

AUTOREFERAT

Opis osiągnięć naukowych

dr inż. Dariusz Waldemar Gościewski

Olsztyn 2017

SPIS TREŚCI

1. Imię i nazwisko.....	2
2. Posiadane dyplomy i stopnie naukowe z podaniem nazwy, miejsca i roku ich uzyskania	2
3. Informacje o dotychczasowym zatrudnieniu w jednostkach naukowych.....	2
4. Wskazanie osiągnięcia wynikającego z art. 16 ust. 2 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. z 2016 r. poz. 882 ze zm.).....	3
4.1. Tytuł osiągnięcia naukowego.....	3
4.2. Omówienie celu naukowego i osiągniętych wyników wraz z omówieniem ich ewentualnego wykorzystania	3
Wstęp	3
4.2.1. Pozyskiwanie i charakterystyka danych masowych	5
4.2.2. Struktura GRID.....	7
4.2.3. Cel i ogólna charakterystyka pracy	9
4.2.4. Dobór narzędzi informatycznych	15
4.2.5. Założenia metodyki i systemu	16
4.2.6. Konstrukcja systemu	19
4.2.7. Kontrola i analiza dokładności.....	22
4.2.8. Przetwarzanie danych	25
4.2.8.1. Przetwarzanie punktów pomiarowych.....	26
4.2.8.2. Generowanie struktury GRID	33
4.2.8.3. Przetwarzanie struktury GRID	45
4.2.9. Podsumowanie etapów badań i uzyskane wyniki.....	52
4.2.10. Możliwości praktycznego wykorzystania wyników badań.....	54
Bibliografia	55
5. Omówienie pozostałych osiągnięć naukowo-badawczych	61
5.1. Charakterystyka wybranych prac naukowo-badawczych.....	61
5.2. Wybrane publikacje	68
6. Perspektywy badawcze	70

1. Imię i nazwisko

Dariusz Waldemar Gościewski

2. Posiadane dyplomy i stopnie naukowe z podaniem nazwy, miejsca i roku ich uzyskania

- a) Tytuł zawodowy magistra inżyniera geodezji urządzeń rolnych, uzyskany na Wydziale Geodezji i Urzędzeń Rolnych Akademii Rolniczo-Technicznej w Olsztynie po obronie pracy magisterskiej pt. *System „Wcięcia geodezyjne” - pakiet komputerowych programów edukacyjnych*. Olsztyn 1991.
- b) Stopień doktora nauk technicznych w zakresie Geodezji i Kartografii nadany przez Radę Wydziału Geodezji i Gospodarki Przestrzennej Akademii Rolniczo-Technicznej w Olsztynie. Tytuł rozprawy doktorskiej: *Ocena dokładności i efektywności pozyskiwania informacji graficznych w procesie tworzenia mapy numerycznej*. Olsztyn 1998.
- Promotor: prof. zw. dr hab. inż. Maria Krystyna Szacherska
- Recenzenci: prof. zw. dr hab. inż. Ewa Krzywicka-Blum
 dr hab. inż. Alojzy Wasilewski, prof. nadzw.

3. Informacje o dotychczasowym zatrudnieniu w jednostkach naukowych

od 1 października 1991 r. do 31 maja 1998 r.:

Akademia Rolniczo-Techniczna w Olsztynie (od 1 września 1999 r. - Uniwersytet Warmińsko-Mazurski w Olsztynie)

Wydział Geodezji i Urzędzeń Rolnych (od 1 września 1992 r. - Wydział Geodezji i Gospodarki Przestrzennej)

Instytut Geodezji i Fotogrametrii

Stanowisko: asystent

od 1 czerwca 1998 r. do 31 maja 2013 r.:

Uniwersytet Warmińsko-Mazurski w Olsztynie (do 31 sierpnia 1999 r. - Akademia Rolniczo-Techniczna w Olsztynie)

Wydział Geodezji i Gospodarki Przestrzennej

Instytut Geodezji

Stanowisko: adiunkt

od 1 czerwca 2013 r. do teraz

Uniwersytet Warmińsko-Mazurski w Olsztynie

Wydział Geodezji, Inżynierii Przestrzennej i Budownictwa (do 31 grudnia 2014 r. - Wydział Geodezji i Gospodarki Przestrzennej)

Instytut Geodezji

Stanowisko: starszy wykładowca

4. Wskazanie osiągnięcia wynikającego z art. 16 ust. 2 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. z 2016 r. poz. 882 ze zm.)

4.1. Tytuł osiągnięcia naukowego

Osiągnięciem naukowym wynikającym z art. 16 ust. 2 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki jest monografia pt. *Metodyka i system przetwarzania danych pozyskiwanych masowo w celu numerycznego modelowania powierzchni z wykorzystaniem struktury GRID* wydana przez Wydawnictwo Uniwersytetu Warmińsko-Mazurskiego w Olsztynie, Olsztyn 2017 (541 s., Ark. wyd. 42, ISBN 978-83-8100-074-1).

Recenzenci: prof. dr hab. inż. Edward Osada

dr hab. inż. Krzysztof Kroszczyński, prof. WAT

4.2. Omówienie celu naukowego i osiągniętych wyników wraz z omówieniem ich ewentualnego wykorzystania

Wstęp

Stały wzrost zapotrzebowania na informację implikuje ciągłe podnoszenie wydajności współczesnych systemów informacji przestrzennej (ang. SIS - Spatial Information System) (Chrisman 2003, Bielecka 2006, Chang 2007, Goodchild 2009, Longley i in. 2011). Jednak mimo stałego rozwoju SIS, nadal problemem pozostają ograniczenia w bieżącym i dynamicznym przetwarzaniu dużej ilości danych (Arctur, Zeiler 2004, Huang i in. 2006, Cheney i in. 2009, Finley i in. 2009). Ustawiczny przyrost różnego rodzaju zasobów, które są rejestrowane przez pracujące często w czasie rzeczywistym coraz wydajniejsze systemy pomiarowe, stawia przed SIS wciąż nowe wyzwania (Higdon 2002, Keim i in. 2004, Thomas, Cook 2005, Yan, Thill 2009). Stosowane coraz powszechniej nowe technologie pozyskiwania danych pozwalają na rejestrację wyników obserwacji w sposób ciągły i automatyczny (Axelsson 2000, Kobler i in. 2007, Liu 2008, Vázquez, Pascual 2008). Pozwala to na pozyskiwanie dużej ilości obserwacji masowych bez bezpośredniego udziału człowieka (w omawianej pracy do określenia tego rodzaju danych stosowane jest pojęcie megadane). Megadane charakteryzują się szczególnymi cechami, które należy uwzględnić podczas projektowania systemów ich obsługi. Pozyskiwane automatycznie pomiary pozwalają na bardzo dokładne opisanie obiektów, jednak bezpośrednie wykorzystanie takich zasobów do wygenerowania numerycznego modelu powierzchni mierzonego obiektu jest często niemożliwe (Fahringer 2005, Travostino 2006, Bu i in. 2010). Megadane wymagają wcześniejszego przeanalizowania oraz przetworzenia do postaci pozwalającej na umieszczenie ich w bazach SIS głównie ze względu na ograniczenia dotyczące ilości informacji. Przykładowo szacowana objętość plików bazy danych (wielkość plików podano w przybliżeniu, dokładnie 1 KB to 1024 B, tu w zaokrągleniu 1000 B) (Longley i in. 2005):

- 1 MB = 10^6 - pojedynczy zbiór danych dotyczący małego projektu;
- 1 GB = 10^9 - cały system ulic dużego miasta lub małego państwa;
- 1 TB = 10^{12} - rzeźba terenu całej Ziemi w cięciu warstwicowym co 30 m;
- 1 PB = 10^{15} - zdjęcie satelitarne całej Ziemi z rozdzielczością 1 m;
- 1 EB = 10^{18} - trójwymiarowy model rzeźby terenu całej Ziemi z rozdzielczością 10 m.

Już same procesy pozyskiwania i archiwizacji takiej ilości danych stanowią wyzwania i są możliwe wyłącznie dzięki stałemu rozwojowi technik komputerowych (Bose, Frew 2005, Miller, Han 2009, Liu, Yang 2012). Dysponując dużą ilością danych pomiarowych można wygenerować bardzo dokładny model danego obiektu, jednak przetwarzanie takiej ilości informacji, a zwłaszcza jej zmian zachodzących w czasie rzeczywistym, jest praktycznie niemożliwe (Li i in. 2005, Cavalli i in. 2008, Longley i in. 2011). Obecnie pomiarom podlegają również nowe zjawiska fizyczne, społeczne i gospodarcze, które są rejestrowane z coraz większą rozdzielczością i częstotliwością czasową (Nyerges 2001, Ballas i in. 2004, Kwiatkowska 2007). Także w tych przypadkach szczególny nacisk jest kładziony na zapewnienie możliwości prowadzenia analiz w czasie rzeczywistym. Uchwycenie zmian mierzonych obiektów w czasie wymaga wielokrotnej rejestracji danych z tego samego obszaru w kolejnych epokach pomiarowych, co jeszcze bardziej zwiększa wymagania stawiane przed SIS. Zwiększa to zapotrzebowanie zarówno na przestrzeń dyskową, niezbędną do archiwizacji generowanych zasobów, jak i implikuje potrzebę wzrostu wydajności metod przetwarzania danych (Griffith 2004, Tzeng i in. 2005, Cressie, Johannesson 2008). Wykorzystanie tego rodzaju informacji wymaga zastosowania odpowiednio dobranych algorytmów przetwarzania, pozwalających na zachowanie wymaganej dokładności opracowania z jednoczesnym ograniczeniem ilości informacji trafiających do baz systemowych (O'Sullivan, Unwin 2003, Arctur, Zeiler 2004, Fischer 2006). Również coraz szersze wykorzystywanie sieciowego transferu informacji przyczynia się do poszukiwania sposobów optymalizacji gromadzenia i przesyłania danych. Zapewnienie wydajnych procesów przetwarzania dużej ilości danych jest obecnie jednym z głównych wyzwań stojących przed systemami informatycznymi (Zhang, Pazner 2004, Guo i in. 2005, Longley i in. 2005, Dudek 2006, Zhou, Chen 2011). Zadania stawiane przed obecnymi systemami przetwarzania informacji są realizowane częściowo dzięki stałemu wzrostowi wydajności infrastruktury sprzętowej. Jednak do zapewnienia kompletnego progresu całego systemu komputerowego niezbędny jest także rozwój odpowiedniego oprogramowania dostosowanego do realizacji wymaganych procesów przetwarzania danych (Banerjee i in. 2004, Banachowski i in. 2006, Johannesson i in. 2007, Nisan, Schocken 2008, Stein 2008, Pelikant 2009, Andrew, Tanenbaum 2010). Dodatkowo podczas przetwarzania dużej ilości danych należy dążyć do ograniczenia roli operatora systemu do minimum, co pozwala na uniezależnienie jakości przetwarzania od doświadczenia osoby obsługującej system. W związku z tym jest pożądana jak największa automatyzacja najważniejszych etapów przetwarzania, podnosząca wydajność i jakość całego procesu. W tym celu powinny ewoluować nie tylko sposoby pozyskiwania i gromadzenia danych, lecz także techniki ich wydajnego przetwarzania gwarantujące odpowiednią jakość i pożądaną

dokładność opracowania, z jednoczesnym zapewnieniem bieżącej aktualizacji zasobów w jak najkrótszym czasie.

4.2.1. Pozyskiwanie i charakterystyka danych masowych

Utworzenie numerycznego modelu rzeźby terenu (ang. DTM - Digital Terrain Model) wymaga uprzedniego zgromadzenia informacji pomiarowych, dzięki którym możliwe jest odtworzenie kształtu mierzonego obiektu i jego cyfrowej reprezentacji (Longley i in. 2005, Masser 2005, Deng 2007, Wilson 2012). Ze względu na to, że dynamika zmian otaczającej przestrzeni wymusza częstą rejestrację dużej ilości informacji w ograniczonym czasie, pożądane stają się metody pozwalające połączyć dokładność pozyskiwanej informacji z jej automatyczną i bieżącą rejestracją. Obecnie do tego celu stosowane są powszechnie specjalizowane, zintegrowane systemy pomiarowe, pozwalające na zautomatyzowane gromadzenie wymaganej ilości danych. Uzupełniają one coraz częściej klasyczne pomiary geodezyjne. Do wykorzystywanych obecnie zintegrowanych systemów pomiarowych, umożliwiających pozyskiwanie danych wykorzystywanych do utworzenia numerycznej reprezentacji powierzchni obiektu należy zaliczyć (Axelsson 2000, Paine, Kiser 2003, Beaudoin i in. 2004, Grządziel 2004, Reutebuch i in. 2005, Van Niel i in. 2008, Mallet, Bretar 2009, Shi i in. 2009, Hamilton, Beaudoin 2010):

- skaning laserowy (ang. LiDAR - Light Detection And Ranging),
- echolokacyjne sondy wielowiązkowe (ang. MBES - Multi Beam Echo Sounder),
- satelitarną interferometrię radarową (ang. InSAR - Interferometric Synthetic Aperture Radar),
- zintegrowany system wykorzystujący do pomiarów pozycjonowanie obiektów w czasie rzeczywistym (ang. RTK - Real Time Kinematic; oraz RTN - Real Time Network),
- urządzenia echolokacji i altimetrii radarowej (np. georadary).

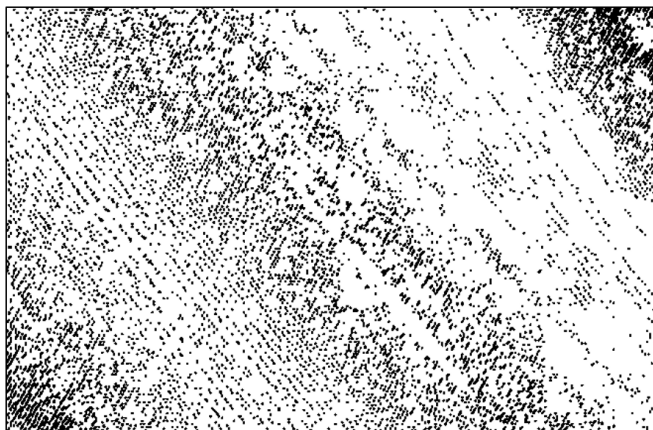
Pośród wielu metod pozyskiwania informacji o terenie, najdokładniejszą i najbardziej wiarygodną grupę wciąż stanowią pomiary bezpośrednie. Ze względu na potrzebę uzyskania maksymalnie aktualnej i dokładnej informacji, pozyskiwanej z dużych obszarów terenu, coraz częściej wykorzystywane są urządzenia pomiarowe pozwalające na bezpośrednią rejestrację wielu milionów punktów pomiarowych w stosunkowo krótkim czasie. Jedną z wiodących technologii zautomatyzowanych pomiarów masowych jest skaning laserowy (LiDAR), który powszechnie wykorzystywany jest w wielu dziedzinach gospodarczych (Wack, Wimmer 2002, Kobler i in. 2007, Liu 2008, Longley i in. 2011). Ze względu na umiejscowienie skanera laserowego w przestrzeni geograficznej, wykorzystujące go technologie, podzielić można na:

- naziemny skaning laserowy – (ang. TLS - Terrestrial Laser Scanning),
- lotniczy skaning laserowy – (ang. ALS - Airborne Laser Scanning),
- satelitarny skaning laserowy – (ang. SLS - Satellite Laser Scanning).

Wszystkie wymienione systemy pomiarowe umożliwiają rejestrację położenia (współrzędne przestrzenne (x,y,z)) oraz dodatkowych atrybutów (np. kolor piksela, charakterystyka widmowa, polirejestracja odbić itp.) wielu milionów punktów w czasie jednej sesji pomiarowej. Rejestracja danych przy użyciu skanera odbywa się dzięki wykorzystaniu

wiązek pomiarowych, których liczba może wahać się od kilku do kilkuset. Są one jednocześnie wysyłane z urządzenia pomiarowego, które pozostaje nieruchome (TLS) lub przemieszcza się wzdłuż mierzonego obiektu (ALS, SLS). Dzięki takiemu rozwiązaniu w jednym cyklu pomiarowym (nadawczo-odbiorczym) pozyskiwanych może być jednocześnie kilkaset punktów w pasie o szerokości kilkuset metrów. Ze względu na to, że obecnie w systemach skanujących częstotliwość wysyłania impulsu wynosi od 50 kHz do 400 kHz (a w skrajnych przypadkach nawet 520 kHz), daje to efektywność pomiaru na poziomie od $2,5 \cdot 10^4$ pkt./s do ponad $2 \cdot 10^5$ pkt./s (czyli do $7,2 \cdot 10^8$ pkt./h) dla jednej głowicy skanującej (RIEGL. Laser Measurement System, <http://www.riegl.com> (dostęp: 15.03.2016)). Przy takiej częstotliwości oraz wielu głowicach skanujących rejestracja położenia obiektów pomiarowych może być dokonywana z prędkościami przekraczającymi półtora miliona punktów na sekundę (Reutebuch i in. 2005, Cavalli i in. 2008, Bretar i in. 2009). Ze względu na to, że wydajność lotniczego skaningu laserowego oceniana jest na kilkadziesiąt kilometrów kwadratowych mierzonej powierzchni w ciągu jednej godziny lotu, ilość zgromadzonych danych w czasie jednej, kilkugodzinnej sesji pomiarowej sięga kilku (N) miliardów rekordów ($N \cdot 10^9$), a wielkości generowanych plików liczone są w TB. Zagęszczenie punktów pomiarowych na jednostkę powierzchni dochodzi do 20 pkt./m² (czyli $2 \cdot 10^7$ pkt./km²), a w skrajnych przypadkach (wiele głowic) może przekraczać 150 pkt./m² (czyli $1,5 \cdot 10^8$ pkt./km²).

