

Zbigniew Sitek

## **ELEMENTY PROJEKTOWANIA ZINTEGROWANEGO GIS DLA POTRZEB MONITORINGU ŚRODOWISKA<sup>1</sup>**

### **1. Wstęp**

W monitoringu środowiska korzysta się z różnych źródeł danych takich jak mapy topograficzne i tematyczne, pomiary terenowe dostarczające wskaźników zanieczyszczeń powietrza, gleb i wody oraz z fotogrametrii i teledetekcji dostarczających informacji za pośrednictwem kamer fotograficznych, skanerów i radiometrów. Integracja tych danych oraz cyfrowego (wysokościowego) modelu terenu z danymi geograficznego systemu informacyjnego (GIS) zależy w głównej mierze od wyposażenia, oprzyrządowania i oprogramowania komputerowego, a także od zaplecza danej instytucji, która zajmuje się monitoringiem (Sitek 1994). Zależy również od tego jak zaprojektuje się taki GIS dla potrzeb badania skażenia środowiska. Dlatego do zadań projektu badawczego KBN pt. "Monitoring środowiska z wykorzystaniem GIS" włączono również rozważania nad zagadnieniem projektowania zintegrowanego GIS.

Bardzo wiele nazw używano i używa się nadal na określenie zintegrowanego i geodezyjnie zorientowanego systemu opracowania map cyfrowych. Należą do nich np. (Acharya, Bell 1992): System informacji o terenie (LIS- Land Information System), Geograficzny System Informacyjny (GIS-Geographic Information System), System Informacji Przestrzennej (SIS-Spatial Information System), System informacji odniesiony do terenu (LRIS - Land Related Information System), Kataster wielozadaniowy (Multipurpose Cadastre), System zapisu terenu (Land Record System), Wspomagane komputerowo systemy opracowania map (CAMS-Computer Aided Mapping Systems), Kataster nowoczesny (Modern Cadastre), Zapis terenu i system informacji o zasobach (Land Record and Resource Information System), System danych o terenie (Land Data System), Bank danych geograficznych (Geographic Data Bank), Bank danych o terenie (Land Data Bank), System zapisu informacji o terenie (Land Record Information System), Geograficzna baza danych o środowisku (Community Geographic Database), System automatycznego opracowania map (Automated Mapping System), System automatycznego zarządzania informacjami (Automated Information Management System), Geomatyka (Geomatics), Geoinformatyka (Geoinformatics), Geometronika (Geometronics), Opracowanie map i analizy z wykorzystaniem informacji przestrzennej (SIMA-Spatial Information Mapping and Analysis).

Najpopularniejsze są dwie pierwsze nazwy GIS i LIS. Największą ilość informacji zawiera oraz największy obszar działalności obejmuje geograficzny system informacyjny, gdyż oprócz informacji terenowych (katastralnych, przyrodniczych, infrastrukturalnych)

---

<sup>1</sup> Praca realizowana w ramach projektu KBN „Monitoring środowiska z wykorzystaniem GIS”.

ujmuje również informacje społeczne i ekonomiczne, a ponadto gromadzi dane statystyczne, umożliwia przeprowadzanie analiz przestrzennych oraz modelowanie procesów i zjawisk.

## **2.Rodzaje i składowe GIS**

Można wyróżnić trzy rodzaje GIS-ów: globalny, regionalny i lokalny.

Globalny GIS jest zwykle drobnoskalowy, wykorzystywany do badań kształtu Ziemi (poła grawitacyjnego), badań naukowych, do celów nawigacyjnych, kształtowania środowiska itp.

Regionalny GIS jest zakładany w średniej skali i może być użyteczny w rozwiązywaniu wielu zadań inżynierskich, planistycznych, przy wykorzystaniu zasobów naturalnych, w badaniach nauk o Ziemi oraz przy wszelkiego rodzaju studiach środowiskowych, ekologicznych itp.

Lokalny GIS należy do GIS-ów wielkoskalowych i jest wykorzystywany do opracowywania projektów planów przestrzennego zagospodarowania i rozwoju dla ograniczonej powierzchni terenu.

W podręcznikach omawiających GIS, w tym także mojego autorstwa (Sitek, 1992) podawano, że składowymi GIS są: położenie sytuacyjne, atrybuty i czas. Wydaje się dużo lepsze określenie składowych GIS podane w publikacji (Acharya i Bell 1992) stwierdzające, że współczesne GIS-y zawierają:

- składowe geograficzne lub geodezyjne,
- składowe informacyjne,
- składowe systemu,
- standardy dokładnościowe.

### **2.1. Składowe geodezyjne**

Geodezyjne składowe GIS umożliwiają przekształcanie informacji do jednoznacznie zdefiniowanego układu odniesienia a ponadto służą do kontroli dokładności rezultatów działania GIS. Do składowych geodezyjnych zalicza się:

- układ powierzchni odniesienia,
- odwzorowanie kartograficzne powierzchni na płaszczyźnie mapy,
- dane geodezyjne służące do powiązania płaszczyzny mapy z układem powierzchni odniesienia,
- skalę map i system numeracji arkuszy map.

Dla celów pomiarowych i opracowania map na świecie stosuje się około 14 głównych elipsoid odniesienia. W 1992 i 1993 r polską część sieci EUREF (European Reference Frame), włączono do europejskiej, jednolitej, zintegrowanej trójwymiarowej sieci geodezyjnej EUREF-89. Powierzchnią odniesienia nowego układu w Polsce powinna być

elipsoida GRS 80 (Geodetic Reference System 80) [Ekspertyza, 1993], która stała się elipsoidą europejskiego systemu odniesienia (European Terrestrial Reference Frame) ETRF-89 (EUREF-89).

Wybór odpowiedniej elipsoidy odniesienia odtwarzającej sytuację i wysokości wszystkich danych wpływa na dokładność informacji dostarczanych przez GIS. Ze względu na niejednorodność rozkładu danych geodezyjnych na różnych kontynentach jeden układ powierzchni odniesienia nie zapewniłby dobrego przylegania regionalnych obszarów na wszystkich kontynentach. Stąd też dla użytku GIS (globalnego, regionalnego i lokalnego) zaleca się stosowanie siedmiu powierzchni odniesienia: sześć dla sześciu kontynentów (Afryka, Azja, Australia, Europa, Ameryka Północna i Ameryka Południowa) i jednej dla pokrycia globalnego. Natomiast powierzchnią odniesienia wysokości powinna być jedna geoida dla całej Ziemi i obliczenia wysokości powinny wykorzystywać tę właśnie geoidę.

