężyckiego (komórki nowego rastra wyliczone zostaną następująco: głębokość + 0 = głębokość, głębokość + NoData = NoData; Rys. 12 K).



Rys. 10. Rejestracja danych rastrowych (A, B), digityzacja izobat (C) i rejestracja danych wektorowych (D).



Rys. 11. Rejestracja danych wektorowych (E), konwersja izobat do punktów (F), wizualizacja danych wektorowych (G) i interpolacja danych (H).



Rys. 12. Filtracja danych (I), tworzenie maski (J) i maskowanie danych, wizualizacja danych rastrowych (K).

Punkt 12. **Obliczenie statystyk**

Po przeprowadzeniu powyższych etapów z danych wejściowych: wektorowy obrys jezior Pomorza i graficzny schemat batymetrii Jeziora Stężyckiego, otrzymano mapę batymetryczną Jeziora Stężyckieo. W celu uzyskania żądanej informacji mapa ta może zostać poddana dalszej analizie. Zakładając, że użytkownika interesują statystyki: wartość średnia głębokości, powierzchnia jeziora i objętość wody w jeziorze, można zastosować wbudowaną funkcję programu ArcGIS *Spatial analyst/Zonal statistics*, która dla całego obszaru jeziora wyliczy z mapy batymetrycznej wartości głębokości: minimalną, maksymalną, średnią, odchylenie standardowe, sumę, powierzchnię obszaru analizy i inne. Aby uzyskać wartość objętości wody w jeziorze należy powierzchnię jeziora przemnożyć przez jego średnią głębokość. Alternatywną metodą będzie przemnożenie mapy batymetrycznej przez powierzchnię pojedynczego piksela, wówczas do danej komórki rastra przypisana zostanie wartość znajdującej się nad nim objętości wody, a następnie znalezienie sumy wartości pikseli.

Podsumowanie

Powyższe przykłady ilustrujące zastosowanie podstawowych funkcji analizy wektorowej pokazują ogromne możliwości edycji i transformacji danych w ramach Geograficznych Systemów Informacyjnych. Analiza danych może przebiegać zarówno pod kątem wykorzystania informacji o lokalizacji obiektów, wzajemnych relacji przestrzennych obiektów/warstw, jak i z wykorzystaniem właściwości danych – atrybutów. Tworzenie i przekształcanie danych z jednoczesnym zachowaniem topologii warstw, rejestracja i digityzacja danych, różnorodność tworzenia buforów i możliwości łączenia wszystkich funkcji analizy w celu uzyskania poszukiwanej informacji, to zalety wykorzystania GIS mogące mieć duże zastosowanie w tworzeniu map bazujących na obiektach zawartych w komórkach ENC oraz warstwach AML.

Literatura:

1. Urbański J., Zrozumieć GIS. Analiza informacji przestrzennej, 1997, PWN
2. Longley P.A., i in., GIS. Teoria i praktyka, 2006, PWN
3. www.esri.com