Pomimo sprecyzowanej rozdzielczości i parametrów lokalizacyjnych punktów pomiarowych, ostatecznie zasób pomiarowy jest bardzo zróżnicowany i w wielu miejscach niekompletny (rys. 1). Rozmieszczenie punktów pomiarowych jest z reguły niejednakowe w różnych częściach obszaru i zależy dodatkowo od specyficznych cech mierzonego obiektu. Zarejestrowane punkty tworzą niejednorodne zbiory z niejednorodnym zagęszczeniem na różnych obszarach pomiarowych. Dotyczy to w głównej mierze wszelkich niepożądanych odbić wiązki pomiarowej od elementów infrastruktury technicznej (budynki, słupy energetyczne itp.) oraz pokrycia bonitacyjnego danego terenu (drzewa, krzewy uprawy polowe o zwartej strukturze itp.). Po wyeliminowaniu niepożądanych punktów w spektrum pomiarowym powstają luki i przerwy powodowane przesłanianiem wiązki pomiarowej przez różne formy morfologiczne terenu, roślinność lub wystające elementy infrastruktury technicznej, co czyni zasób pomiarowy jeszcze bardziej zróżnicowanym.



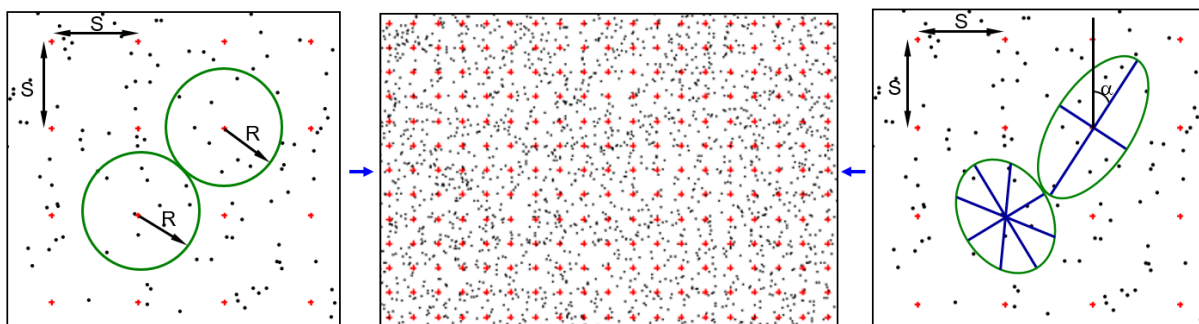
Rys. 1. Przykładowa lokalizacja punktów pomiarowych

Dodatkowo sama kolejność archiwizacji poszczególnych zbiorów często utrudnia dalsze ich przetwarzanie i analizę. Sekwencyjna struktura zapisu poszczególnych rekordów, występująca powszechnie redundancja informacji (pokrywanie się fragmentów obszarów pomiarowych), różne zagęszczenie punktów pomiarowych oraz wielkość generowanych zbiorów nie pozwalają na ich bezpośrednie wykorzystanie w systemach informacji przestrzennej do wygenerowania numerycznego modelu terenu. W celu usprawnienia administrowania dużą ilością danych powszechnie dąży się do uporządkowania struktury przestrzennej i ograniczenia liczby punktów tworzących numeryczny model powierzchni. Jednym ze sposobów uporządkowania struktury przestrzennej punktów tworzących model numeryczny jest wykorzystanie regularnej siatki kwadratów (ang. Regular Grid Squares) tworzących strukturę typu GRID (Maune 2001, Takefusa 2001, Wechsler 2003, Tay i in. 2005, Raaflaub, Collins 2006, Chen i in. 2011).

4.2.2. Struktura GRID

Podczas generowania struktury GRID do wyznaczania wartości w jej węzłach wykorzystywane są algorytmy interpolacyjne. Kluczowym celem algorytmu interpolacyjnego jest wyznaczenie wartości z w regularnych węzłach siatki na podstawie nieregularnie rozłożonych punktów pomiarowych, przy takim sprecyzowaniu parametrów przetwarzania, aby uzyskany wynik był jak najdokładniejszy. Punkty pomiarowe definiowane są w tym przypadku jako zmienne niezależne o współrzędnych (x,y) posiadające wartości funkcji $z=f(x,y)$. Generalnie tworzenie zbiorów struktury GRID, polega na zamianie bazy rekordów pomiarowych (nieregularnie rozmieszczonych punktów pomiarowych) na bazę rekordów obliczeniowych (regularnie rozmieszczonych węzłów) przy wykorzystaniu określonych algorytmów interpolacyjnych. Punkty pomiarowe odnajdywane są w zdefiniowanej przestrzeni poszukiwań wokół nowotworzonych węzłów o ustalonej rozdzielczości S (rys. 2). Przestrzeń poszukiwań może tworzyć okrąg lub elipsę i dodatkowo może zostać podzielona na sektory, w których następuje odszukanie wymaganej liczby punktów. Na podstawie tak zlokalizowanych punktów pomiarowych wybrany algorytm interpolacyjny dokonuje wyznaczenia wartości w każdym punkcie węzłowym. Cały proces wyznaczania węzłów, który nazywany jest również „griddingiem” (Wechsler 2003, Tay i in. 2005, Raaflaub, Collins 2006), można określić jako interpolację wartości funkcji w węzłach regularnej siatki (o współrzędnych x,y) na podstawie wartości funkcji w punktach pomiarowych (o współrzędnych (x,y,z)). Zastosowanie griddingu, z wykorzystaniem znanych wartości obserwacji oraz algorytmów interpolacji, pozwala na określenie wartości tworzących powierzchnię w miejscach, w których nie istnieją oryginalne dane pomiarowe. Po dokonaniu interpolacji wysokości $h = z$ w punktach węzłowych siatki kwadratów, do dalszych działań, zmierzających do wygenerowania DTM, wykorzystuje się jedynie same węzły struktury GRID uzupełnione ewentualnie o linie szkieletowe lub elementy obiektowe. Korzyści płynące z zastosowania struktury GRID są wyraźnie widoczne zwłaszcza podczas przetwarzania megadanych. W tym przypadku znaczne ograniczenie ilości przetwarzanej informacji oraz jej uporządkowanie jest szczególnie pożądane, ponieważ zbyt duża ilość danych pomiarowych nie pozwala na bezpośrednie wygenerowanie DTM na ich podstawie. Ze względu na to, że żaden współczesny system

informacji przestrzennej nie może przetworzyć jednocześnie takiej ilości informacji (dochodzącej do ponad 10^9 rekordów w jednym zbiorze), konieczne staje się dzielenie jej na części. To z kolei nie zawsze jest korzystne ponieważ dodatkowe podzbiory stwarzają problemy podczas ich ponownej agregacji i utrudniają analizowanie większych obiektów powierzchniowych.



Rys. 2. Generowanie struktury GRID; opis w tekście

Ponadto generowanie DTM na podstawie megadanych z pomiarów LiDAR (Kraus, Pfeifer 2001, Reutebuch i in. 2005, Cavalli i in. 2008) czy MBES (Beaudoin i in. 2004, Grządziel 2004, Lihong i in. 2013) jest utrudnione z powodu kierunkowej, liniowej lokalizacji punktów pomiarowych. Powoduje to dodatkowe problemy występujące podczas analizy lokalnych ekspozycji, bilansu mas ziemnych czy interpolacji warstw. W przypadku megadanych utrudnione jest również porównywanie pomiarów dokonanych na tym samym obiekcie w różnym czasie. Zastosowanie struktury GRID do przetwarzania danych pozyskiwanych automatycznie w sposób masowy umożliwia wyeliminowanie większości problemów związanych z ich wykorzystaniem pozwalając w szczególności na:

- ograniczenie nadmiarowości informacji,
- ograniczenie lub całkowitą eliminację redundancji informacji,
- ograniczenie objętości zbiorów przechowywanych w bazach danych,
- kompensację błędów pomiarowych,
- eliminację kierunkowej struktury danych pomiarowych tworzonych przez skanery wielowiązkowe,
- zmniejszenie liczby punktów opisujących powierzchnię podczas generowania DTM,
- płynną regulację rozdzielczości siatki i dostosowanie jej do morfologii analizowanej powierzchni,
- zwiększanie rozdzielczości węzłów na analizowanej powierzchni w celu precyzyjnych analiz,
- dokonywanie analiz w kolejnych epokach pomiarowych przy różnej lokalizacji i zagęszczeniu punktów pomiarowych,
- analizy porównawcze na tych samych punktach węzłowych określonych matematycznie,
- zdynamizowanie i przyspieszenie procesu modelowania danych i ich wizualizacji,

- zwiększenie dynamiki przetwarzania numerycznego danych przez optymalizację algorytmów,
- uproszczenie kierunkowego (profilowego) modelowania i analizowania danych,
- stworzenie przejrzystej organizacji przestrzennej przetwarzanych danych,
- precyzyjne uporządkowanie i organizację struktury topologicznej przechowywanych danych,
- stworzenie wygodniejszego i szybszego dostępu do pożądaných fragmentów informacji w procesach eksploracji danych,
- umożliwienie dogodnej i wydajnej archiwizacji danych,
- możliwość archiwizacji wieloepokowej i odtwarzania modeli na dowolny moment czasowy,
- zwiększenie wydajności przetwarzania informacji przestrzennych w czasie rzeczywistym,
- zwiększenie wydajności w Systemach Zarządzania Bazą Danych (ang. DBMS - DataBase Management System) (Connolly, Begg 2004, Dudek 2006, Soares 2006, Pelikant 2009),
- przyspieszenie procesów Eksploatacji Danych Przestrzennych (ang. SDM - Spatial Data Mining) (Carrier, Povel 2003, Anselin i in. 2006, Huang i in. 2006, Hastie i in. 2009, Miller, Han 2009),
- przyspieszenie procesów przesyłania danych pomiędzy Systemami Informacji Przestrzennej (ang. SIS - Spatial Information System) (O'Sullivan, Unwin 2003, Worboys, Duckham 2004, Harris i in. 2005, Makowski 2005),
- skrócenie czasu dostępu do danych przez sieć komputerową podczas wykorzystania Internetowych Serwerów Danych Przestrzennych (ang. ISDS - Internet Spatial Data Server) (Peng, Tsou 2003, Ha i in. 2006, Dean, Ghemawat 2008),
- uproszczenie struktury zapisu i odczytu informacji w modelach Przestrzennych Baz Danych (ang. SDB - Spatial DataBase) (Longley, Batty 2003, Pigozzi 2004, Masser 2005, Fischer 2006),
- zwiększenie dynamiki Bieżącego Przetwarzania Danych Przestrzennych (ang. SOLAP - Spatial On Line Application Processing) (Rivest i in. 2005, Bédard i in. 2007, Proulx, Bédard 2008),
- zwiększenie możliwości przetwarzania informacji w Dynamicznych Systemach Informacji Przestrzennej (ang. DSIS - Dymamic SIS) (Stroud i in. 2001, Andrienko, Andrienko 2006, Johannesson i in. 2007).

Wymienione właściwości struktury GRID stały się podstawą jej wszechstronnego wykorzystania w proponowanej metodyce na różnych etapach przetwarzania megadanych.

4.2.3. Cel i ogólna charakterystyka pracy

Omawiana praca stanowi propozycję metodyki przetwarzania i analiz danych pozyskiwanych masowo, które następnie są wykorzystywane do numerycznego modelowania powierzchni z zastosowaniem struktury GRID. W celu praktycznej realizacji założeń przedstawionej metodyki oraz uściślenia jej koncepcji opracowano autorski system aplikacji

użytkowych **SYPMEG** (*System of Processing of Megadata with using the GRID structure - System przetwarzania megadanych z wykorzystaniem struktury GRID*). Jest on złożony z 41 autonomicznych modułów (niezależnych aplikacji) umożliwiających zestawienie 81 jednostek funkcjonalnych. Utworzenie całego systemu od podstaw pozwoliło na jego pełną integrację z prezentowaną metodyką, a poszczególne rozwiązania i dostępne opcje systemowe stanowią jej szczegółowe rozwinięcie. W proponowanej metodyce zaprezentowano kompletny tok postępowania podczas przetwarzania megadanych. Przedstawiono zbiór zasad pozwalających na przetworzenie w praktyce dużej ilości danych oraz sposoby ich redukcji i kompresji z zachowaniem wymaganej dokładności opracowania.

Prezentowane autorskie rozwiązania stanowią innowacyjne podejście do przetwarzania dużych zasobów danych oraz analiz dokładności i jakości modeli powierzchni tworzonych przy wykorzystaniu struktury GRID. W systemie *SYPMEG* jest możliwe zarówno wykorzystanie zbiorów wirtualnych w celu wykonania wielu analiz dokładnościowych generowanych modeli powierzchni, jak również przetwarzanie rzeczywistych danych pomiarowych pozyskanych z faktycznych pomiarów. Zgodnie z przyjętym założeniem wszystkie prace badawcze w prezentowanym opracowaniu przebiegają dwutorowo. Dzięki możliwościom systemu *SYPMEG* pozwoliły one objąć zarówno sferę eksperymentalną, jak i praktyczną. W sferze badawczej system *SYPMEG* udostępnia wirtualne środowisko pozwalające na przeprowadzenie kompleksowych analiz dokładności i jakości generowanych wyników obliczeniowych w każdej fazie przetwarzania megadanych. Pozwala to na dobór i ustalenie parametrów przetwarzania megadanych oraz wygenerowanie struktury GRID, która umożliwia utworzenie DTM z założoną dokładnością oraz określoną wydajnością. Możliwa jest także kontrola i definiowanie parametrów uzupełniania, kompresji i archiwizacji utworzonej struktury, co pozwala na wydajne administrowanie wygenerowanym zasobem wynikowym. W każdej fazie procesu wszystkie ustalone parametry przetwarzania mogą być zastosowane w praktyce. W sferze zastosowań praktycznych system *SYPMEG* pozwala na kompleksowe przetworzenie wielu zasobów megadanych według założeń opracowanych w części badawczej. Cały proces przetwarzania danych podzielono na kilka etapów, poczynając od przygotowania i selekcji danych pomiarowych, przez utworzenie struktury GRID wybranymi algorytmami przetwarzania, a kończąc na kompletowaniu, zarządzaniu i archiwizacji utworzonej struktury, która umożliwia ostatecznie wygenerowanie DTM. Dzięki wykorzystaniu zaprezentowanej metodyki oraz zaprojektowanym praktycznym rozwiązaniom systemowym możliwe staje się takie określenie parametrów przetwarzania danych, które prowadzi do uzyskania modelu powierzchni o pożądanej dokładności i jakości.

W pracy skoncentrowano się na powszechnym wykorzystaniu struktury GRID i zastosowaniu jej do przetwarzania megadanych w wielu etapach proponowanej metodyki. Podstawowym zadaniem realizowanym z zastosowaniem struktury GRID była redukcja dużych zbiorów nieuporządkowanych danych pomiarowych oraz zamiana ich w uporządkowane układy punktów węzłowych tworzących numeryczny model powierzchni. W pracy skoncentrowano się na ograniczeniu do minimum liczby punktów tworzących DTM, z jednoczesną kontrolą i zachowaniem ustalonej dokładności generowanego modelu. Dzięki