Ważną rolę w planowaniu GIS odgrywa odwzorowanie kartograficzne powierzchni na płaszczyznę (projekcja mapy). Na świecie w opracowaniu map korzysta się z około 27 różnych odwzorowań kartograficznych. Planując GIS można dla danego regionu lub obszaru dobrać optymalne odwzorowanie, w którym będą przedstawiane wszystkie mapy i zobrażenia. W mniejszych projektach trzeba się jednak dostosowywać do istniejących aktualnych map danego obszaru. Niemniej wówczas należy dokładnie sprawdzać (jeśli obszar leży na granicy pasów odwzorowanych) jakiego rodzaju zniekształcenia geometryczne powierzchni są tam wprowadzane. Nieprzystające powierzchnie mogą spowodować znaczne błędy w opracowaniach GIS-owskich. W Polsce od 1928 roku było stosowane odwzorowanie Gaussa - Krugera. Ekspertyza (1993) sugeruje, że *"odwzorowaniem kartograficznym funkcjonującym w Państwowym Systemie Współrzędnych Geodezyjnych w Polsce powinno być odwzorowanie Gaussa-Krugera w pasach 3<sup>o</sup> z odpowiednio dobranym współczynnikiem skali na południku środkowym"*. Spełnia ono warunki odwzorowania lokalnego.

Obecnie w kraju są dostępne mapy w różnych skalach ale też w różnych układach o nazwach "1942", "1965", "GUGiK 80". Biorąc pod uwagę rodzaj projektowanego GIS-u a więc skalę, aktualność treści dostępnych map, odwzorowanie tych map i zniekształcenia na granicach pasów oraz program i zadania tworzonego GIS-u, można go zakładać na bazie któregoś z w/w układów lub też jak wcześniej wspomniano dobrać optymalne odwzorowanie.

Do powiązania płaszczyzny projekcji mapy z powierzchnią odniesienia potrzebne są co najmniej 3 punkty kontrolne o znanych współrzędnych geodezyjnych. W zależności od skali mapy liczba stanowisk kontrolnych będzie się znacznie zmieniała. Np dla mapy w skali 1:1000 trzeba zakładać punkty kontrolne co 400 - 500 m. Zatem wielkoskalowy GIS będzie wymagał bardzo dużo stanowisk kontrolnych. W opracowaniu Acharya i Bella (1992) zestawiono niezbędną dla potrzeb GIS-u regionalnego liczbę punktów kontrolnych dla wszystkich kontynentów. Przy gęstości 1 punktu na 25 km<sup>2</sup> dla Europy potrzeba 1.007.177 punktów a jest dostępnych 2.385.000 z tym, że na terytorium byłego Związku Radzieckiego - tylko 200 tysięcy, a potrzeba tam dodatkowo około 700 tysięcy. Europa jest jedynym kontynentem z nadmiarem geodezyjnych stanowisk kontrolnych. Te poziome i pionowe stanowiska kontrolne umożliwiają ponadto wprowadzenie poprawnej skali, właściwe utrzymanie zależności sytuacyjnej i orientacyjnej pomiędzy obiektami, tworząc strukturę na której kompiluje się elementy mapy.

Skala map i system numeracji arkuszy map jest bardzo istotnym czynnikiem w budowie zintegrowanego GIS-u, ale zintegrowanego w takim sensie, że wszystkie instytucje

w obrębie kraju, województwa, regionu czy lokalnej strefy będą z niego korzystały. Wówczas koszt budowy takiego GIS-u jest rozłożony na te instytucje i wtedy staje się on również optymalnym narzędziem pomiarowym oraz systemem do opracowania map. Dlatego warunkiem zasadniczym takiego zintegrowanego GIS-u jest aby agencje i instytucje będące użytkownikami GIS-u w obrębie regionu lub strefy używały tego samego układu odniesienia i tych samych map podstawowych o tej samej numeracji. Wprowadzenie niektórych systemów GIS-owskich wyposażono w moduły do transformacji współrzędnych, ale wymaga to dodatkowej pracy i odpowiednich danych. Jak wiadomo każdy GIS niezależnie od tego do czego ma służyć musi być wyposażony w warstwy informacyjne lokalizujące obiekty w przestrzeni geograficznej. Powinien też zawierać wysokościowy model terenu. Zbieranie tych i innych informacji dla potrzeb jednej tylko instytucji zwiększa bardzo koszty budowy GIS-u. Dlatego przyszłość GIS-ów w znacznym stopniu zależy od integracji danych podstawowych przydatnych w różnych instytucjach i pochodzących z map podstawowych opracowanych w tym samym układzie odniesienia. Integracja danych terenowych, mapowych i obrazowych wymaga digitalizacji, umiejscowienia ich w odpowiednim układzie odniesienia, oraz rejestracji. Każdemu z tych procesów muszą towarzyszyć działania wymagające identyfikacji, pomiaru i transformacji. Zbieranie, aktualizowanie i uzupełnianie tych danych jest najbardziej kosztownym i pracochłonnym procesem w budowaniu i eksploatacji GIS.