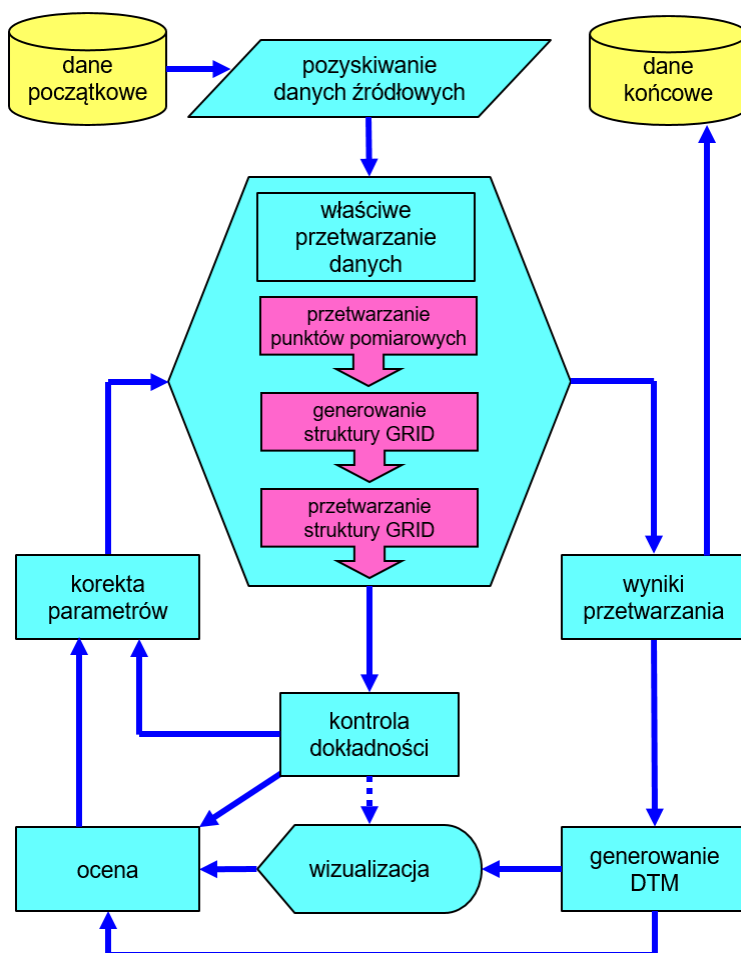
wykorzystaniu struktury GRID możliwe było nie tylko ograniczenie liczby przetwarzanych punktów ale także wyeliminowanie redundancji i nadmiarowości danych. Pozwoliło to na dynamiczne administrowanie wygenerowanymi zasobami oraz uporządkowanie zapisu informacji, co zapewniło wydajną archiwizację i kompresję danych tworzących model. Generowana struktura GRID może być zastosowana wszędzie tam, gdzie najważniejsze jest dynamiczne i szybkie modelowanie przestrzeni, bieżąca aktualizacja jej parametrów oraz prowadzenie analiz w czasie rzeczywistym w kolejnych epokach pomiarowych. Dzięki niej ułatwiony i przyspieszony został proces przeszukiwania baz danych oraz wybór pożądanej i jednocześnie uporządkowanej informacji. Zastosowanie struktury GRID pozwoliło także na zwiększenie dynamiki przetwarzania numerycznego danych przez optymalizację algorytmów oraz przyspieszenie wzajemnego przekazywania danych między modułami obliczeniowymi systemu *SYPMEG*. Wykorzystanie zalet uporządkowanej struktury węzłów pozwoliło na ustalenie spójnych zasad przetwarzania informacji przez poszczególne jednostki funkcjonalne prezentowanego systemu oraz określenie jednolitej komunikacji między nimi. Tego rodzaju działania umożliwiły z kolei osiągnięcie dużej wydajności obliczeniowej oraz sprawności analitycznej na każdym etapie przetwarzania. Strukturę GRID wykorzystano w systemie *SYPMEG* zarówno na etapie wstępnego przygotowania danych pomiarowych, właściwego generowania węzłów, jak i późniejszego ich przetwarzania oraz archiwizacji. Na każdym z etapów pozwoliło to zwiększyć dynamikę i przyspieszyć proces modelowania danych oraz ich wizualizacji. Dodatkowe korzyści płynące z wykorzystania struktury GRID są widoczne również podczas prowadzenia różnego typu analiz dokładności dokonywanych w pracy na poszczególnych etapach przetwarzania. Tak zorganizowana struktura danych umożliwia łatwiejsze generowanie powierzchni i wyznaczanie żądanych wielkości w dowolnym punkcie czy przekroju w stosunkowo krótkim czasie. Dało to możliwość zastosowania wydajnych algorytmów modelowania, wizualizacji 3D oraz wszechstronnych analiz porównawczych wykonywanych w tych samych punktach węzłowych. Strukturę GRID wykorzystano także podczas oceny jakości i dokładności numerycznych modeli powierzchni. Dzięki utworzeniu diagramów różnicowych i wygenerowaniu współczynników statystycznych możliwe było zbadanie dokładności generowanych DTM oraz dokonanie analiz procesów iteracyjnych. Wykorzystanie struktury GRID umożliwiło również eliminację kierunkowego układu danych pomiarowych pozyskiwanych przez skanery wielowiązkowe oraz płynną regulację rozdzielczości siatki, pozwalając na dostosowanie jej do lokalizacji przestrzennej punktów pomiarowych i ukształtowania morfologicznego analizowanej powierzchni. Regulacja rozdzielczości siatki węzłów tworzących powierzchnię umożliwiła przeprowadzenie szczegółowych badań z wykorzystaniem swobodnie definiowanych wirtualnych przestrzeni roboczych. Powszechne wykorzystanie struktury GRID we wszystkich fazach procesu pozwoliło na zachowanie spójnych metod postępowania na każdym etapie przetwarzania. Umożliwiło to ustandaryzowanie wymiany danych między poszczególnymi jednostkami funkcjonalnymi systemu oraz pozwoliło na ujednolicenie kontroli dokładności we wszystkich fazach procesu.

Podzielenie całego procesu przetwarzania na wiele elementów składowych pozwoliło na wydzielenie głównych etapów postępowania oraz stworzenie koncepcji metodycznej i

technologicznej rozwiązania określonych zadań. W złożonym procesie przetwarzania megadanych wyodrębniono kilka kluczowych etapów, które składają się na cały cykl przetwarzania:

- pozyskiwanie danych źródłowych;
- właściwe przetwarzanie danych zawierające trzy fazy realizacji procesu:
 - przetwarzanie punktów pomiarowych,
 - generowanie struktury GRID,
 - przetwarzanie struktury GRID;
- kontrola i ocena dokładności wyników etapów pośrednich;
- dobór i korekta parametrów przetwarzania;
- tworzenie i formatowanie wyników przetwarzania;
- generowanie numerycznego modelu terenu;
- wizualizacja wyników przetwarzania;
- ocena końcowa procesu przetwarzania.

W celu zaprezentowania metodyki oraz toku postępowania w omawianej pracy wykorzystano schematy szczegółowo objaśniające kolejne fazy przetwarzania. Ogólny schemat metodyki przetwarzania danych (rys. 3) pokrywa się ze schematem realizacji etapów procesu przetwarzania megadanych w systemie *SYPMEG* (rys. 7).



Rys. 3. Schemat proponowanej metodyki przetwarzania danych

Integracja metodyki i systemu pozwoliła na ustalenie zależności między poszczególnymi etapami oraz wypracowanie kompletnych metod przetwarzania megadanych jak również spójnego sposobu analiz i kontroli otrzymanych wyników. Wykorzystanie jednostek funkcjonalnych systemu *SYPMEG* umożliwiło prezentację wielu przykładów, które pozwalają zilustrować zastosowanie przedstawianej metodyki do badań naukowych. Z kolei ścisłe powiązanie metodyki postępowania ze schematem funkcjonalnym systemu umożliwiło opracowanie wielu praktycznych rozwiązań. Zaproponowane rozwiązanie metodyczne i aplikacyjne obejmuje szereg czynności szczegółowych, których realizacja dokonywana jest w kolejnych etapach:

1) na etapie pozyskiwania danych źródłowych:

- wykorzystanie zbiorów pomiarów zarejestrowanych na obiektach rzeczywistych,
- generowanie zbiorów punktów pseudopomiarowych jako danych do wyznaczenia parametrów przetwarzania i analiz dokładności,
- generowanie raportów;

2) na etapie przetwarzania punktów pomiarowych (pp):

- import danych pomiarowych,
- kontrolę poprawności rekordów pomiarowych,
- konwersję danych źródłowych,
- zmianę formatu archiwizacji danych;
- stworzenie roboczych zbiorów poglądowych,
- wybór obszarów przetwarzania i zasięgu opracowania,
- analizę lokalizacji *pp*,
- eliminację i filtrację niepożądanych *pp*,
- przetwarzanie iteracyjne i wsadowe wielu zbiorów danych,
- generowanie raportów;

3) na etapie generowania struktury GRID:

- ustalanie stref lokalizacji węzłów,
- wyznaczenie obszarów do uzupełnień,
- ustalenie parametrów konstrukcyjnych siatki,
- dobór rozmieszczenia i rozdzielczości węzłów w zależności od lokalizacji *pp*,
- dobór rozmieszczenia i rozdzielczości węzłów w zależności od morfologii terenu,
- ustalenie przestrzeni poszukiwań *pp* wokół węzła,
- dobór metod lokalizacji sektorowej *pp* wokół węzła,
- generowanie lokalizacji siatki bazowej struktury GRID,
- dobór algorytmów interpolacyjnych,
- wyznaczenie struktury GRID przez zastosowanie globalnej aproksymacji powierzchni,
- generowanie struktury GRID przez dokonanie właściwej interpolacji lokalnej,
- wykorzystanie kombinacji algorytmów interpolacyjnych i sekwencji interpolacji,
- eliminację i filtrację niepożądanych węzłów,
- iteracyjne przetwarzanie danych,

- dobór parametrów przetwarzania,
 - przetwarzanie wsadowe wielu zbiorów danych,
 - generowanie raportów;
- 4) na etapie przetwarzania struktury GRID:
- porządkowanie i sortowanie siatki węzłów,
 - wyznaczanie i uzupełnianie brakujących węzłów wewnętrznych struktury GRID,
 - wyznaczanie i uzupełnianie węzłów brzegowych,
 - kompresję i dekompresję struktury GRID,
 - przetwarzanie zbiorów z wielu epok pomiarowych,
 - rejestracja zmian powierzchni z wielu epok w czasie,
 - archiwizację wielu epok pomiarowych,
 - odtwarzanie kształtu powierzchni na dowolny moment czasowy,
 - kontrolę poprawności archiwizowanych danych,
 - ustalenie formatu zapisu informacji,
 - konwersję danych wynikowych,
 - eksport danych do SIS,
 - iteracyjne przetwarzanie danych,
 - dobór parametrów przetwarzania,
 - przetwarzanie wsadowe wielu zbiorów danych,
 - generowanie raportów;
- 5) na etapie kontroli dokładności:
- tworzenie teoretycznej siatki węzłów,
 - obliczenie błędów dla szeregu zbiorów praktycznych,
 - klasyfikację przedziałową liczebności błędów,
 - generowanie danych do diagramów różnicowych,
 - wyznaczanie współczynników wykorzystywanych do analiz dokładności,
 - generowanie danych do analiz procesów iteracyjnych,
 - generowanie raportów;
 - zestawienie analiz w celu kompleksowej oceny dokładności DTM,
 - iteracyjne przetwarzanie danych,
 - manualną i automatyczną korektę parametrów przetwarzania,
 - dobór parametrów przetwarzania,
 - przetwarzanie wsadowe wielu zbiorów danych;
- 6) na etapie wizualizacji wyników przetwarzania:
- generowanie numerycznych modeli powierzchni w dowolnej konwencji przetwarzania przestrzennych współrzędnych punktów (trójpolowe rekordy XYH),
 - generowanie i wizualizację DTM dla modeli wirtualnych lub rzeczywistych,
 - generowanie i wizualizację obiektów punktowych XYZ,
 - generowanie i wizualizację obiektów powierzchniowych,
 - generowanie i wizualizację obiektów hybrydowych (punktowo-powierzchniowych),

- generowanie i wizualizację warstwic, spadków, nachyleń, ekspozycji itp.,
- generowanie i wizualizację przekrojów (profilu),
- wizualizację diagramów różnicowych,
- wizualizację współczynników dokładnościowych,
- wizualizację wykresów do analiz procesów iteracyjnych,
- wizualizację zbiorów z wielu epok pomiarowych,
- wizualizację zmian wartości w czasie,
- przygotowanie danych do analiz szeregiem czasowymi.

Podczas definiowania metodyki i konstruowania systemu *SYPMEG* poczyniono wiele założeń, które starano się zrealizować za pomocą kolejnych jednostek funkcjonalnych. W tym celu właściwy proces przetwarzania megadanych podzielono na trzy niezależne fazy, w których dla różnych grup danych następuje realizacja założeń prezentowanej metodyki. W każdej fazie wykorzystano określone jednostki funkcjonalne przedstawianego systemu, pozwalające zarówno na przetwarzanie, jak i analizy wybranej grupy danych. Cały proces przetwarzania megadanych może być wykonany dla swobodnie definiowanej wirtualnej (lub rzeczywistej) przestrzeni roboczej oraz może zachodzić w różny sposób, zależny od ukształtowania morfologicznego powierzchni, różnych parametrów konstrukcyjnych struktury GRID czy wielu odmiennych parametrów algorytmów interpolacyjnych. W związku z tym liczba przypadków możliwych do przeanalizowania z zastosowaniem systemu *SYPMEG* jest praktycznie nieograniczona. Wykorzystane w omawianym opracowaniu przykłady miały na celu pokazanie sposobów realizacji wybranych analiz oraz wykonanie interpretacji określonych wyników dla wyselekcjonowanych przypadków, które są reprezentatywne dla głównych procesów realizowanych w systemie. Prezentowane w omawianej pracy analizy miały na celu ukazanie wybranych możliwości z wielu istniejących wariantów postępowania podczas przetwarzania różnych danych systemem *SYPMEG*. Umożliwiają one prześledzenie całego procesu badawczego oraz pozwalają na zademonstrowanie sposobów doboru głównych parametrów pracy poszczególnych algorytmów obliczeniowych w kolejnych fazach procesu przetwarzania megadanych.

4.2.4. Dobór narzędzi informatycznych

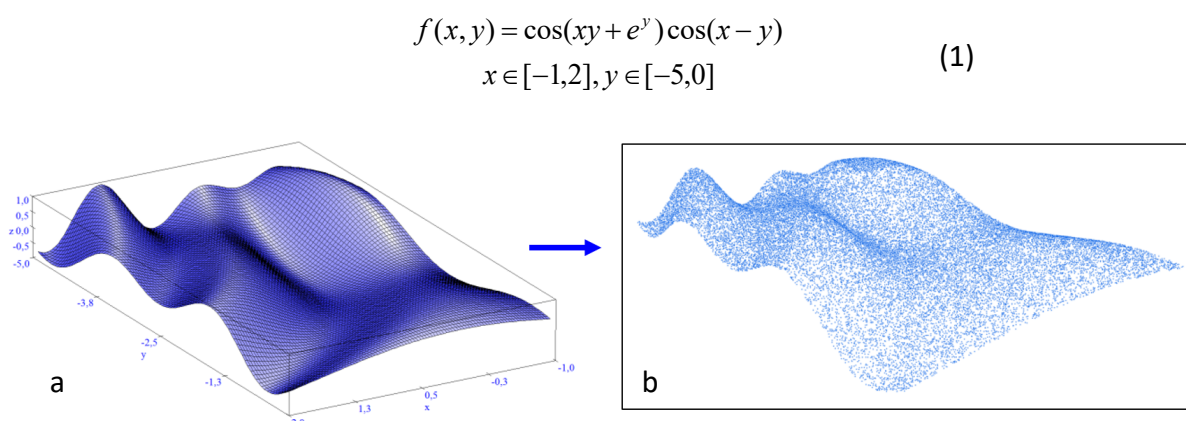
W celu realizacji zamierzonych zadań istotnym elementem był dobór narzędzi informatycznych do przetwarzania wyjątkowo dużej ilości danych. W wyborze tym kierowano się założeniami funkcjonalnymi, które koncentrowały się na wydajności obliczeniowej zapewniającej szybkie przetworzenie megadanych z maksymalną automatyzacją procesu przetwarzania. Ponadto podczas wyboru narzędzi informatycznych zwrócono również uwagę na to, że zastosowane rozwiązania powinny pozwalać na zaimplementowanie tworzonej aplikacji do większości dostępnych wydajnych platform programowo-sprzętowych. Za najwydajniejsze współczesne rozwiązania informatyczne (zarówno pod względem obliczeniowym jak i wielkości przestrzeni adresowej) powszechnie uważa się trzy technologie (Blythe 2005, Travostino 2006, Hindman i in. 2009, Bu i in. 2010, Voorsluys i in. 2011, Amies i in. 2012):

- Cluster Computing (CC),
- Grid Computing (GC),
- Cloud Computing (CLC).

W celu przygotowania systemu do wykorzystania wymienionych technologii dobór narzędzi uwzględniał kryterium możliwości projektowania złożonych struktur obliczeniowych oraz wykorzystanie programowania wieloprocesorowego, wielowątkowego i równoległego, z jednoczesnym operowaniem jak największą ilością danych umieszczonych w pamięci operacyjnej. Zastosowanie języka Fortran (Kott, Walczak 1991, Trykozko 1999, Chrobak 2003, Rościszewski, Witt 2003, Sobieski 2008, Press i in. 2014) do zaprogramowania głównych operacji obliczeniowych pozwoliło na spełnienie wymaganych założeń. Wybór tego języka programowania został podyktowany również powszechną dostępnością kompilatorów dla wersji 32- i 64-bitowej. Pod uwagę wzięto również ustawiczny rozwój i unowocześnianie implementacji języka z zachowaniem wstecznej kompatybilności, co pozwala na przenoszenie istniejących aplikacji na coraz nowsze platformy programowo-sprzętowe. Dzięki temu opracowany system *SYPMEG* może być nadal rozwijany i unowocześniany. Jego modułowa konstrukcja pozwala na sukcesywną rozbudowę, aktualizowanie, zastępowanie i dodawanie poszczególnych jednostek wykonawczych. Za pomocą języka Fortran utworzono wszystkie jednostki obliczeniowe systemu *SYPMEG*, a wykorzystanie tych samych przestrzeni adresowych umożliwiło wydajne przekazywanie informacji między nimi. Dzięki wzbogaceniu języka programowania środowiskiem Visual Studio Shell było możliwe stworzenie graficznych interfejsów komunikacji z użytkownikiem. Dzięki pracy w formularzach graficznych generujących pliki sterujące możliwe jest określanie parametrów przetwarzania poszczególnych modułów obliczeniowych systemu *SYPMEG* w czytelny sposób. Podzielenie systemu na wydzielone jednostki obliczeniowe (środowisko Fortran) oraz formularze konfiguracyjne (środowisko Visual) pozwoliło zachować kompromis między wydajnością przetwarzania, a wygodą definiowania parametrów sterujących przetwarzaniem. W procesie programowania i doboru algorytmów do rozwiązania kluczowych problemów wykorzystane zostało ponad 30 letnie doświadczenie autora w tworzeniu aplikacji użytkowych służących przetwarzaniu danych numerycznych.

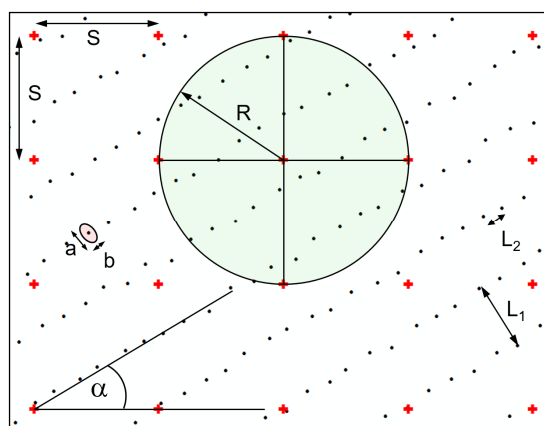
4.2.5. Założenia metodyki i systemu

Prezentowana metodyka uwzględnia kompletny proces przetwarzania, zarówno rzeczywistych danych pomiarowych, jak i symulowanych. Przetwarzanie danych pomiarowych daje możliwość zastosowania systemu *SYPMEG* do realizacji określonych zadań produkcyjnych, natomiast wykorzystanie danych symulujących pomiar rzeczywisty pozwala na wykonywanie wszechstronnych analiz. Wyspecjalizowany podzespół pozyskiwania danych źródłowych umożliwia wygenerowanie ustalonej liczby wirtualnych punktów pomiarowych (punktów pseudopomiarowych) na podstawie swobodnie definiowanych funkcji matematycznych dwóch zmiennych $z = f(x, y)$ (przykładowo funkcja (1); rys. 4 a). Wykorzystanie w badaniach funkcji matematycznych daje nieograniczone możliwości kształtowania wirtualnej przestrzeni roboczej (rys. 4 b).