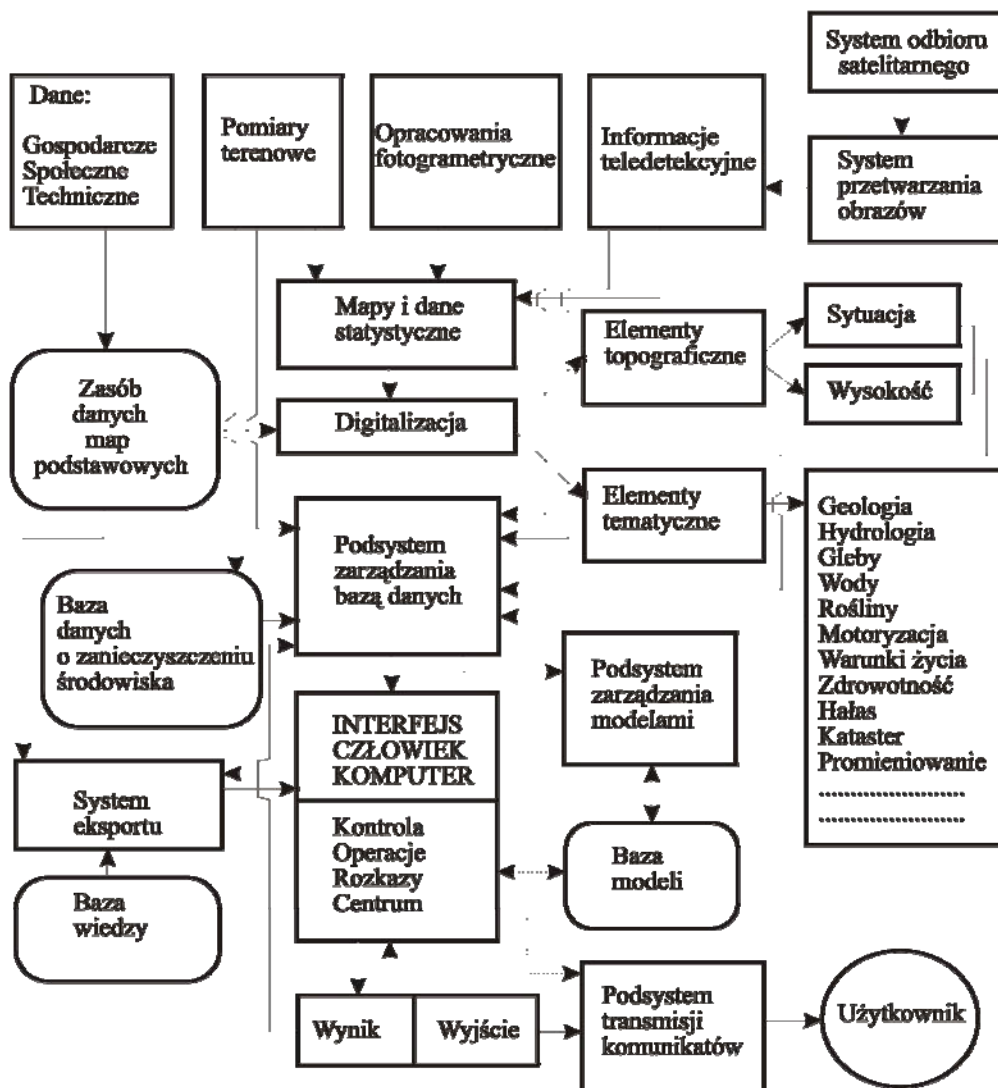
## **2.2. Składowe informacyjne**

Dane przestrzenne są faktami świata rzeczywistego. Z danych można wyciągać trzy rodzaje informacji: informacje przestrzenne dotyczące umiejscowienia, informacje atrybutowe (opisowe nieprzestrzenne) oraz informacje czasowe (dotyczące zmian danych spowodowanych upływem czasu). Geograficzne systemy informacyjne przetwarzają dane wykorzystując modele matematyczne dostarczając oczekiwanych informacji w środowisku zarządzanym komputerem. Można wydzielić dwie kategorie danych: podstawowe i dodatkowe. Do danych podstawowych zalicza się: identyfikatory, dane sytuacyjne (przestrzenne), dane atrybutowe (nieprzestrzenne). Dane dodatkowe to dane czasowe, parametry jakościowe, wskaźniki, oznaczenia, itp., które są informacjami o faktach i o danych.

Składowe informacyjne są głównymi składowymi GIS. Są one pozyskiwane z danych zbieranych metodami: 1) pomiarów terenowych, 2) fotogrametrycznymi lub 3) teledetekcyjnymi z pułapu lotniczego lub satelitarnego. W przypadku monitoringu środowiska gromadzi się niezwykle bogactwo informacji z przeróżnych źródeł, w tym także z zestawień statystycznych i przeróżnych map. Zasadnicze dane dotyczące monitoringu skażenia środowiska są pozyskiwane z instytucji prowadzących w ramach swych zadań statutowych badania stanu środowiska naturalnego. Dane te są udostępniane lub sprzedawane zazwyczaj w postaci zestawień tabelarycznych lub niekiedy dyskietek z danymi. Postać tych danych jak również danych zbieranych innymi metodami pomiarowymi bardzo różni się od siebie i dlatego dane muszą być wprowadzane do przestrzennej bazy danych różnymi ścieżkami. Ilustruje to rys.1.1. Jak widać przepływ danych rozpoczyna się od przeprowadzenia przestrzennie zorientowanych obserwacji, pomiarów i badań rzeczywistego środowiska, dokonywanych zdalnie lub bezpośrednio.

Pomiary bezpośrednie i badania terenowe można wprost wprowadzać do bazy danych (łącznie z błędami pomiaru, kategoryzacji i umiejscowienia). Mogą też być one

zinterpretowane i wykorzystane do opracowania map, a po dokonaniu digitalizacji wprowadzone do bazy. Analogowe zobrazowania teledetekcyjne muszą być również poddane interpretacji przez człowieka, a wynik interpretacji zdigitalizowany. Do konwersji map analogowych na postać cyfrową stosuje się dwie metody digitalizacji: ręczną i skanowanie elektroniczne.



Rys. 1.1. Strukturalny schemat składowych informacyjnych dynamicznego GIS wraz ze ścieżkami wprowadzania danych do bazy (linie ciągłe ilustrują przepływ danych, linie przerywane - przepływ rozkazów, prostokąty przedstawiają funkcjonujące modele, a prostokąty owalne - bazy danych lub modeli).

Poprawność, sprawność, postać oraz szybkość przeprowadzania digitalizacji ręcznej zależy od dostępnego sprzętu i programu nadzorującego digitalizację. W wyniku digitalizacji mapy rzeczywistej można uzyskać zbiory cyfrowe w dwóch postaciach (Mierzwa 1994):

- (1) - mapy cyfrowej, tj. takiego zapisu mapy konwencjonalnej, który umożliwi wyświetlenie jej na ekranie komputera i wydrukowanie (automatyczne wykreślenie),
- (2) - bazy danych przestrzennych, która również ma własności jak zbiory w postaci (1), ale przez zastosowanie kodowania topologicznego oraz połączenia z bazą atrybutów - umożliwia przeprowadzanie analiz GIS-owskich.

Zbiory (1) najczęściej gromadzone są w standardzie DXF natomiast w trakcie formowania zbiorów (2) podczas digitalizacji musi być wprowadzona topologia obiektów obrazowanych na mapie.

Skanowanie elektroniczne mapy lub rysunku przekształca postać graficzną (analogową) na postać cyfrową w formacie rastrowym. Następnie obrazy zeskanowanych map są zazwyczaj przekształcane na cyfrową mapę wektorową, do której dołącza się informacje z bazy danych (takie jak np. o użytkowaniu ziemi, gatunkach roślin, typach gleb itp.).

Po przeprowadzeniu automatycznego skanowania elektronicznego konwersję obrazu rastrowego na wektorowy można przeprowadzić jednym z trzech następujących sposobów : 1) kursora ekranowego, 2) wektoryzacją półautomatyczną i 3) wektoryzacją automatyczną.