Rys. 4. Generowanie punktów pseudopomiarowych (b) na podstawie funkcji (a)

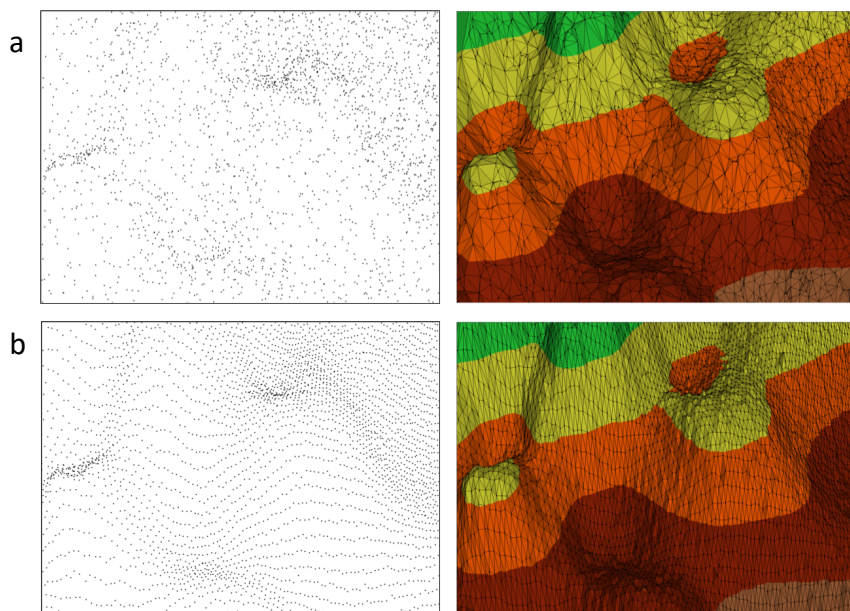
Z kolei możliwość generowania punktów pseudopomiarowych o takiej samej charakterystyce lokalizacyjnej, jak oryginalne punkty pomiarowe uzyskane techniką LiDAR czy MBES (rys. 5), pozwala na przeprowadzenie kompleksowych analiz. W systemie *SYPMEG* możliwe jest między innymi definiowanie kąta nachylenia rzędów zawierających punkty pseudopomiarowe, regulowanie odległości rzędów (L_1) oraz punktów w rzędzie (L_2) czy określanie zaburzenia (a, b) współrzędnych poziomych x, y oraz współrzędnej pionowej h .



Rys. 5. Parametry definiujące lokalizację punktów pseudopomiarowych; opis w tekście

W celu zademonstrowania metodyki analiz w pracy wykorzystano wiele różnorodnych wirtualnych przestrzeni roboczych, pozwalających wygenerować dane symulacyjne o różnej charakterystyce i lokalizacji przestrzennej. W zależności od ustawień parametrów przetwarzania numeryczny model powierzchni może charakteryzować się określonymi cechami wynikającymi z różnej lokalizacji punktów pseudopomiarowych. W przypadku wyboru punktów rozproszonych (rys. 6 a) powstaje charakterystyczna nieregularna siatka triangulacyjna struktury TIN. Po wygenerowaniu uszeregowanych punktów pseudopomiarowych (rys. 6 b) model powierzchni posiada morfologię taką, jak w przypadku tworzenia DTM na podstawie pomiarów LiDAR. Wykorzystanie wirtualnych przestrzeni roboczych pozwoliło na uzyskanie danych do precyzyjnej oceny i kontroli dokładności na

każdym etapie przetwarzania megadanych. Porównanie wartości praktycznej wyznaczonej na podstawie punktów pseudopomiarowych w określonym procesie obliczeniowym (interpolacji, aproksymacji, kompresji itp.) z obliczoną z funkcji wartością teoretyczną w dowolnie wybranych punktach wirtualnej przestrzeni roboczej dało możliwość uzyskania błędów prawdziwych. Pozwoliło to na jednoznaczne określenie dokładności generowanych modeli powierzchni.



Rys. 6. Przykłady różnej lokalizacji punktów pseudopomiarowych i wygenerowanych na ich podstawie modeli powierzchni; a – punkty rozproszone, b – punkty uszeregowane

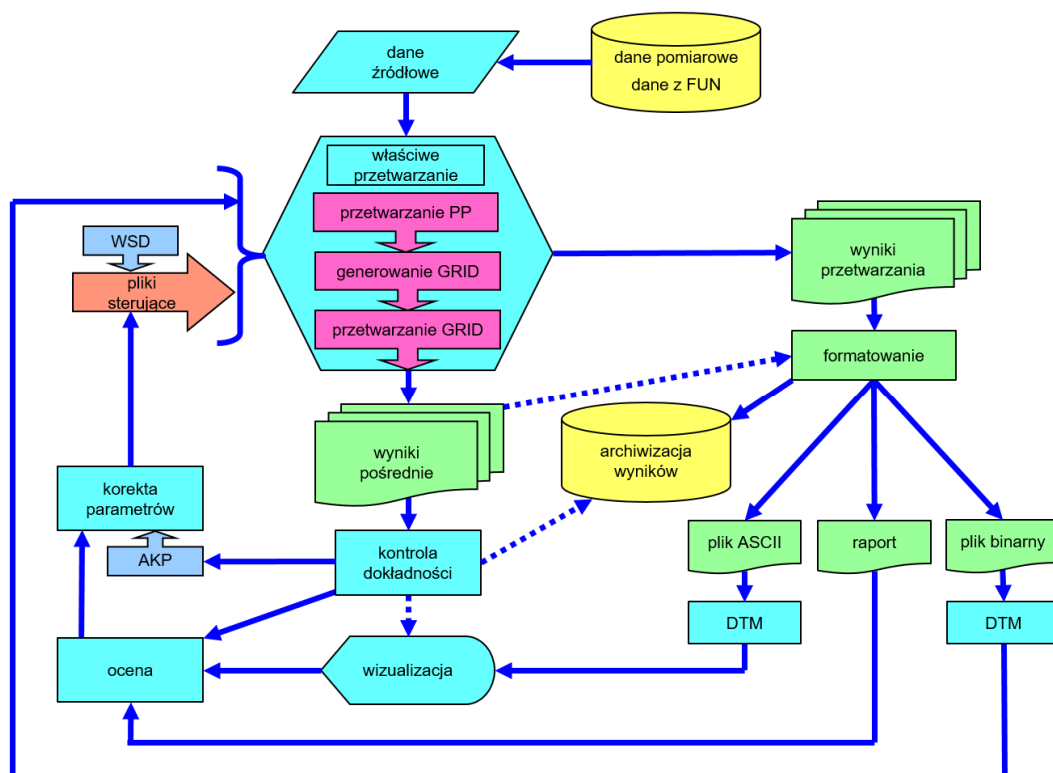
Wykorzystanie punktów pseudopomiarowych do wygenerowania struktury GRID różnymi algorytmami interpolacyjnymi pozwoliło ustalić wpływ wielu elementów (np. rozmieszczenia punktów pomiarowych czy morfologii powierzchni) na parametry konstrukcyjne tworzonej siatki węzłów. Wykorzystanie modeli wirtualnych umożliwiło również ustalenie parametrów przetwarzania określonych algorytmów. Parametry te mogą być następnie wykorzystane podczas obliczeń wykonywanych na danych rzeczywistych. Dzięki kombinacji parametrów funkcji analitycznej oraz przedziałów jej zmienności można dobrać wirtualną przestrzeń roboczą w taki sposób, aby reprezentowana przez nią powierzchnia była podobna do kształtu rzeczywistego terenu. Ostatecznie zaprezentowane rozwiązanie umożliwia wygenerowanie wirtualnej przestrzeni roboczej (dowolnie zróżnicowanej lub podobnej do przestrzeni rzeczywistej), a następnie wykorzystanie jej do przeanalizowania określonych parametrów przetwarzania ustalonych algorytmów w kolejnych fazach procesu określonych w metodyce. Pozwala to na wykonanie analizy poszczególnych faz procesu i określenie, w jakim stopniu dany etap wpływa na końcową jakość generowanego DTM.

Możliwość precyzyjnej kontroli modeli wirtualnych na każdym etapie procesu pozwala na opracowanie spójnych metod postępowania, które ostatecznie mogą być wykorzystane do przetwarzania informacji pochodzącej z rzeczywistych pomiarów. Dzięki kontroli dokładności wykonywanej w wybranym etapie jest możliwe wypracowanie właściwych sekwencji

obliczeniowych dla tego etapu, które następnie zostaną wykorzystane do przetwarzania danych rzeczywistych. W procesie kontrolnym może być również zweryfikowana dokładność rzeczywistych powierzchni DTM wygenerowanych na podstawie faktycznych pomiarów. W takim przypadku jako dane wzorcowe, zamiast danych wygenerowanych z funkcji matematycznej, mogą być wykorzystane dane z rzeczywistych pomiarów, które są dokładniejsze od danych wykorzystanych w procesie interpolacji. Dane tego typu mogą pochodzić z obliczeń wykonanych na przykład dla większego zagęszczenia punktów pomiarowych lub z innego, dokładniejszego pomiaru wykonanego inną metodą i potraktowanego jako zasób wzorcowy. Można je także uzyskać z pomiarów istniejących obiektów na danym obszarze, których lokalizację określono dokładniej (np. przez pomiary klasyczne lub pomiary TLS) niż dane wykorzystywane do interpolacji GRID. Mogą także pochodzić z innych przestrzennych baz danych lub z innego DTM utworzonego na analizowanym obszarze w innym czasie i inną dokładniejszą metodą.

4.2.6. Konstrukcja systemu

W celu realizacji założeń metodyki oraz przeprowadzenia stosownych analiz w systemie *SYPMEG* wykorzystano jego specjalną konstrukcję, która umożliwiła sekwencyjną pracę komponentów, przetwarzanie wsadowe oraz przetwarzanie iteracyjne. W celu podniesienia wydajności systemu podczas przetwarzania megadanych oraz uwzględnienia dużego stopnia automatyzacji procesu przetwarzania i kontroli wyników zastosowano wiele autorskich rozwiązań odpowiadających za realizację określonych zadań. System jest złożony z wielu niezależnych jednostek wykonawczych (modułów), z których każda odpowiada za pewien wycinek realizowanego procesu przetwarzania (rys. 7).



Rys. 7. Realizacja etapów procesu przetwarzania megadanych w systemie *SYPMEG*

W celu realizacji wydzielonych etapów cały system został podzielony na cztery specjalizowane grupy zadań funkcjonalnych, z których pierwsze trzy wykonywane są przez zintegrowane podzespoły systemu, a czwarta realizowana jest przez aplikację zewnętrzną:

1) Pozyskiwanie danych źródłowych.

Na tym etapie następuje rejestracja i archiwizacja pomiarów rzeczywistych (*pp*) lub wygenerowanie punktów pseudopomiarowych (*ppp*) w oparciu o model wirtualny.

2) Właściwe przetwarzanie (zawierające 3 fazy realizacji procesu):

- [przetwarzanie PP],
- [generowanie GRID],
- [przetwarzanie GRID].

Na tym etapie każda faza realizowana jest przez jednostki obliczeniowe oparte na algorytmach wykorzystujących strukturę GRID i pracujące w sekwencyjnych układach segmentów złożonych z wielu modułów systemu.

3) Kontrola dokładności.

Na tym etapie wyniki kontroli dokładności pozwalają na kompleksową ocenę i analizę jakości generowanych modeli i są wykorzystywane do weryfikacji parametrów przetwarzania w wybranych modułach.

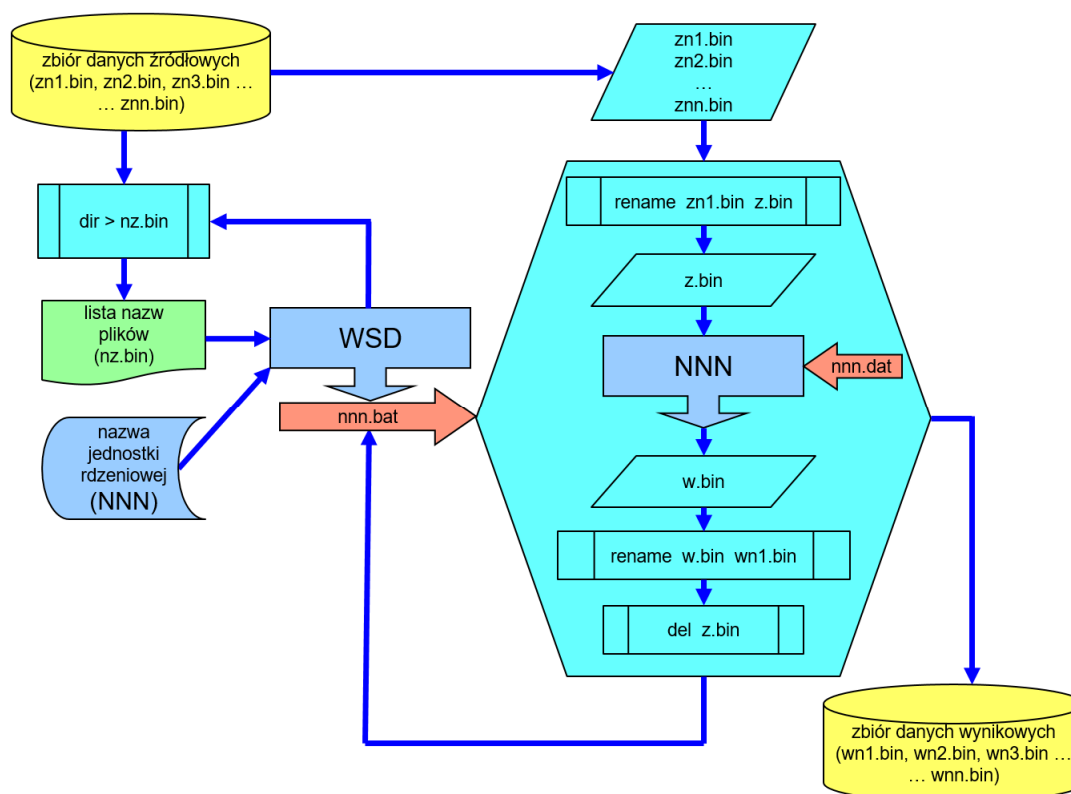
4) Generowanie DTM i wizualizacja wyników przetwarzania.

Na tym etapie proces tworzenia numerycznego modelu powierzchni oraz jego prezentacji (wizualizacji) zależy od możliwości wykorzystywanej aplikacji zewnętrznej.

Ze względu na to, że poszczególne moduły stanowią autonomiczne jednostki funkcjonalne, można je wykorzystywać w sposób niezależny lub zestawiać w sekwencje, w których przypadku kilka modułów realizuje określone zadanie obliczeniowe. Dodatkowo modułowa konstrukcja systemu umożliwia wielokrotne użycie wybranego modułu w określonej sekwencji przetwarzania. Ujednolicony format plików wymiany informacji między modułami pozwala na dowolną kolejność pracy poszczególnych elementów składowych. Modułowa i otwarta budowa systemu umożliwia także zastępowanie wybranych jednostek funkcjonalnych przez aplikacje zewnętrzne, co pozwala na wykorzystanie ich możliwości obliczeniowych w wybranych fazach procesu przetwarzania. Ponadto wykorzystanie autonomicznych modułów do budowy systemu umożliwia niezależną pracę wszystkich podzespołów, co z kolei daje możliwość wykonywania zdefiniowanych działań w tym samym czasie na jednym lub kilku stanowiskach komputerowych. Wyniki przetwarzania określonego modułu mogą stanowić wyniki pośrednie lub końcowe oraz być wykorzystane jako dane źródłowe następnego modułu w sekwencji. W ten sposób zostaje zapewniona iteracyjna praca systemu *SYPMEG*, a dzięki kontroli dokładności w każdym cyklu jest możliwa sukcesywna korekta parametrów przetwarzania.

Wykorzystanie w systemie przetwarzania wsadowego pozwoliło na znaczną automatyzację wielu procesów obliczeniowych (rys. 8). Mogą one przebiegać jednocześnie w niezależnych przestrzeniach roboczych, wydzielonych dla dowolnej sekwencji jednostek obliczeniowych. Przetwarzanie wsadowe pozwala na sukcesywne uruchamianie kolejnych modułów pracujących w sekwencji, a dzięki znormalizowaniu nazw zbiorów źródłowych i

wynikowych umożliwia spójne przekazywanie danych między nimi. Wykorzystanie poleceń wsadowych pozwala również na automatyczne przekształcanie nazw wielu plików oraz tworzenie sekwencji ich przetwarzania przez wybrany element systemu. W kolejnych cyklach dokonywana jest zamiana nazw poszczególnych plików roboczych, na stałe nazwy plików wykorzystywanych przez jednostkę rdzeniową [NNN], oraz realizowane jest ich sukcesywne przetwarzanie. Wygenerowany w taki sposób plik wsadowy [nnn.bat] pozwala na całkowitą automatyzację procesu przetwarzania wielu zbiorów roboczych przez określoną jednostkę rdzeniową przy zachowaniu niezmiennych ustawień jej parametrów przetwarzania. W efekcie sekwencje kolejnych poleceń zapisanych w pliku wsadowym realizują automatycznie cały proces przetwarzania dowolnej ilości zbiorów danych przekształcając je w pliki wynikowe. Zbiór tak wygenerowanych plików wynikowych, może stanowić dane źródłowe dla kolejnego etapu przetwarzania danych systemem *SYPMEG*. Umożliwia to całkowitą automatyzację procesu przetwarzania dowolnej ilości zbiorów zlokalizowanych w tej samej przestrzeni roboczej. Dostępne w systemie *SYPMEG* połączenie wsadowego i sekwencyjnego przetwarzania danych umożliwia automatyczne przetworzenie praktycznie nieograniczonej ilości danych.