W metodzie pierwszej operator śledzi mapę i zestrzaja jej obraz z obrazem na ekranie. Następnie odnajduje obraz linii na ekranie używając myszy do jej identyfikowania i śledzenia kursorem na monitorze. Jest to dobra metoda dla wektoryzacji map złożonych z wielu warstw informacyjnych.

W sposobie półautomatycznej wektoryzacji komputer usiłuje śledzić linię automatycznie, ale zatrzymuje się i czeka na interwencję operatora gdy dalszy przebieg linii jest wieloznaczny (linie się przecinają lub rozgałęziają). Jest to dobry sposób wektoryzacji, ale dla map nieskomplikowanych z ograniczonym tekstem opisowym i niezbyt dużą liczbą linii.

W pełni automatycznej wektoryzacji komputer sam śledzi wszystkie linie, często całonocnym trybem przetwarzania wsadowego. Niemniej jeśli pozostawi nierozwiązane powierzchnie, trzeba to uzupełnić sposobem pierwszym lub drugim. Pracuje poprawnie przy wektoryzacji prostych map, takich jak wielkoskalowe mapy katastralne czy mapy warstwowe.

Dane cyfrowe można generować wprost z opracowań fotogrametrycznych przez stereokompilację z modelu stereoskopowego w autografie lub metodą fotogrametrii ekranowej bądź też z ortofotografii cyfrowej. Stereodigitalizacja wykonywana na autografie analitycznym lub wspieranym komputerowo autografie analogowym wykorzystuje zespół programów edycji graficznej typu CAD np. System MicroStation f-my Intergraph lub AutoCAD. Autograf i system połączone są specjalnym interfejsem. Autograf spełnia rolę przestrzennego trójwymiarowego digitalizatora, a system w czasie rzeczywistym umożliwia wprowadzanie i edycję danych na monitorze. Dane pozyskiwane w taki sposób metodą stereodigitalizacji mogą być przesyłane do stacji GIS i innych systemów CAD za pośrednictwem np. formatu DXF.

Stosunkowo najłatwiej można pozyskiwać dla GIS-ów składowe informacyjne z satelitarnych lub lotniczych obrazów cyfrowych. Niektóre geograficzne systemy informacyjne przyjmują bowiem wprost "surowe" obrazy teledetekcyjne, a ponadto wyposażono je w moduły do przetwarzania i klasyfikacji obrazów. Jak widać z powyższego obrazy cyfrowe są kierowane do źródłowej bazy danych GIS, jako rezultat łącznej decyzji analizującego je operatora - człowieka, procesów przetwarzania oraz algorytmów statystycznych biorących udział w przetwarzaniu i klasyfikacji. Są to zawsze skomplikowane procesy uzależnione od interpretatora nawet wówczas gdy jest przeprowadzona automatyczna czy półautomatyczna analiza obrazów cyfrowych.

Poważne konsekwencje w dalszych opracowaniach, wynikające ze zbierania danych cyfrowych dla potrzeb GIS mają głównie:

- jakość danych cyfrowych niezbędnych do opracowania map topograficznych i umiejscowienia danych, oraz
- wymiana informacji cyfrowych między różnymi komputerowo wspieranymi systemami opracowania map.

Dane do opracowania map cyfrowych mają służyć do kompilacji map w różnych skalach i dlatego są one wyrażane we współrzędnych terenowych bez przypisywania im założonych dokładności jak to miało miejsce przy opracowaniu map konwencjonalnych w różnych skalach. Niemniej dokładność określenia współrzędnych terenowych ma istotne znaczenie zarówno przy kompilacji map cyfrowych jak i w rezultatach analiz dokonanych GIS-em.

Używane od lat i stosowane obecnie powszechnie liczne komputerowo wspierane systemy opracowania map są wykorzystywane do gromadzenia danych cyfrowych oraz do manipulowania nimi. Każdy z tych systemów ma niestety własną strukturę bazy oraz własną możliwość wymiany danych zgromadzonych w innych systemach. Dlatego wymiana danych cyfrowych jest kosztowna i niepraktyczna. Jest to powodem, że w wielu krajach opracowano lub opracowuje się standardy dla opisu danych przestrzennych, w tym standardy dotyczące struktury danych przestrzennych, standardy do wymiany danych cyfrowych, standardy jakości kartograficznych danych cyfrowych itp.. Na przykład w USA jest to standard przestrzennego transferu danych SDTS (Spatial Data Transfer Standard), w Niemczech ATKIS, w Wielkiej Brytanii NTS w Francji EdiGeo w Danii DSFL czy DIGEST w NATO lub CCSM (Canadian Council Surveying and Mapping) w Kanadzie. Standardy i rozwiązania opracowane i przyjęte w jednych krajach są adaptowane dla własnych potrzeb przez inne kraje. Standardy te dostarczają narodowego formatu do wymiany danych topograficznych wprowadzając jednostkę podstawową formatu (często o składowych graficznych), dowolne atrybuty oraz informacje o przestrzennych zależnościach z innymi cechami. Opracowanie takich standardów również w Polsce staje się rzeczą bardzo pilną.

Do składowych informacyjnych zaliczyć należy także modele logiczne i modele matematyczne, które wykorzystuje się do przetwarzania danych. Mogą być one uproszczone, a nawet bardzo przybliżone. W większości profesjonalnych systemów GIS modele są zazwyczaj wprowadzane na czas użycia, a nie instalowane. Modele wykorzystuje się do procesów analitycznych i manipulacyjnych łącznie z danymi zgromadzonymi w GIS w celu generowania określonych informacji. Modele logiczne manipulują nieciągłymi zmiennymi i mogą to być np. moduły umożliwiające modelowanie rozprzestrzeniania się zanieczyszczeń pyłowych lub gazowych w atmosferze albo skażeń i zanieczyszczeń w glebie lub wodzie. Modele matematyczne manipulują danymi ciągłymi (np. współrzędnymi  $x, y$ ) i stałymi (np. półosiąmi elipsoidy lub też wynikającymi z innych warunków matematycznych) w celu

określenia np zmian rzutowania. Modele można sprawdzić określając ich jakość, ale wymaga to pracy w terenie.

### 2.3. Składowe systemu

Do składowych systemu należy zaliczyć oprogramowanie i oprzyrządowanie. Głównym przyrządem, który stymuluje rozwój GIS jest komputer. Komputery sklasyfikowano (Acharya, Bell 1992) jako:

- bardzo duże superkomputery (Cray X-MP lub Y-mp),
- duże superkomputery (IBM 370),
- superminikomputery (VAX 11/780 i MicroVAX II),
- minisuperkomputery (NPL i VAX 8978),
- minikomputery (PAP 11/70),
- mikrokomputery (IBM PC lub inne PC).