Rys. 8. Schemat przetwarzania wsadowego realizowanego dla wielu zbiorów

Ponadto system *SYPMEG* zaprojektowano w taki sposób, aby wybrane jednostki funkcjonalne umożliwiały pracę iteracyjną. Dzięki kontroli dokładności w każdym cyklu przetwarzania iteracyjnego jest możliwa sukcesywna korekta parametrów na poszczególnych etapach przetwarzania. Pozwoliło to na uzyskanie za każdym razem rezultatów spełniających

wymogi założeń dokładnościowych poczynionych na wstępie przetwarzania. Wyniki kontroli dokładności mogą być wykorzystane do sukcesywnej korekty parametrów przetwarzania poszczególnych jednostek wykonawczych systemu *SYPMEG* przechowywanych w plikach sterujących. Umożliwia to sterowanie całym przebiegiem przetwarzania megadanych i pozwala na określenie najkorzystniejszego zestawu głównych parametrów w danym procesie obliczeniowym. Korekta parametrów, wykonywana dzięki kontroli dokładności, może przebiegać w sposób manualny lub automatyczny. Podczas przetwarzania kolejne iteracje pozwalają na wygenerowanie wielu wyników pośrednich, z których jest wybierane najlepsze rozwiązanie. Dzięki znormalizowaniu plików sterujących w całym systemie *SYPMEG* (wszystkie pliki mają taką samą organizację zapisu przechowywanych parametrów przetwarzania) jest możliwy uniwersalny sposób bezpośredniej zmiany każdej z przechowywanych w nich wartości. Pozwoliło to na jednorodne sterowanie całym procesem przetwarzania i ułatwiło uzyskanie wielu danych kontrolnych umożliwiających ocenę dokładności generowanych modeli powierzchni. W ocenie tej były również wykorzystywane raporty, które mogą być utworzone przez poszczególne moduły systemu podczas każdej operacji obliczeniowej, co pozwoliło na dodatkową analizę wszystkich wyników pośrednich.

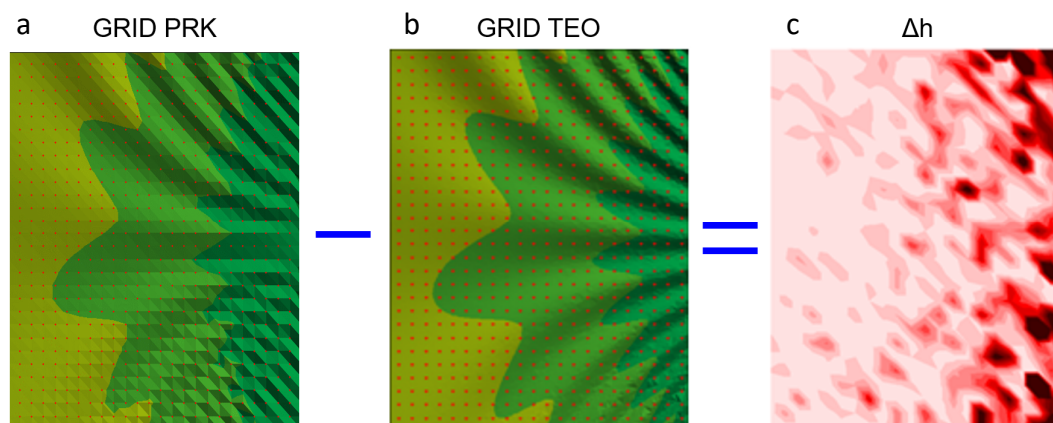
W systemie *SYPMEG* każdy segment jest wydzielonym blokiem logicznym i może pracować niezależnie od pozostałych lub być ujęty w określoną sekwencję przetwarzania. Jednostki obliczeniowe we wszystkich segmentach realizują zadania dzięki sekwencyjnym układom komponentów, które mogą być zbudowane z wielu modułów wykonawczych. Cały proces przetwarzania megadanych może być realizowany dla swobodnie definiowanej wirtualnej lub rzeczywistej przestrzeni roboczej. Może on być dokonywany w różny sposób, zależnie od ukształtowania morfologicznego powierzchni oraz od parametrów konstrukcyjnych struktury GRID. Dodatkowo proces przetwarzania może być zrealizowany iteracyjnie dla każdego parametru zdefiniowanego w formularzu konfiguracyjnym wykorzystywanego algorytmu, przy dowolnie określanych kombinacjach modułów lub ich grup. W związku z powyższym ilość przypadków możliwych do przeanalizowania przy wykorzystaniu systemu *SYPMEG* jest praktycznie nieograniczona.

4.2.7. Kontrola i analiza dokładności

W celu przebadania parametrów pracy algorytmów przetwarzających megadane wykorzystano ujednoliconą dla wszystkich etapów kontrolę dokładności. Jednakowy format danych pozwolił na zastosowanie takich samych procedur kontrolnych dla każdego weryfikowanego przypadku. Uniwersalna kontrola dokładności umożliwiła zarówno ocenę jakości tworzonych modeli powierzchni, jak i dobór najkorzystniejszych parametrów przetwarzania w wybranych fazach procesu. Dzięki kontroli dokładności generowanych modeli na każdym etapie możliwe stało się wypracowanie założeń zmierzających do uzyskania jak najkorzystniejszego rozwiązania analizowanych problemów. Zastosowane w prezentowanej metodyce mechanizmy kontroli pozwoliły na wygenerowanie danych umożliwiających realizację trzech rodzajów analiz.

Pierwszą grupę stanowiły diagramy różnicowe, które pozwoliły na wizualną ocenę jakości generowanych modeli oraz określenie obszarów o różnym stopniu dopasowania do

powierzchni wzorcowej. W celu skonstruowania diagramu na każdym węźle struktury GRID obliczane były różnice $\Delta h = \text{GRID PRK} - \text{GRID TEO}$ jako różnice pomiędzy wartością teoretyczną (wyznaczoną z funkcji), a wartością interpolacyjną (wyznaczoną przez dany algorytm) (rys. 9). Dzięki wizualizacji wartości błędów w poszczególnych punktach struktury GRID możliwe było przeanalizowanie zasięgu obszarów o różnej wielkości zniekształceń, występujących na wygenerowanych modelach powierzchni. Pozwoliło to na określenie szczególnych założeń dotyczących przetwarzania takich obszarów lub wyeliminowanie ich z procesu obliczeniowego wykonywanego danym algorytmem.



Rys. 9. Zasada tworzenia diagramów różnicowych; a – model interpolacyjny, b – model teoretyczny, c – diagram różnicowy

Drugą grupę stanowiły współczynniki statystyczne umożliwiające liczbowe określenie dokładności tworzonych modeli powierzchni. Ze względu na dużą ilość przetwarzanych danych do analiz wykorzystujących współczynniki wykorzystano podział wszystkich wyników na kolejne przedziały klasowe. Pozwoliło to na zastosowanie szczególnej postaci wzorów umożliwiających analizy dokładnościowe oraz przyspieszenie procesu kontroli. Zastosowane współczynniki pozwalają na wybór najdokładniej wygenerowanego modelu powierzchni spośród modeli tworzonych przez wybraną do badań procedurę obliczeniową. W celu oceny dokładności generowanych powierzchni interpolacyjnych w systemie *SYPMEG* wykorzystano szereg współczynników statystycznych (Jóźwiak, Podgórski 2000, Paradysz 2005, Schabenberger, Gotway 2005, Sobczyk 2006, Suchecki 2010):

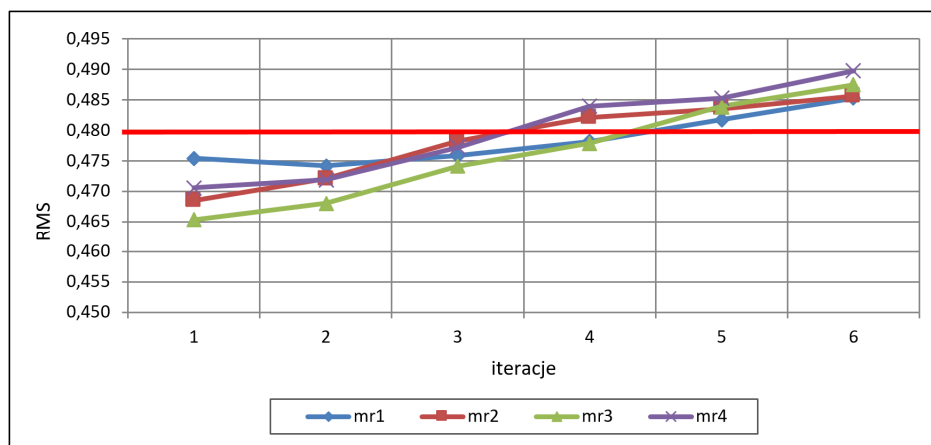
- liczebność skumulowana (dla wartości rzeczywistych i bezwzględnych),
- skumulowane częstości względne (dystrybuanta empiryczna),
- średnia arytmetyczna zwykła,
- średnia arytmetyczna wagowana,
- łączna wartość błędów ,
- kwartyle,
- rozstęp ćwiartkowy,
- mediana,
- dominanta,
- wariancja,

- estymator odchylenia standardowego,
- standaryzowane wartości błędów,
- trzeci moment centralny,
- asymetria rozkładu wartości błędów,
- pozycyjny współczynnik asymetrii,
- współczynnik skośności,
- współczynnik RMS (Root Mean Square Error).

W pracy dokonano szczegółowej interpretacji wartości poszczególnych współczynników w celu porównania dokładności interpolacyjnych modeli powierzchni. Dodatkową weryfikację wyników można wykonać, porównując jednocześnie diagramy różnicowe utworzone dla poszczególnych algorytmów. Dzięki wykorzystaniu diagramów różnicowych można wyeliminować z obliczeń obszary charakteryzujące się ekstremalnymi błędami, które zakłócają analizę prowadzoną wyłącznie z zastosowaniem omawianych współczynników. Z przeprowadzonych badań wynika, że połączenie obu metod pozwala na obiektywne porównanie różnych algorytmów obliczeniowych i wybór najlepszego rozwiązania.

Trzecią grupę stanowiła analiza procesów iteracyjnych, która umożliwiła ustalenie tendencji zmian dokładności modeli podczas cyklicznej modyfikacji parametrów przetwarzania. Możliwości iteracyjnej pracy systemu *SYPMEG* pozwalają na powtarzanie w cyklu operacji obliczeniowych przez wybrane algorytmy. Podczas takich operacji jest możliwa zmiana określonych parametrów wpływających na dokładność obliczeń, która może być wykonywana manualnie lub w sposób zautomatyzowany. Dzięki temu możliwe było uzyskanie wielu modeli powierzchni, z których każdy miał odmienne parametry przetwarzania i charakteryzował się różną dokładnością. Pozwoliło to z kolei na wygenerowanie zestawu wartości współczynników dokładnościowych, które można przedstawić na wykresie zbiorczym oraz porównać i określić trend zmian. Procesy iteracyjne umożliwiły przebadanie określonych parametrów przetwarzania danego algorytmu dla wszystkich pośrednich modeli powierzchni, co ostatecznie pozwoliło na wybór najlepszego rozwiązania, z jednoczesnym ustaleniem parametrów pozwalających utworzyć dany model. Możliwości pracy iteracyjnej oraz sukcesywny dobór poszczególnych parametrów przetwarzania pozwoliły na uzyskanie wielu zbiorczych zestawień umożliwiających porównanie dokładności obliczeń w różnych konfiguracjach. Ułatwiło to ustalenie właściwych parametrów pracy określonych algorytmów na różnych etapach procesu przetwarzania. Na rysunku 10 przedstawiono przykładowe wykresy zmian wartości współczynnika RMS, wyznaczonego po interpolacji powierzchni tym samym algorytmem dla czterech różnych modeli morfologicznych (*mr1* – *mr4*), w sześciu iteracjach. W kolejnych iteracjach zmniejszano rozdzielczość GRID poprzez zwiększanie (o 1 m) wielkości boku kwadratu bazowego siatki. Dopuszczalną granicę błędu ustalono dla RMS = 0,48 m. Dla poszczególnych modeli morfologicznych wykonywano kolejne iteracje (zwiększając wielkość kwadratu bazowego), aż do przekroczenia ustalonej granicy błędu. W przypadku modeli *mr2* i *mr4* (bardziej zróżnicowanych morfologicznie) przekroczenie granicy błędu nastąpiło w czwartej iteracji, a dla modeli *mr1* i *mr3* (mniej zróżnicowanych

morfologicznie) w piątej. Dalsze zwiększanie siatki (zmniejszanie rozdzielczości GRID) jest w przypadku określonych modeli niepożądane. Zestawienie wyników kontroli dokładności z kolejnych iteracji pozwoliło wybrać ustawienia wielkości GRID dla każdego z modeli, spełniające poczynione na wstępie założenia dotyczące wygenerowania jak najdokładniejszego modelu dla danej powierzchni przy użyciu jak najmniejszej ilości danych opisujących model.



Rys. 10. Zmiany wartości współczynnika RMS podczas kolejnych iteracji

Połączenie wszystkich metod oceny: wizualizacyjnej (diagramy różnicowe), wskaźnikowej (współczynniki statystyczne) oraz analiz iteracyjnych (trendy), pozwala na kompleksowe porównanie dokładności i jakości wygenerowanych modeli powierzchni oraz wybór najkorzystniejszego wariantu przetwarzania danych. Ostatecznie dzięki możliwościom systemu *SYPMEG* można przeanalizować wpływ określonego parametru przetwarzania na dokładność generowanego modelu powierzchni. W procesie analiz i kontroli dokładności każdy z analizowanych parametrów może być korygowany. Kolejne wyniki przetwarzania są kontrolowane, co pozwala na utworzenie szeregu diagramów różnicowych oraz obliczenie wielu współczynników statystycznych umożliwiających ocenę dokładności generowanych modeli powierzchni. Przetwarzanie wielu zbiorów pozwala na wykonanie analiz iteracyjnych dla wielu obiektów roboczych i utworzenie zestawów danych porównawczych umożliwiających przeanalizowanie zmian dokładności modeli na wykresach zbiorczych. Analiza tego typu zestawień pozwala ocenić wpływ procesu zmian danego parametru na dokładność generowanego modelu i wybór najkorzystniejszego rozwiązania.

4.2.8. Przetwarzanie danych

Zasadnicza grupa zadań funkcjonalnych prezentowanej metodyki obejmuje właściwe przetwarzanie danych. Danymi źródłowymi (wejściowymi) tego etapu przetwarzania mogą być zarówno oryginalne rzeczywiste zbiory pomiarowe jak i zbiory punktów pseudopomiarowych. Zadanie to realizowane jest przez wydzielony podzespół właściwego przetwarzania i stanowi zasadniczy etap przetwarzania megadanych (rys. 7). Proces realizowany w tym etapie jest wielowątkowy i może obejmować szerokie spektrum przetwarzania i analiz różnego rodzaju danych, pośród których można wyodrębnić kilka zasadniczych grup. W celu dostosowania

systemu *SYPMEG* do założeń metodyki oraz realizacji określonych założeń funkcjonalnych cały proces etapu właściwego przetwarzania został podzielony na trzy niezależne fazy realizujące oddzielne zadania związane z inną grupą danych:

- 1) faza przetwarzania i przygotowania punktów,
- 2) faza tworzenia struktury GRID,
- 3) faza administrowania utworzoną strukturą GRID.

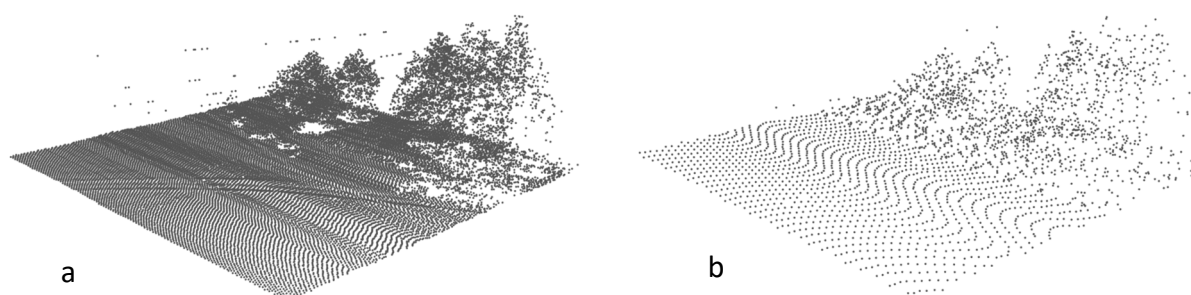
Wyodrębnienie poszczególnych faz pozwoliło na ścisłe zrealizowanie w praktyce założeń metodyki postępowania i opracowanie wyspecjalizowanych segmentów systemu biorących udział w realizacji każdego etapu właściwego przetwarzania. Końcowe wyniki przetwarzania poszczególnych elementów systemu mogą stanowić dane źródłowe następnej jednostki obliczeniowej. Pozwala to na zautomatyzowanie realizacji kompletnego procesu obliczeniowego poczynając od analizy danych źródłowych do wygenerowania ostatecznego DTM i administrowania nim. Wszystkie wyniki przetwarzania mogą być przed archiwizacją każdorazowo formatowane do postaci plików binarnych lub tekstowych plików ASCII (rys. 7). Oba rodzaje plików mogą być następnie wykorzystane jako dane źródłowe w kolejnym etapie przetwarzania oraz posłużyć do wygenerowania DTM w dowolnej aplikacji zewnętrznej. Wszystkie wyniki pośrednie podlegają kontroli dokładności i są wykorzystywane do sterowania korektą parametrów przetwarzania. Sukcesywne dokonywanie kontroli dokładności na każdym z etapów pozwala na wybór najkorzystniejszych parametrów przetwarzania, które następnie mogą zostać wykorzystane do przetwarzania rzeczywistych danych pomiarowych.

4.2.8.1. Przetwarzanie punktów pomiarowych

W pierwszej fazie procesu były przetwarzane punkty pomiarowe. Na wstępie wykonano analizę zarejestrowanych zbiorów obserwacji, co pozwoliło na eliminację błędów rejestracji poszczególnych rekordów oraz umożliwiło ustalenie jednorodnej struktury zapisu przechowywanej informacji. W celu minimalizacji ilości przetwarzanych danych i osiągnięcia maksymalnej wydajności w systemie *SYPMEG* są wykorzystywane wyłącznie współrzędne (x,y,h) każdego punktu pomiarowego. Z tego powodu dane pomiarowe archiwizowane z zastosowaniem różnej struktury rekordów były konwertowane do ujednoliconej postaci stosowanej w systemie we wszystkich jednostkach funkcjonalnych. Jednakowa struktura danych pozwoliła na usprawnienie przekazywania informacji między jednostkami systemu oraz umożliwiła zastosowanie ujednoliconej kontroli dokładności na różnych etapach przetwarzania. Na tym etapie dzielono również duże zbiory na fragmenty o zdefiniowanej wielkości. Możliwość deklarowania liczby jednocześnie przetwarzanych rekordów pozwoliła na ograniczenie ilości danych przechowywanych w jednym pliku. Z kolei regulacja wielkości zasobów przekazywanych do przetwarzania w następnych etapach pozwoliła na sterowanie wydajnością procesu.

W następnym etapie przewidziano wstępną analizę zbiorów obserwacji, które ze względu na dużą ilość danych i ich charakterystykę lokalizacyjną nie mogą być wczytane bezpośrednio do systemu ich wizualizacji. Wstępna analiza zasobów pozwoliła na określenie rozmieszczenia danych w przestrzeni pomiarowej, organizację ich zapisu w poszczególnych

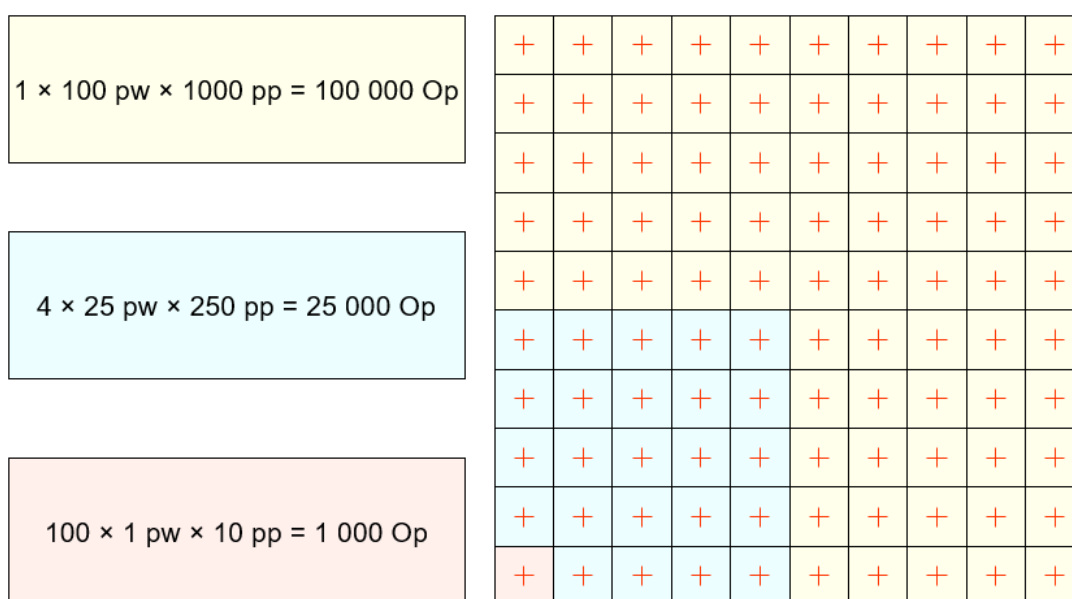
zbiorach oraz identyfikację miejsc nieciągłości rejestracji danych. Zadania te zrealizowano przez utworzenie zbiorów poglądowych, w których zredukowano liczbę rekordów pomiarowych, z zachowaniem zakresu obszarowego oraz logiki i kolejności archiwizacji danych (rys. 11). Zaproponowane metody pozwalają na wczytanie złożonych zasobów składających się z wielu plików do aplikacji wizualizującej i wykonanie dowolnych analiz lokalizacji danych pomiarowych. Umożliwia to także skontrolowanie układu przestrzennego sekwencji wielu pomiarów na tym samym obszarze oraz pozwala na wybranie z całego spektrum pomiarowego wyłącznie pożądaných zasobów. Robocze zbiory poglądowe umożliwiają również wstępną identyfikację infrastruktury technicznej, pokrycia bonitacyjnego oraz ogólną analizę morfologii powierzchni. Ułatwia to dobór odpowiednich algorytmów przetwarzania dla kompletnych zbiorów danych.



Rys. 11. Redukcja liczby punktów pomiarowych; a – przed redukcją, b – po redukcji

W kolejnym etapie zwrócono uwagę na ustalenie zakresu obszarowego opracowania. Ze względu na to, że informacja o danym obszarze może być przechowywana w różnych zbiorach pomiarowych, często zachodzi konieczność przeskanowania całego zasobu danych pod kątem odnalezienia rekordów pomiarowych spełniających zdefiniowane wymogi przestrzenne. Dodatkowo ograniczenia dotyczące ilości przetwarzanych danych często uniemożliwiają odczytanie wszystkich informacji z pliku (lub wielu plików) jednocześnie. Zastosowana w systemie *SYPMEG* automatyzacja procesu wyboru punktów pochodzących z tego samego obszaru pomiarowego i przechowywanych w wielu zbiorach roboczych pozwala na wybranie z całego spektrum pomiarowego określonej przestrzeni roboczej zawierającej komplet informacji pomiarowej na danym obszarze. Punkty są wyszukiwane niezależnie dla każdej ze współrzędnych (x,y,h) . Daje to możliwość ograniczenia ilości danych, które są jednocześnie przetwarzane, oraz podziału całego zasobu na podobszary robocze o ustalonej wielkości. Podział dużego zasobu danych na fragmenty i kolejne ich przetwarzanie podczas generowania struktury GRID pozwala na przeprowadzenie wydajniejszych operacji obliczeniowych niż w przypadku takich operacji wykonanych na jednym zbiorze. Wydzielenie wielu podobszarów roboczych pozwala na ograniczenie liczby operacji wyszukiwawczych wykonywanych przez algorytmy obliczeniowe w jednym przebiegu i znaczne skrócenie czasu obliczeń. Na rysunku 12 przedstawiony zostały teoretyczne fragmenty obszarów podlegających przetwarzaniu w jednym cyklu obliczeniowym. Zakładając że na badanym obszarze zarejestrowano 1000 punktów pomiarowych (1000 *pp*) i przewidziano 100 punktów

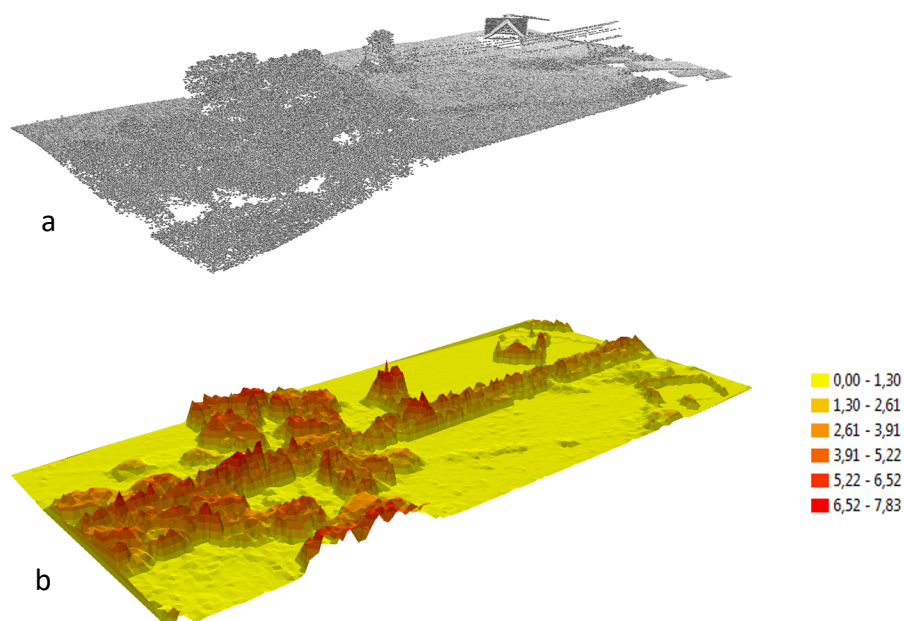
węzłowych (100 *pw*), przeanalizowanie wszystkich punktów pomiarowych w stosunku do każdego węzła wymaga $1 \cdot 10^5$ operacji obliczeniowych (*Op*). Po podzieleniu całego fragmentu na cztery mniejsze podobszary (po 25 węzłów każdy), liczba operacji obliczeniowych potrzebnych do przeanalizowania wszystkich danych zostaje zredukowana do $2,5 \cdot 10^4$. Kolejne zmniejszenie wielkości podobszarów jednoczesnego przetwarzania danych (100 podobszarów z 1 węzłem) redukuje liczbę operacji obliczeniowych do 1000. Dodatkowo zastosowanie w systemie *SYPMEG* możliwości sekwencyjnego przetwarzania danych zgodnie z ich naturalnym porządkiem rejestracji daje możliwość ich bezpośredniego przetwarzania bez wstępnego wybierania zasięgu opracowania. Pozwala to na swobodne dostosowanie ilości jednocześnie przetwarzanej informacji przez określoną jednostkę systemu oraz bieżące kontrolowanie wydajności całego procesu.



Rys. 12. Ustalenie podobszarów przetwarzanych w jednym cyklu obliczeniowym

Analiza rozmieszczenia punktów pomiarowych wykonywana w następnym etapie pozwala na ustalenie wskaźników określających charakterystykę przestrzenną ich lokalizacji. Podczas generowania siatki węzłów na podstawie punktów pomiarowych istotne jest określenie ich zagęszczenia oraz zróżnicowania pionowego. Analiza tych wartości ułatwia dobór podstawowych parametrów konstrukcyjnych struktury GRID oraz wpływa na ustalenie parametrów przetwarzania algorytmów obliczeniowych. W pracy do sprecyzowania równomiernego rozmieszczenia przestrzennego wskaźników określających charakterystykę przestrzenną lokalizacji punktów na badanym obiekcie wykorzystano siatkę węzłów o dowolnie regulowanej rozdzielczości. Wykorzystanie systemu *SYPMEG* pozwoliło na wykonanie wielu analiz dotyczących rozmieszczenia punktów pomiarowych wokół wybranych grup węzłów, których lokalizację można następnie wykorzystać do wygenerowania docelowego numerycznego modelu powierzchni. Wyznaczenie wskaźników zagęszczenia (przyporządkowanych do każdego węzła) pozwoliło na sporządzenie diagramów umożliwiających wykonanie analiz ich lokalizacji na badanym obiekcie. Ułatwiły one taki dobór

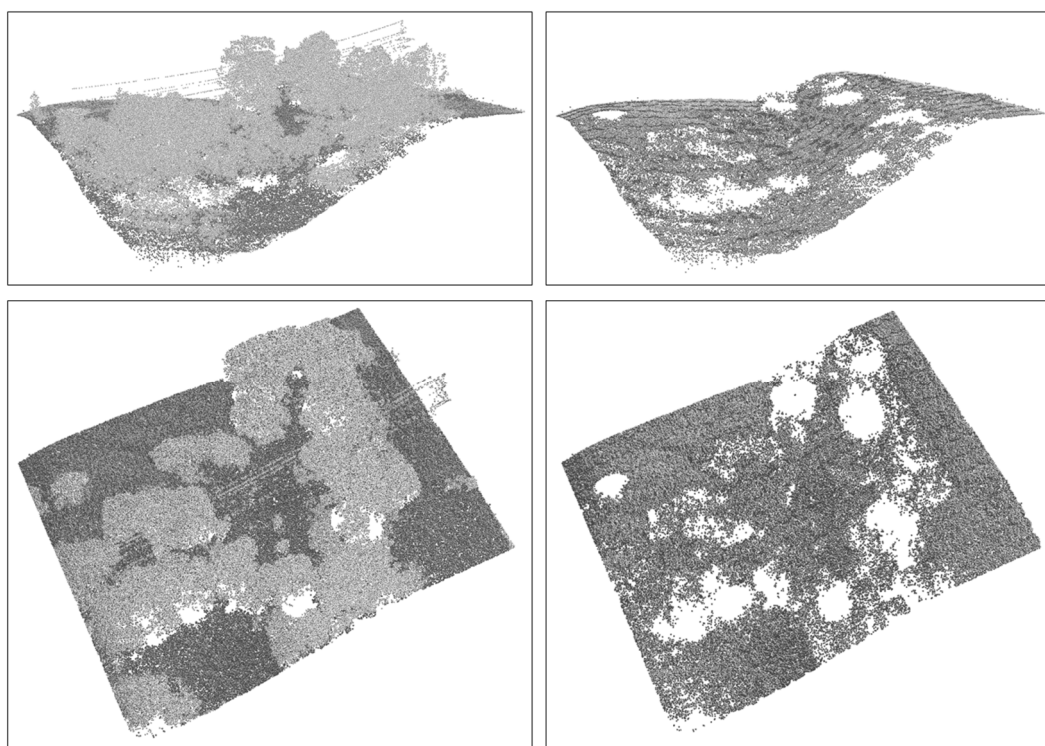
parametrów przetwarzania, który umożliwia osiągnięcie jednorodnej dokładności modelu na całym obszarze. Wskaźniki zagęszczenia mogą również posłużyć do określenia lokalizacji zdublowanych obszarów pomiarowych, a ich przyporządkowanie do siatki węzłów pozwala na wydajną wizualizację takich stref. Odpowiednie zdefiniowanie wskaźników zagęszczenia umożliwia również wyznaczenie obszarów, w których mogą być zastosowane określone algorytmy interpolacyjne, do obliczeń potrzebujące odpowiedniej liczby właściwie rozmieszczonych punktów pomiarowych. Z kolei wyznaczenie wskaźnika dyspersji jako miary rozproszenia pionowego i przypisanie go do swobodnie definiowanej lokalizacji węzłów umożliwia klasyfikację obszarów o różnym zróżnicowaniu pionowym punktów pomiarowych. Wyniki przykładowych analiz przedstawia rysunek 13, na którym zaprezentowano odpowiadającą badanemu obszarowi mapę hipsometryczną. Ustalenie przedziałów klasowych pozwoliło na zobrazowanie kolorami wartości wskaźników dyspersji. Pozwala to na ustalenie obszarów, w których zróżnicowanie pionowe punktów wykorzystywanych do interpolacji węzłów struktury GRID jest największe. Dzięki temu możliwa jest identyfikacja infrastruktury technicznej i pokrycia bonitacyjnego badanego obiektu, co z kolei ułatwia dobór parametrów filtracji niepożądanych punktów ze zbioru pomiarowego.



Rys. 13. Punkty pomiarowe (a) i analiza wskaźnika dyspersji (b) na badanym obiekcie

Kolejny etap przetwarzania megadanych stanowiła filtracja zbiorów pomiarowych. Pozwoliła na wyselekcjonowanie z zasobów pomiarowych punktów leżących na mierzonej powierzchni topograficznej i jednocześnie wyeliminowanie punktów zarejestrowanych po odbiciu wiązki pomiarowej od obiektów naziemnej infrastruktury technicznej czy pokrycia bonitacyjnego. W prezentowanej pracy, zgodnie z założeniami prezentowanej metodyki, zaproponowano metody filtracji pozwalające na maksymalne zautomatyzowanie tego procesu. W systemie *SYPMEG* realizację tego zagadnienia podzielono na dwa etapy.