Technologia komputerowa ustawicznie się zmienia i to bardzo szybko. Zacytuję tu dane zaczerpnięte z pracy Acharya i Bella (1992). Nowe struktury jednostek centralnych biurkowych stacji roboczych mają 64 bitową długość słowa, a więc podobną do superkomputerów i są w stanie adresować olbrzymią ilość pamięci. Pojemność głównej pamięci stale rośnie, 16 MB w mikrokomputerach jest niczym nadzwyczajnym, a wiele biurkowych stacji roboczych ma pamięć ponad 64 MB. Prędkość przetwarzania także rośnie średnio 1,5 razy/rok. Niektóre procesory jednostek centralnych komputerów osiągają ponad 70 MIPS (Millions Instructions Per Second). Oczekiwana jest szybkość przetworzenia 1000 MIPS i większa przy 400 mega FLOPS (operacji zmiennie-przecinkowych na sekundę). Wpływa to na znaczne przyspieszenie operacji obliczeniowych w tak zasadniczy sposób, że obliczenia, które przed czterema laty trwały 8 godzin, w roku przyszłym będą przeprowadzone w przeciągu 3 minut..

Objętości pamięci również wzrastają. Są dostępne dyski 5.25" z 2.5 GB. Prędkość przenoszenia z dysku do pamięci głównej także zwiększyła się niezwykle do 10 MB/sekundę. Ale jest ona nadal zbyt mała dla niektórych zastosowań. Systemy rozwinięte takie jak zsynchronizowane równoległe działające dyski twarde wspomagają prędkość przesyłania i są stosowane w niektórych fotogrametrycznych cyfrowych stacjach roboczych (Sitek, 1992). Rozwinęły się także urządzenia taśmowe. Technologia cyfrowa D2 umożliwia na jednej taśmie cyfrowej D2 zapis 165 GB i przekazuje dane z prędkością 16 MB/sekundę..

Innym oprzyrządowaniem używanym w GIS są różnego rodzaju drukarki, plotery, skanery i urządzenia do digitalizacji.

Najważniejszą składową systemu jest jego oprogramowanie. Według *The 1991 GIS World Software Survey* było dostępnych na rynku 117 pakietów oprogramowania GIS. Pomijając systemy uniwersyteckie, które są wielokrotnie tańsze, to ceny pakietów komercyjnych mieściły się od jednego do 65 tysięcy USD i zależały głównie od struktury i liczby funkcji w jakie je wyposażono. Najwięcej pakietów wykorzystuje systemy operacyjne DOS i UNIX. Ten ostatni jest instalowany w komputerach klasy *work station*, dlatego jego sprawność jest lepsza niż w systemach działających w środowisku DOS. Pakiety mogą wymieniać dane w ponad 20 formatach, ale najczęściej stosowany jest format ASCII oraz



DXF. Niemalże wszystkie pakiety mogą współpracować z takimi urządzeniami jak digitalizatory, skanery, ale tylko 70% z ploterami, natomiast ze stacjami fotogrametrycznymi czy z GPS tylko około 30%. Prawie połowa systemów umożliwiła przekształcenie zapisu wektorowego na rastrowy i nieco więcej systemów *vice versa*. Takie konwersje danych podobnie jak wymiana informacji cyfrowych między różnymi komputerowo wspieranymi systemami opracowania map, należą do problemów technicznych GIS i mają wpływ na jego stronę ekonomiczną. 90 % GIS-ów wyposażono w moduły do przetwarzania i analizy obrazów skanerowych. Prawie wszystkie systemy wyposażono w pakiety do obrazowania kartograficznego, moduły do digitalizacji map, programy do zarządzania bazą danych, moduły do przeprowadzania analiz statystycznych i analiz geograficznych.

Pakiety GIS są przydatne w analizach danych dotyczących skażenia środowiska. W olbrzymiej większości realizują one te same zadania, ale sposób realizacji może się znacznie różnić, a ponadto mogą one być mniej lub bardziej "przyjazne" użytkownikowi.

Istotną rolę w GIS odgrywają systemy zarządzania bazą danych. W ostatnich latach poprawiła się ich wydajność. Dostępne są bazy rozprawdzające z dwufazową dostępnością umożliwiające gromadzenie danych w wielu miejscach. Większość GIS-ów wyposażono we własne bazy graficzne oddzielone od stowarzyszonych z nimi atrybutów, co nie jest dobrym rozwiązaniem, gdyż utrudnia to w dużych bazach integrację danych a ułatwia ich desynchronizację.

Niektóre GIS-y oparte na systemie operacyjnym DOS wymagają wprowadzenia ciągu rozkazów tzn. wpisania nazwy modułu i parametrów potrzebnych do przeprowadzenia obliczeń. Większość systemów posiada rozwijalne *menu*, które prowadzi użytkownika krok za krokiem podczas rozwiązywanego zadania oferując różne opcje postępowania, które musi wybrać użytkownik. Musi on sam decydować o wyborze opcji lub parametru, gdyż program sygnalizuje że nie ma (*default*) automatycznego wyboru. Niektóre, te bardziej zaawansowane pakiety GIS działają w środowisku graficznym WINDOWS.

## 2.4. Standardy dokładności

Standardy dokładności określają minimalne dokładności, które należy zachować przy realizacji zadań, projektów czy rozwiązań. W geograficznych systemach informacyjnych podobnie jak przy opracowaniu map i pracach pomiarowych standardy dokładności umożliwiają realizację zadań z określonym stopniem doskonałości i niezawodności. Koniecznym staje się wprowadzenie nawyku oceny dokładności danych wejściowych, aby można było ujawnić ich wpływ na rezultatach uzyskanych w wyniku zastosowania modułów GIS. Dlatego w tablicach bazy danych (relacjach) - danym podstawowym powinny towarzyszyć dane pomocnicze zarówno ilościowe jak i jakościowe podające oprócz identyfikatorów, wartości liczbowych i oznaczeń, również wskaźniki nieciągłości zmiennych, jednostki miar zestawionych wartości, błędy średnie danych, wskaźniki prawdopodobieństwa poprawności zestawionych danych (atrybutów) itp. Jest to szczególnie ważne, kiedy do przetwarzania danych są wykorzystywane modele, gdyż zestawione w tablicach bazy danych błędy i wskaźniki umożliwiają ocenę jakości modelu. Podobnie jak to ma miejsce w geodezji i fotogrametrii również w GIS rozkład wariancji może być wykorzystywany do oszacowania jakości informacji wygenerowanej przez geograficzny system informacyjny, kiedy te informacje powstają przy udziale modeli matematycznych. Procesy matematycznego modelowania zmiennych losowych typu ciągłego (np. współrzędnych) i wartości stałych (np. parametrów modelu) dostarczają