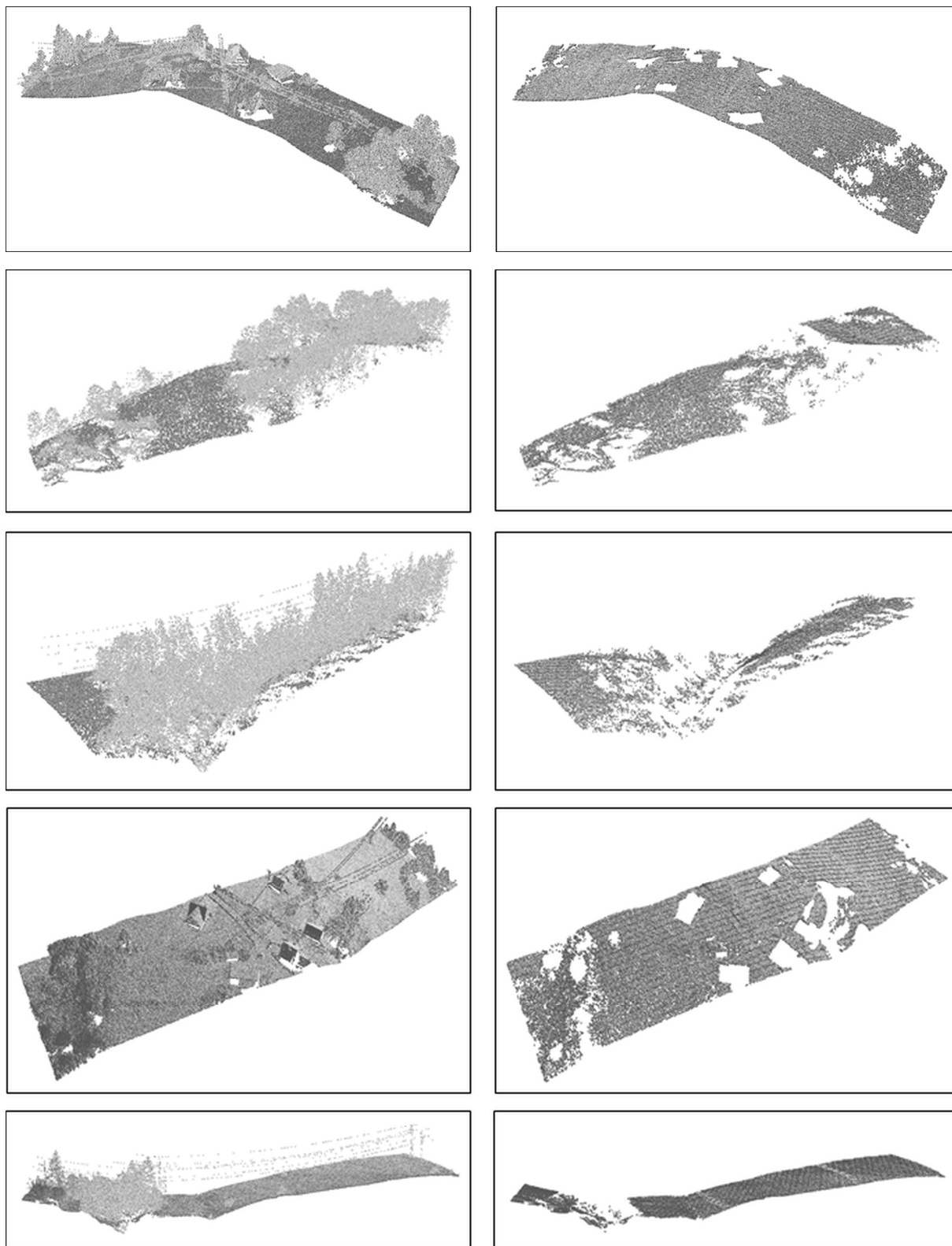
W pierwszym z nich zastosowano filtrację globalną, która pozwoliła na wydajne przetwarzanie dużych obszarów pomierzonego obiektu w jednym procesie obliczeniowym. Zakłada ona wykorzystanie funkcji globalnej, tj. płaszczyzny lub wielomianu o automatycznie dobieranym stopniu, do aproksymacji kształtu analizowanej powierzchni. Zastosowanie płaszczyzny jest ograniczone do niewielkich obszarów, które można traktować jak teren płaski. Pozwala to jednak na najszybszą eliminację niepożądanych punktów pomiarowych. W przypadku powierzchni bardziej zróżnicowanych morfologicznie (rys. 14) zastosować można wielomian aproksymacyjny. Umożliwia on wykonanie aproksymacji powierzchni zakrzywionych na obszarach, które można zakwalifikować do podstawowych form morfologicznych, np. stok, wzgórze, niecka, siodło. Obszary podstawowe można również wydzielić z całego spektrum pomiarowego. We wszystkich przypadkach są eliminowane punkty, których odległość od powierzchni aproksymacyjnych przekracza ustaloną wartość.



Rys. 14. Obszar pomiarowy przed filtracją (lewa strona) i po filtracji (prawa strona) w różnych ujęciach perspektywicznych

Eliminacja globalna umożliwia w wydajny sposób przefiltrowanie zbioru danych pomiarowych i wyeliminowanie z niego określonych grup punktów, które nie leżą na powierzchni topograficznej. Pozwala to na wstępną filtrację zasobu pomiarowego, która w przypadkach małego skomplikowania morfologicznego powierzchni jest wystarczająca. Automatyzacja procesu z zastosowaniem funkcji globalnej pozwoliła na uzyskanie wysokiej wydajności przetwarzania oraz umożliwiła wstępną filtrację zasobu pomiarowego. Filtracja taka w przypadkach małego skomplikowania morfologicznego powierzchni jest wystarczająca. Jeżeli mierzona powierzchnia była bardziej zróżnicowana, stosowano filtrację lokalną.

W drugim etapie filtracji była wykonywana właściwa i precyzyjna filtracja zasobu pomiarowego. Pozwoliła ona na uwzględnienie lokalnych zniekształceń analizowanego obszaru, który w takim przypadku może być zróżnicowany morfologicznie w dowolnym stopniu. W tym celu cała analizowana powierzchnia jest dzielona na podobszary robocze, których wielkość jest określana przez zdefiniowanie długości boku kwadratu bazowego struktury GRID. Możliwość ustalenia dowolnej rozdzielczości siatki węzłów pozwala na dokładne dopasowanie lokalnie tworzonych powierzchni aproksymacyjnych do kształtu morfologicznego mierzonego obiektu. Dzięki podzieleniu procesu eliminacji odstających punktów pomiarowych na podetapy (realizowane w jednej sekwencji przetwarzania) osiągnięto dużą dokładność filtracji, z zachowaniem wysokiej wydajności obliczeniowej oraz minimalnego udziału operatora. Dodatkowo dzięki pracy iteracyjnej oraz przetwarzaniu wsadowemu system *SYPMEG* umożliwia automatyczne przetwarzanie wielu zbiorów, co jeszcze bardziej podnosi wydajność procesu filtracji. Jednoczesne uwzględnienie w procesie filtracji parametrów lokalizacji pionowej oraz dopuszczalnej wartości maksymalnej dyspersji punktów pomiarowych umożliwia precyzyjny wybór punktów zlokalizowanych na szukanej powierzchni topograficznej lub w jej bezpośrednim sąsiedztwie. Na ich podstawie było możliwe wygenerowanie lokalnych (zależnych od rozdzielczości GRID) powierzchni aproksymacyjnych, które stanowiły podstawę eliminacji niepożądanych punktów pomiarowych. W każdym przypadku wokół każdego węzła była wyznaczana powierzchnia aproksymacyjna z automatycznie dobieranym stopniem wielomianu. Stopień wielomianu zależał od lokalizacji i liczby odnalezionych wokół węzła punktów pomiarowych. Automatyczny dobór najlepszego rozwiązania dla danego układu przestrzennego punktów pomiarowych pozwolił w każdym przypadku na dokładne dopasowanie wielomianu do lokalnej powierzchni topograficznej. Sam proces filtracji przebiegał dwustopniowo. W pierwszym etapie wartość dopuszczalnej odległości punktów pomiarowych od powierzchni aproksymacyjnej ustalano globalnie, co pozwoliło na wstępną eliminację najbardziej oddalonych punktów. W drugim etapie na bazie zbiorów przetworzonych w pierwszym etapie wykonywano ostateczną filtrację. Umożliwia ona automatyczne ustalenie maksymalnej dopuszczalnej odległości punktu pomiarowego od powierzchni aproksymacyjnej na podstawie wyliczanych lokalnie współczynników rozproszenia pionowego. Wszystkie punkty pomiarowe, które nie spełniały zdefiniowanych założeń, były eliminowane ze zbioru wynikowego. Pokazane na rysunku 15 przykłady w różnych ujęciach perspektywicznych pozwalają ocenić jakość dokonanej filtracji na obiektach pomiarowych o różnej morfologii, zróżnicowanym pokryciu roślinnością i obiektami technicznymi oraz różnym zagęszczeniu punktów pomiarowych. Zaletą prezentowanej metody dostępnej w systemie *SYPMEG* jest jej całkowite zautomatyzowanie podczas przetwarzania dużej ilości danych pomiarowych. Jednorazowe ustawienie parametrów przetwarzania pozwala na automatyczną filtrację zbiorów pomiarowych zarejestrowanych na obszarach o różnym stopniu zróżnicowania morfologicznego. Wykorzystanie takiego uniwersalnego sposobu przetwarzania pozwala na realizację procesu przy różnym zagęszczeniu punktów pomiarowych zarejestrowanych dla obiektów o zróżnicowanym pokryciu bonitacyjnym oraz różnorodnej infrastrukturze technicznej.

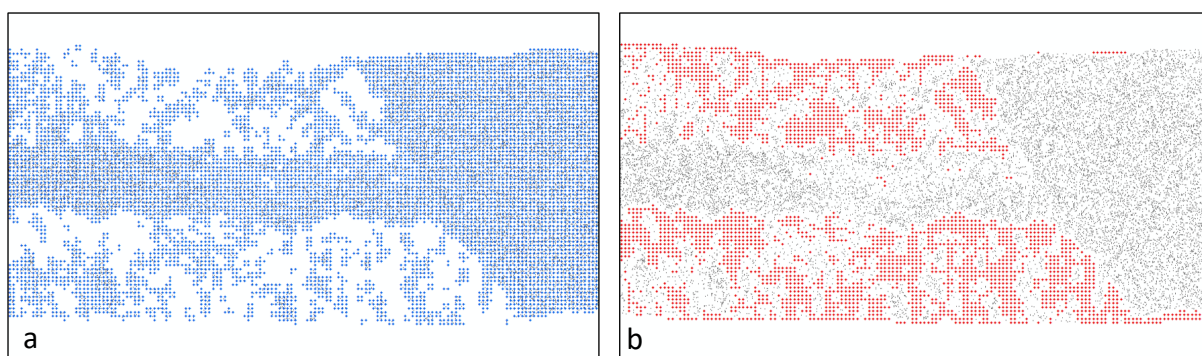


Rys. 15. Przykłady zastosowania automatycznej filtracji na różnych obiektach pomiarowych (lewa strona - przed filtracją, prawa strona - po filtracji)

Filtracja zbiorów pomiarowych kończy pierwszą fazę procesu przetwarzania megadanych. Dane wynikowe operacji obliczeniowych pierwszej fazy stanowią dane źródłowe dla drugiej fazy procesu.

4.2.8.2. Generowanie struktury GRID

W drugiej fazie procesu przetwarzania megadanych były generowane struktury GRID. Analizy prowadzone w tej fazie pozwoliły na ustalenie parametrów konstrukcyjnych siatki węzłów oraz na dobór algorytmów interpolacyjnych o ustalonych kryteriach przetwarzania. W pierwszym etapie nastąpiło ustalenie stref potencjalnej lokalizacji węzłów oraz obszarów do uzupełnień. Węzły tworzące strukturę GRID można najdokładniej wygenerować w strefach występowania punktów pomiarowych. W związku z tym analiza ich rozmieszczenia pozwala określić potencjalne obszary lokalizacji węzłów oraz strefy, w których punkty pomiarowe nie występują (rys. 16). Sprawia to, że dokładność obliczeń może być w nich mniejsza lub interpolacja w ogóle nie będzie możliwa do przeprowadzenia. Możliwość zdefiniowania już na tym etapie przetwarzania dowolnej rozdzielczości siatki węzłów oraz lokalizacji punktów do potencjalnej interpolacji pozwala wydzielić strefy, w których można wyznaczyć węzły algorytmem o określonych wymaganiach co do lokalizacji punktów pomiarowych.



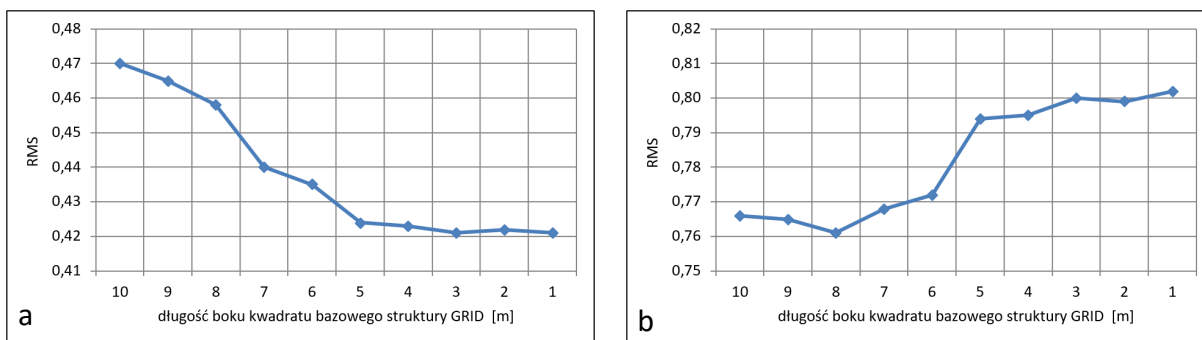
Rys. 16. Wyznaczenie siatki węzłów w strefach występowania (a) i braku (b) punktów pomiarowych

Automatyzacja tego procesu w systemie *SYPMEG* pozwala na wykonywanie analiz na obiektach pomiarowych w różnych ustawieniach konfiguracyjnych generowania struktury GRID. Daje to możliwość lokalizacji obszarów o potencjalnie małym lub dużym błędzie interpolacji. Wydzielenie obszarów pozwala ocenić potencjalne pokrycie obszaru siatką węzłów o zdefiniowanych parametrach oraz strefy, w których należy wykonać uzupełnienia. Wyznaczenie takich stref można wykorzystać do zaplanowania miejsc dodatkowego pomiaru oraz określić jego technologię (np. tachimetrię w przypadku powierzchni lub analizę zasobów mapy zasadniczej w przypadku infrastruktury technicznej), która pozwoli na dostarczenie dodatkowych danych niezbędnych do zachowania ciągłości tworzonego modelu powierzchni.

W drugim etapie przeanalizowano czynniki mające wpływ na ustalenie parametrów konstrukcyjnych struktury GRID. Podstawowym parametrem konstrukcyjnym siatki węzłów jest długość boku kwadratu bazowego, która odpowiada za rozdzielczość generowanej struktury. Ustalenie właściwej rozdzielczości siatki węzłów zależy od wielu czynników, których

wpływ można przeanalizować w systemie *SYPMEG*. Ustalenie odległości między węzłami tworzącymi model terenu w głównej mierze zależy od liczby punktów pomiarowych przypadających na jednostkę powierzchni. Większa liczba punktów położonych wokół wyznaczanego węzła pozwala algorytmom wyszukiwawczym na ustalenie odpowiedniej ich lokalizacji, co zapewnia wykluczenie ekstrapolacji obliczanej w nim wartości. Znajdzenie w otoczeniu węzła większej niż potrzebna do interpolacji liczby punktów pomiarowych umożliwiło wykorzystanie algorytmów pozwalających na aproksymację powierzchni. Z kolei im więcej punktów jest zlokalizowanych wokół węzła w mniejszym promieniu poszukiwań, tym powszechniej można wykorzystać uproszczone i szybkie algorytmy interpolacji. Zastosowanie analizy rozproszenia pionowego punktów pomiarowych pozwoliło na realizację założeń dotyczących wyboru sposobu interpolacji. Na obszarach o dużej dyspersji należy stosować algorytmy wykorzystujące aproksymację pozwalającą na kompensację rozproszenia punktów pomiarowych. Z kolei węzły wyznaczone w strefach o małej dyspersji mogą być interpolowane z zastosowaniem uproszczonych i szybkich algorytmów interpolacyjnych, co ma szczególne znaczenie w podnoszeniu wydajności przetwarzania megadanych.

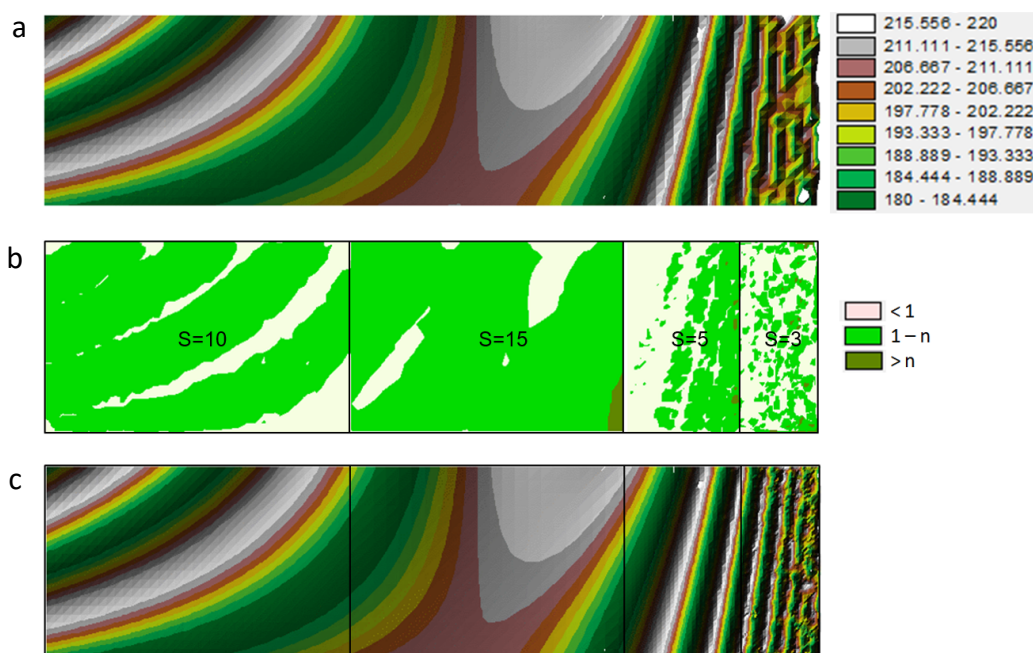
Analizy wykonywane dzięki możliwościom systemu *SYPMEG* pozwalają na zbadanie wpływu wielu czynników na rozdzielczość struktury GRID. Jednym z nich jest dobór wielkości siatki bazowej do różnego zagęszczenia punktów pomiarowych. Z przeprowadzonych badań wynika, że przy dużym zagęszczeniu punktów pomiarowych podnoszenie rozdzielczości GRID prowadzi do sukcesywnego zwiększania dokładności modelu. Ustalono również, że po przekroczeniu określonej wielkości siatki bazowej nie następuje już istotny wzrost dokładności (rys. 17 a).



Rys. 17. Zmiana dokładności interpolacyjnego modelu powierzchni podczas zwiększania rozdzielczości GRID przy zagęszczeniu 0,42 ppp/m² (a) i 0,042 ppp/m² (b)

Z kolei w przypadku małego zagęszczenia punktów pomiarowych sukcesywne zwiększanie rozdzielczości GRID powoduje pogorszenie dokładności modelu, które dla analizowanej metody interpolacji trwa do momentu wyrównania liczby węzłów i liczby punktów pomiarowych przypadających na jednostkę powierzchni (rys. 17 b). Analiza przebiegu zmian dokładności modeli w obu przypadkach pozwoliła na ustalenie granicznej rozdzielczości siatki węzłów, której zwiększanie nie powoduje już podnoszenia dokładności modelu interpolacyjnego.