nowych informacji. Można przy tym gromadzić w tablicach bazy danych błędy średnie tych nowo wygenerowanych informacji korzystając z rozkładu wariancji danego modelu matematycznego. Rozkład wariancji znany w praktyce jako rozkład błędów umożliwia dla modelu matematycznego  $a = f(b, c)$  obliczenie odchyłki standardowej  $\sigma_a^2$ , jeśli znane są odchyłki  $\sigma_b^2$  i  $\sigma_c^2$  (Mikhail 1976):

$$\sigma_a^2 = \sigma_b^2 + (\hat{a} / \hat{b})^2 + \sigma_c^2 + (\hat{a} / \hat{c})^2 + 2\sigma_b\sigma_c(\hat{a} / \hat{b})(\hat{a} / \hat{c})$$

gdzie:  $b$  i  $c$  - dane zgromadzone w tablicach bazy danych GIS wraz z odchyłkami standardowymi  $\sigma_b$  i  $\sigma_c$ ;  $a$  - nowe informacje; różniczki cząstkowe  $\sigma_a$ ,  $\sigma_b$  i  $\sigma_c$  musi dostarczyć użytkownik albo obliczyć system. Jeśli nie ma korelacji między  $b$  i  $c$  - ostatni człon równania można pominąć.

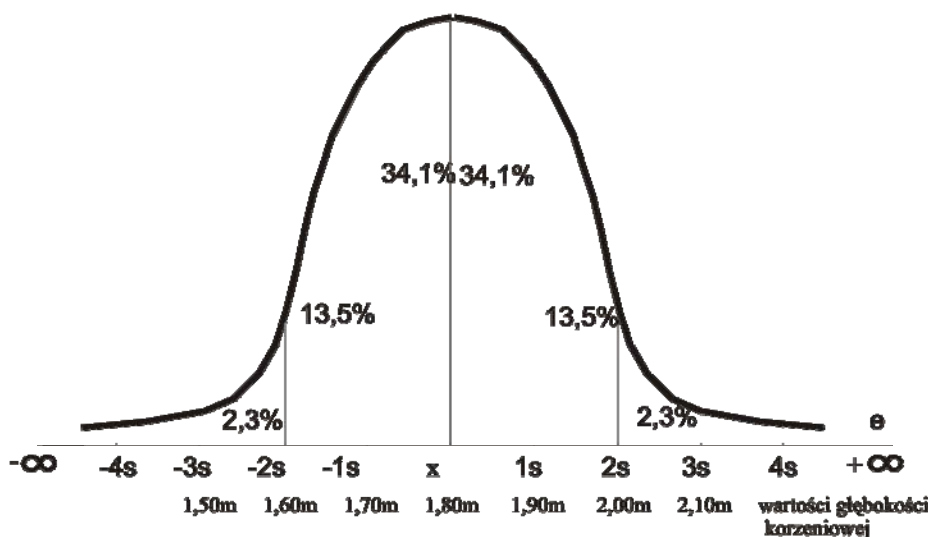
W przypadku kiedy GIS przetwarza dane przy wsparciu modelu liczbowego do oceny jakości otrzymanej informacji można wykorzystać teorię mnogości. W celu zapoznania czytelnika z praktycznym zastosowaniem tej teorii do rozkładu błędów (kiedy rezultaty przetwarzania GIS-em powstały przy wsparciu modelu liczbowego) przytoczę przykład z publikacji (Ramlal, 1992). Niech modelem liczbowym będzie np. WŁAŚCIWOŚĆ PASTWISKOWA Nr 1 kiedy:

- 1) typ gleby należy do E (gleby bagiennie) i
- 2) głębokość korzeniowa wynosi 1.50 m do 2.00 m.

Ograniczymy model do tych dwóch cech chociaż można by go poszerzyć o: stan drenażu, urodzajność i zapasy wilgotności. W tabeli bazy danych oprócz identyfikatora obszarów glebowych i wskaźników klasyfikacyjnych gleb podano także odchyłki standardowe współrzędnych  $x, y$  dla węzłów i łuków, oznaczenia ciągłości i nieciągłości zmiennych, prawdopodobieństwo poprawności określenia typów gleb i głębokości korzeniowe gleby wraz z jednostkami pomiarowymi i odchyłkami standardowymi ich pomiaru.

Rozpatrzmy obszar glebowy o identyfikatorze np. 0023, w którym prawdopodobieństwo klasy gleby E (bagienna) wynosi 0.65, głębokość korzeniowa równa się 1.80 m - tzn. prawdopodobieństwo, że głębokość korzeniowa mieści się w granicach 1.50m do 2.00m - jest 0.98. Takiego obliczenia można dokonać na podstawie rozkładu normalnego błędu pomiaru głębokości korzeniowej, dla głębokości 1.80m i odchyłki standardowej  $\sigma=0,10$  m (rys.2), wiedząc (co jest również ujęte w tablicy bazy danych), że ma się do czynienia ze zmienną losową typu ciągłego. Z wykresu na rys.1.2 widać, że dolny kraniec klas (1.50 - 2.00 m) głębokości korzeniowej przypada na  $3\sigma$  (1.50 m) od średniej (1.80 m), natomiast górny kraniec klas jest na odległości  $2\sigma$  (2.00 m) powyżej średniej. Zatem prawdopodobieństwo, że błąd zawiera się między liczbami 1.50 i 2.00m (jak pokazano pod osią odciętych  $\varepsilon$  będącej osią zasięgu odchyłek standardowych  $\sigma$ ) można wyrazić powierzchnią pola ograniczonego: od dołu osią odciętych  $\varepsilon$ , od góry krzywą prawdopodobieństwa a po bokach rzędnymi  $\varepsilon=1.50m$  oraz  $\varepsilon=2.00m$  (która wynosi 97,5%), w stosunku do całego pola ograniczonego osią  $\varepsilon$  i krzywą prawdopodobieństwa (100%). Stąd dla przytoczonego przykładu prawdopodobieństwo wynosi 0.98. Przyjmując zatem, że model jest doskonały (prawdopodobieństwo sprawności modelu wynosi 100%), zatem prawdopodobieństwo obszaru glebowego 0023 posiadającego WŁAŚCIWOŚCI PASTWISKOWE Nr 1 wynosi 0.64 (lub 0.65 x 0.98). Natomiast gdyby prawdopodobieństwo sprawności modelu wynosiło 80%, wtedy prawdopodobieństwo tego

obszaru glebowego o WŁAŚCIWOŚCIACH PASTWISKOWYCH Nr 1 wyniosłoby 51% (czyli 0.64 x 80). Jest to problem przecięć teorii mnogości.



Rys.1.2. Rozkład normalny błędu pomiaru głębokości korzeniowej dla średniej głębokości pomiaru równej 1.80m i odchyłki standardowej  $\sigma = 0.10\text{m}$  (Ramlal,1992)

Przytoczony przykład dotyczył oceny skuteczności modelu, w którym występowały tylko dwie cechy: obszar glebowy z prawdopodobieństwem występowania tej gleby wynoszącym 0.65 i głębokość korzeniowa. Zarówno w tym jak i w innych modelach liczbowych tych cech może być więcej i jeśli są dla nich podane wskaźniki dotyczące poprawności ich określenia lub odchyłki standardowe pomiarów jakimi je wyznaczono, wtedy również możliwa jest ocena skuteczności generowanych przez model i GIS wyników - stosując przytoczone powyżej jak też inne metody rachunku prawdopodobieństwa.

Analiza ilościowa błędów może być przeprowadzona podczas każdego etapu przetwarzania danych. W zasadzie wartości błędów powinny być określane po każdym etapie tego przetwarzania, ale w opracowaniach z udziałem danych teledetekcyjnych błędy są obliczane po zakończeniu przetwarzania i analiz danych i zazwyczaj są ukierunkowane na dokładność sytuacyjną i tematyczną. Podobnie jak w analizach dokładnościowych prowadzonych w geodezji i fotogrametrii, tak i w opracowaniach GIS - wielkość próbki na której dokonuje się oszacowania wyników odgrywa ważną rolę. Istnieją różne zalecenia i wytyczne do określania wielkości próbki - takiej aby nadawała się do przeprowadzenia analiz statystycznych. Rozmieszczenie badanych punktów pól testowych w próbce spełnia istotną rolę w ocenie dokładności, zwłaszcza, gdy dotyczy to obrazu teledetekcyjnego. Rozmieszczenie to musi być reprezentatywne dla całego klasyfikowanego obrazu.

Dokładność sytuacyjną obrazów teledetekcyjnych można określić błędami średnimi położenia, ale należy pamiętać, że nie ujmuje ona wszystkich pikseli, gdyż błąd średni dotyczy tylko punktów kontrolnych, które posłużyły do wpasowania obrazu cyfrowego do mapy. Dlatego najdokładniejszym, ale i bardzo kosztownym sposobem sprawdzenia dokładności sytuacyjnej jest kontrola terenowa z wykorzystaniem danych GPS.

Do oceny dokładności wygenerowanych map tematycznych lub przeprowadzonej klasyfikacji obrazów teledetekcyjnych można zastosować metodę zestawienia w postaci rastra (macierzy) kwadratów odchyłek (z pól testowych) uzyskanych z różnic położenia klas wygenerowanych przez GIS i klas aktualnie istniejących w terenie. Kwadraty odchyłek uporządkowane są w taki sposób w macierzy, że kolumny informują o miejscach położenia odchyłek a wiersze określają klasy (cechy) jakie uzyskano z klasyfikacji. Jest to doskonały i efektywny sposób określania dokładności, gdyż umożliwia ocenę dokładności każdej wygenerowanej klasy. Taka macierz zestawionych błędów może być wykorzystana jako dane wyjściowe do różnych oszacowań metodami statystycznymi.

### **3. Strategia optymalizacji GIS**

O optymalizacji geograficznego systemu informacyjnego można mówić wtedy, gdy do jego budowy zostały wykorzystane standardy geodezyjne i standardy stosowane w opracowaniu map, a ponadto wówczas gdy jest efektywny ze względu na nakład kosztów. W publikacji (Acharya, Bell 1992) podano trzy działania strategiczne, które powinny przyczynić się do optymalizacji projektu GIS:

- 1 - jak najwięcej nasycić projekt standardami dokładnościowymi,
- 2 - zminimalizować koszty wykorzystując do tego analizę zależności: koszty-korzyści,
- 3 - zwiększyć liczbę użytkowników systemu.

Standardy dokładności zapewnią pozyskiwanie poprawnych i zadawalająco dokładnych wyników generowanych przez GIS. Nie chodzi tu o zwiększanie precyzji i podnoszenie dokładności, ale o kontrolę i ocenę dokładnościową danych wejściowych oraz o metody oszacowania pozyskiwanych wyników. Współczesna technologia komputerowa zezwala na takie działania nie powodując przy tym zbytecznego zwiększania kosztów.

Zmniejszenie kosztów musi być poprzedzone analizą efektywności kosztów. Dokonuje się tego przez porównanie korzyści i kosztów. Oddzielnie przeprowadza się ocenę korzyści i odrębnie analizę kosztów, przypisując obydwu ocenom wskaźniki. Końcowym etapem analizy efektywności kosztów jest tworzenie kombinacji ze wskaźników i obliczanie stosunków: korzyści/koszty. Koszt projektu GIS można znacznie zmniejszyć jeśli jest on budowany przez wielu użytkowników, różne wydziały, biura, agencje i instytucje. W ten sposób unika się dublowania wielu operacji niekiedy bardzo pracochłonnych, natomiast koszty i dane mogą być odpowiednio rozdzielane. Wyżej wymienieni autorzy sugerują, że takim wielozadaniowym, mającym wielu użytkowników GIS-em jest taki, z którego korzystają Komisja Planowania, Biuro Planowania i Rozbudowy, Służba Geologiczna, Wydział Prac Publicznych, Służba Leśna, Wydział Zasobów Naturalnych, Wydziały: Nawadniania, Edukacji, Policji, Kanalizacji i Zaopatrzenia w Wodę, Geodezji i Katastru, Drogowy, Ewidencji Ludności, Urząd Skarbowy oraz agencje użyteczności publicznej.

Dobrze zaprojektowany GIS może być optymalnie wykorzystany i służyć dla wielu celów u różnych użytkowników po minimalnej modyfikacji struktury podstawowej.

#### **4. Uwagi końcowe**

Podczas projektowania geograficznego systemu informacyjnego oprócz kryterów ekonomicznych trzeba uwzględniać szereg innych wytycznych, takich jak:

- uwarunkowania geodezyjne związane z układem odniesienia, odwzorowaniem kartograficznym, zasadami pomiaru i opracowania map oraz dostępnością do kontroli geodezyjnej;
- zmierzać do zapewnienia integracji danych informacyjnych podstawowych i dodatkowych, niezależnie od tego czy pochodzą z pomiarów terenowych, fotogrametrycznych czy teledetekcyjnych;
- zabezpieczyć, aby system uwzględniał standardy dokładności przy realizacji poszczególnych zadań zapewniając określony stopień poprawności i niezawodności, np. przy opracowaniu map, a ponadto aby był wyposażony w standardy do zamiany formatu
  - łatwe i dostępne dla użytkownika.

System trzeba tak zaprojektować aby sytuacyjne i atrybutowe parametry jakościowe były bezpośrednio powiązane z opisem środowiska zgromadzonym w bazie danych. Dlatego będzie można sukcesywnie zbierać te parametry w tablicach atrybutowych bazy danych. Takie tablice muszą zawierać przynajmniej jeden rekord dla każdej jednostki zapisanej w bazie danych.

Modele zwłaszcza logiczne, które są wykorzystywane okazjonalnie, powinny być osobnym modułem GIS-u - nie instalowane, tylko wprowadzane na czas użytkowania. Dotyczy to takich modeli jak modele ekologiczne np. model oceny zagrożenia ekotoksykologicznego metalami ludzi w rejonach przemysłowych (Borowik, 1992), czy modele do ilościowej oceny erozji gleb, czy też modele do oceny rozprzestrzeniania się zanieczyszczeń powietrza.

Natomiast często używane modele matematyczne oraz model oceny jakości rezultatów otrzymanych z przetwarzania GIS-em - powinny być na stałe włączone do pakietów systemu.

Szczególnie wtedy, gdy jest to GIS przeznaczony do monitorowania środowiska, gdyż jak wiadomo bardzo dużo informacji generowanych GIS-em nie może być sprawdzonych jakimkolwiek innym sposobem.

Dlatego bardzo jest ważna ocena dokładności parametrów i danych, które są danymi wejściowymi procesów i analiz GIS-owskich oraz są wykorzystywane do modelowania procesów i zjawisk dynamicznych. Do oceny tej należy używać rozkładu błędów, teorii mnogości lub innych rozwiązań statystycznych, które są wspierane danymi zgromadzonymi w GIS. Niemniej są niezbędne dalsze badania nad możliwością oceny błędów produktów wytworzonych przez GIS, zwłaszcza tych, które uzyskuje się z integracji danych teledetekcyjnych, GIS i danych pozyskiwanych bezpośrednio pomiarami w terenie (często tabelarycznych). Zachodzi zatem konieczność określania błędów sytuacyjnych, błędów klasyfikacji obrazów cyfrowych i korelacji między nimi. Badania nad strukturą, przenoszeniem się i rozkładem błędów powinny doprowadzać do opracowania metod przestrzennego wyrównania rezultatów pozyskiwanych z GIS.

## **Literatura**

- [1]. Acharya B., Bell W.C.: Designing an Optimal and Scientific GIS Project. Int.Arch. of Ph.and R.S., Vol.XXIX, part B3, Washington, p.627-633, 1992.
- [2]. Borowik W.: Ocena możliwości modelowania ekotoksykologicznego zagrożenia metalami ludzi w rejonach przemysłowych. Rozprawa doktorska, AGH, Kraków, 1992.
- [3]. Ekspertyza dotycząca podstawowych założeń nowego państwowego systemu współrzędnych geodezyjnych w Polsce (pod red.W.Barana). PAN Komitet Geodezji, W-wa, 1993r.
- [4]. Mierzwa W.: Pozyskiwanie danych cyfrowych z istniejących opracowań kartograficznych. (Sprawozdanie z zadania nr 4 z projektu badawczego KBN "Monitoring środowiska z wykorzystaniem GIS". Kraków AGH, 1994 (niepublikowane).
- [5]. Mikhail E.M.: Observations and Least Squares. IEP Series, N.York 1976.
- [6]. Ramlal B.: Durmmond J.E., A GIS Uncertainty Subsystem. Int. Arch.of Ph.and R.S., Vol. XXIX part B3, Washington, p.356-362, 1992.
- [7]. Sitek Z.: Fotogrametryczne cyfrowe stacje robocze. Moniterra, nr 7/8 str.2-9, Kraków 1992.
- [8]. Sitek Z.: Zarys teledetekcji lotniczej i satelitarnej. Wydawnictwa AGH, str 1-303, Kraków 1992 (1).
- [9]. Sitek Z.: Zintegrowany GIS w monitoringu skażonego środowiska. Archiwum Fotogrametrii, Kartografii i Teledetekcji, Vol.1, str.1-1 do 1-10, Kraków 1994.

### ***Elements of Designing Integrated GIS for Environmental Monitoring***

#### ***Summary***

*In recent years GIS technology has matured and offers excellent capabilities for environmental monitoring and protection purposes of the areas under increased ecological hazard. In such case GIS must meet many criterion besides economics to become an optimal GIS. Regional, in a medium scale, GIS for environmental and ecological studies must consist of the following components: geographic or geodetic, information, system (hardware and software) and accuracy standards. All these components are presented and discussed.*

*Geodetic component in p.2.1 gives outline of problems connected with polish conditions.*

*Information component (p.2.2) which are major component of a GIS technology, presents collection of information using ground, photogrammetric and remote sensing data collection methods in either raster or vector mode. Manual digitization and electronic scanning and also some basic standards are discussed. A concept and main GIS operation are presented in the fig.1.1.*

*System component is presented in p.2.3. Brief information connected with main hardware factors, main memory size, processing speed, remarks on DBMS and other are presented.*

*Accuracy standards are discussed at p.2.4. Manipulation of position and attribute quality parameters were also discussed. The variance propagation was used to estimate the quality of GIS generated information when a mathematical model is being used. Therefore the variance propagation of the model  $a = f(b,c)$  and standard deviation  $\sigma_a^2$  of  $a$  was reminded.*

*The example a normal distribution of the one parameter in the model (rooting depth) measurement error is attached (fig.1.2). This figure allowed computation of the area under distribution curve and determined the certainty factor for this parameter. Using such certainty factors accounted for other parameters of the model and Crisp Set Theory one can estimate the probability of the predicted applied model efficiency. But to reach this goal the positional and attribute quality parameters must be stored in attribute database tables.*

*The third and forth part of the paper concerns strategy for GIS optimalization and criteria related to accuracy, cost and other factors are also discussed.*