W kolejnym etapie wykonano analizy pozwalające na ustalenie wielkości boku kwadratu bazowego struktury węzłów w zależności od morfologii analizowanej powierzchni. W zależności od zróżnicowania morfologicznego terenu odległości między poszczególnymi węzłami powinny być dobierane w taki sposób, aby rozdzielczość tworzonej struktury GRID pozwalała na możliwie dokładną charakterystykę rzeźby terenu z wykorzystaniem minimalnej liczby punktów węzłowych. Zaproponowana metoda pozwoliła na znaczną automatyzację procesu doboru wielkości siatki bazowej do ukształtowania mierzonej powierzchni z zachowaniem zrównoważenia dokładnościowego w strefach o różnej morfologii (rys. 18 a). Dzięki wyznaczeniu lokalnego wskaźnika zróżnicowania morfologicznego powierzchni i przypisania go do swobodnie regulowanej lokalizacji węzłów możliwa była analiza zróżnicowania powierzchni i wyznaczenie współczynnika morfologicznego o jednakowej wartości (rys. 18 b). Pozwoliło to na takie ustalenie parametrów struktury GRID, które zapewniło osiągnięcie podobnej dokładności modelu interpolacyjnego na różnie ukształtowanych obszarach obiektu pomiarowego i podniesienie jego jakości (rys. 18 c). Prezentowana metoda pozwala przeanalizować kształt morfologiczny mierzonej powierzchni, również na podstawie rzeczywistych zasobów pomiarowych.

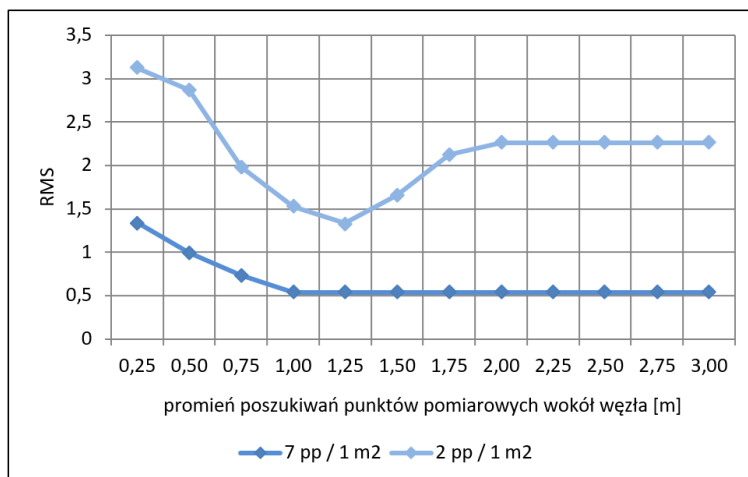


Rys. 18. Interpolacyjny model powierzchni wygenerowany przez klasyczną jednorodną strukturę GRID ($S=10$) w jednej strefie (a) i niejednorodną strukturę GRID w czterech strefach interpolacyjnych (c); b – podział na strefy z zachowaniem jednakowego współczynnika morfologicznego (n)

Wykorzystując system *SYPMEG* analizy tego rodzaju można przeprowadzić dla dowolnie zróżnicowanej powierzchni, w różnym zagęszczeniu punktów pomiarowych i swobodnie definiowanej rozdzielczości siatki wzorcowej. Z kolei możliwości iteracyjnej pracy systemu pozwalają na ograniczenie do minimum roli operatora systemu. Ostatecznie na analizowanym

obszarze jest ustalana niejednorodna struktura węzłów, która umożliwia wygenerowanie modelu powierzchni z założoną dokładnością w różnych strefach morfologicznych.

Trzeci etap obejmował analizę doboru punktów pomiarowych wykorzystywanych do interpolacji siatki węzłów. Każdy z algorytmów interpolacyjnych wymaga określonej liczby punktów pomiarowych oraz ich odpowiedniej lokalizacji wokół poszczególnych węzłów. Głównym elementem pozwalającym na wybranie pożądanych punktów ze spektrum pomiarowego jest określenie przestrzeni poszukiwań wyznaczonej wokół każdego potencjalnego węzła struktury GRID oraz zdefiniowanie liczby odszukiwanych w niej punktów. Z przeprowadzonych analiz wynika, że zasięg przestrzeni poszukiwań ma bezpośredni wpływ na jakość generowanych modeli powierzchni. Może ona zależeć od ustalonej rozdzielczości siatki węzłów, zagęszczenia punktów pomiarowych, dokładności wykonanych pomiarów, wykorzystywanych algorytmów interpolacyjnych czy szerokości stref służących do koordynacji styków niezależnych obszarów pomiarowych. Duże zagęszczenie punktów pomiarowych sprawia, że w odpowiednio dobranej rozdzielczości GRID ograniczanie przestrzeni poszukiwań punktów pomiarowych wokół węzłów nie wpływa znacząco na jakość modelu interpolacyjnego (rys. 19 – kolor niebieski: 7 pp / m²). Zmniejszenie przestrzeni poszukiwań punktów pomiarowych wokół węzła pozwala natomiast na znaczne podniesienie wydajności przetwarzania. Przeprowadzone analizy pozwoliły na ustalenie granicznej wielkości przestrzeni poszukiwań, której dalsze zwiększanie nie podnosiło już dokładności modeli, ponieważ na dokładność kompletnej struktury węzłów miały wpływ najbliższe punkty pomiarowe.

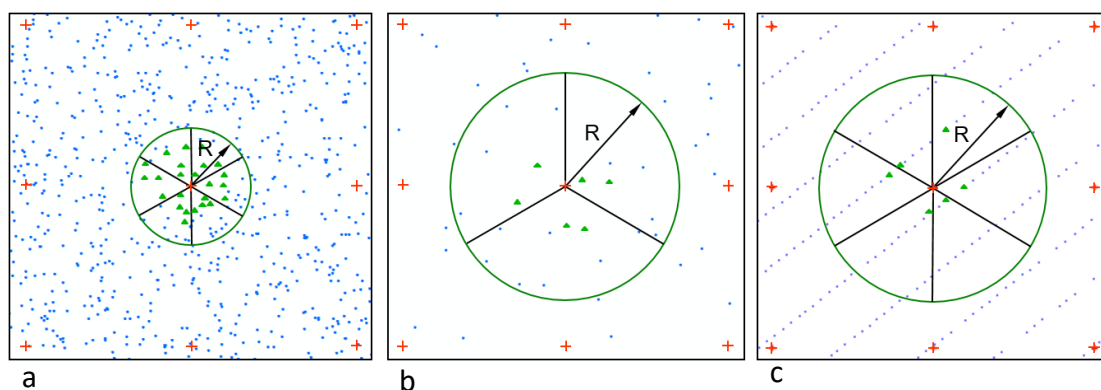


Rys. 19. Zmiana dokładności modeli powierzchni podczas zwiększania przestrzeni poszukiwań punktów pomiarowych wokół węzłów

Z kolei dla małego zagęszczenia punktów pomiarowych sukcesywne zwiększanie przestrzeni poszukiwań ustalonej wokół węzłów prowadziło do poprawy dokładności modelu jedynie w początkowym stadium (rys. 19 – kolor czerwony: 2 pp / m²). Dalsze zwiększanie przestrzeni poszukiwań punktów pomiarowych wokół węzłów powodowało stopniowe pogarszanie dokładności modeli. Również w tym przypadku przeprowadzone analizy pozwoliły na ustalenie

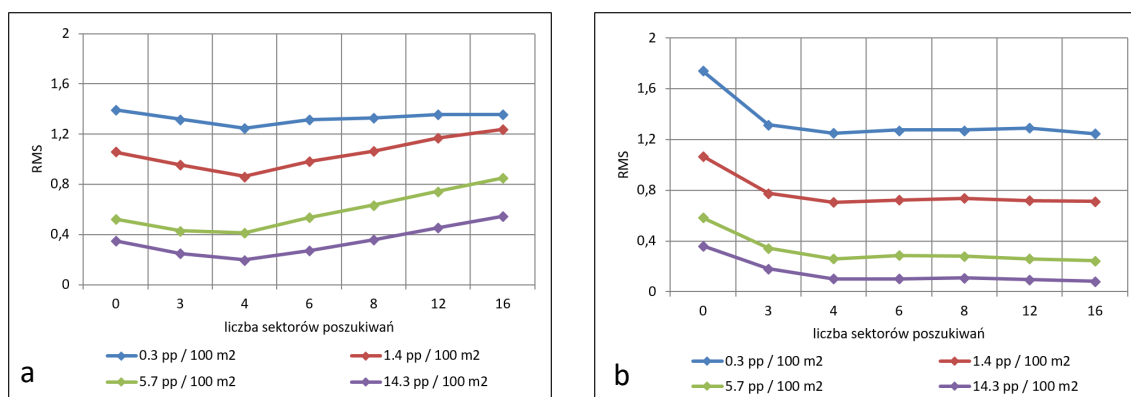
granicznej wielkości przestrzeni poszukiwań, która dla wybranej metody interpolacji pozwala na osiągnięcie najlepszej dokładności modelu.

Oprócz ustalenia wielkości przestrzeni poszukiwań, ważny jest również jej podział na sektory. Podział ten pozwala na odnalezienie odpowiedniej lokalizacji punktów pomiarowych wykluczających ekstrapolację wartości w węźle. Możliwość zdefiniowania liczby odszukiwanych punktów pomiarowych w każdym sektorze oraz ustalenie dopuszczalnej liczby sektorów bez odnalezionych punktów pozwala na precyzyjne określenie ich lokalizacji wokół każdego węzła (rys. 20). Zgodnie z założeniami prezentowanej metodyki w pracy przeanalizowano wydajne algorytmy wyszukiwacze, które gwarantują szybkie odszukanie prawidłowej lokalizacji punktów pomiarowych, mimo ich liniowego rozmieszczenia lub małego zagęszczenia. Odpowiednie zdefiniowanie parametrów przetwarzania w metodach wyszukiwawczych pozwoliło wykluczyć przypadki ekstrapolacji podczas stosowania wydajnych algorytmów interpolacyjnych i efektywnie wygenerować dokładne modele powierzchni. Skoncentrowanie się na wyszukiwaniu minimalnej liczby punktów w ograniczonej przestrzeni poszukiwań pozwoliło odnaleźć najwięcej przypadków interpolacji przy jednoczesnym maksymalnym skróceniu czasu przetwarzania. Gdy zagęszczenie punktów pomiarowych jest duże, istnieje duża swoboda w definiowaniu parametrów decydujących o ich rozmieszczeniu wokół węzła (rys. 20 a). Mała liczba punktów pomiarowych przypadających na jednostkę powierzchni nie zapewnia jednak w każdym przypadku kompletu danych niezbędnych do wykonania obliczeń dowolnym algorytmem interpolacyjnym. W takiej sytuacji przestrzeń wokół każdego węzła była dzielona na sektory przy jednoczesnym ustaleniu liczby punktów, które powinny być odnalezione w każdym z nich (rys. 20 b). W celu uniknięcia ekstrapolacji wyznaczany węzeł powinien znajdować się wewnątrz grupy punktów pomiarowych wykorzystywanych przez dany algorytm, co pozwala na wyznaczenie wartości w węźle na podstawie punktów rozmieszczonych równomiernie z każdej strony. Jest to szczególnie pożądane odnośnie do megadanych, w których przypadku lokalizacja punktów pomiarowych jest często współliniowa (rys. 20 c).



Rys. 20. Wybór punktów pomiarowych (pp) do interpolacji przy sektorowym podziale przestrzeni poszukiwań; a – duże zagęszczenie pp, małe zagęszczenie pp, c – liniowe rozmieszczenie pp

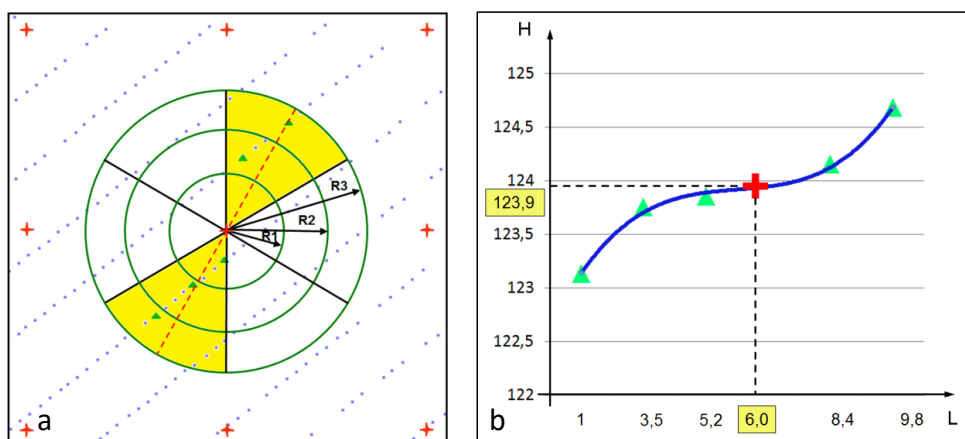
W pracy przeanalizowano dokładność modeli interpolacyjnych w zależności od liczby sektorów, na które podzielono przestrzeń poszukiwań punktów pomiarowych wokół każdego węzła z różnym zagęszczeniem punktów pomiarowych. Zbadano również przypadki zastosowania coraz większej liczby sektorów, gdzie początkowe zwiększanie ich liczby powodowało stopniową poprawę dokładności generowanych modeli. Dokładność modeli interpolacyjnych wzrastała do chwili podziału przestrzeni poszukiwań na cztery sektory (rys. 21). Następnie w zależności od użytego algorytmu interpolacyjnego (rys. 21 a - algorytm wielomianów lokalnych; rys. 21 b - algorytm radialnych funkcji bazowych) dokładność modeli pogarszała się lub pozostawała na niezmiennym poziomie. W przypadku każdego zagęszczenia modele interpolacyjne wygenerowane z udziałem punktów pomiarowych zlokalizowanych w czterech sektorach były najdokładniejsze. Wraz ze zwiększaniem liczby sektorów poszukiwań czas przetwarzania się wydłuża, dlatego podział przestrzeni poszukiwań na cztery sektory w przypadku badanych algorytmów interpolacyjnych należy uznać za najkorzystniejsze rozwiązanie. W przypadku wygenerowania porównywalnych rezultatów wybór ostatecznej metody interpolacji można dodatkowo wesprzeć analizą diagramów różnicowych. Pozwala to na porównanie jakości generowanej powierzchni i określenie charakterystyk zniekształceń wnoszonych przez analizowane algorytmy interpolacyjne. Omawiane analizy, dzięki zautomatyzowaniu procesu przetwarzania i możliwości pracy iteracyjnej systemu *SYPMEG*, mogą być przeprowadzone również w przypadku odmiennej morfologii powierzchni i różnych algorytmów interpolacyjnych, co pozwala na wybór najkorzystniejszego rozwiązania.



Rys. 21. Porównanie dokładności modeli interpolacyjnych utworzonych algorytmem wielomianów lokalnych (a) i radialnych funkcji bazowych (b) przy różnej liczbie sektorów poszukiwań

W czwartym etapie wykonano analizy dokładności i wydajności algorytmów interpolacyjnych oraz ich przydatności do modelowania powierzchni globalnie i lokalnie. Każdorazowo w procesie interpolacji generowano strukturę typu GRID, w której przypadku określone algorytmy interpolacyjne były wykorzystywane do zamiany nieregularnie rozmieszczonych punktów pomiarowych w regularnie rozmieszczone punkty siatki. Dobór odpowiedniego algorytmu interpolacyjnego oraz jego parametrów przetwarzania ma

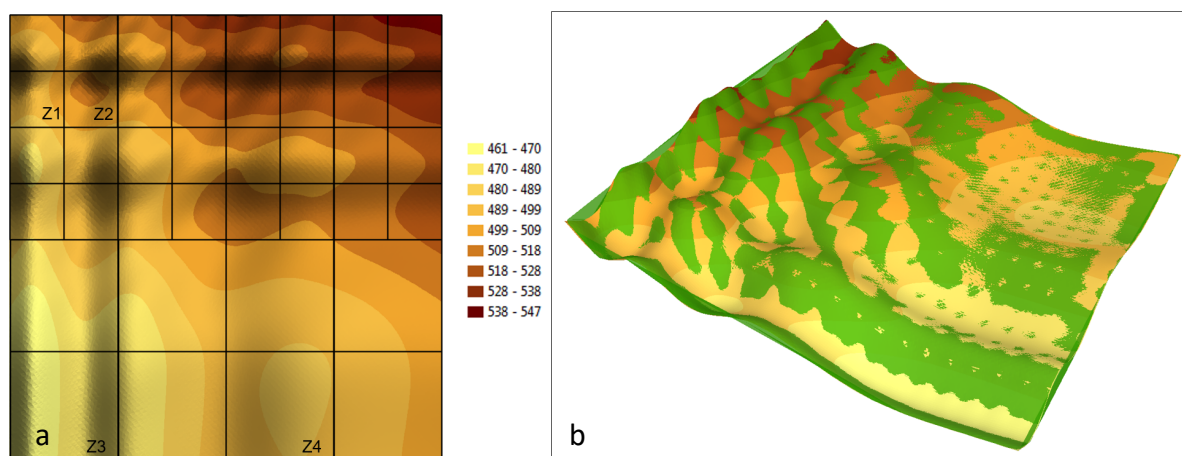
zasadniczy wpływ na dokładność analiz prowadzonych z zastosowaniem wygenerowanych powierzchni. W praktycznych zastosowaniach, do obliczenia wartości w szukanym punkcie interpolacyjnym (węźle siatki GRID) na podstawie otaczających go punktów pomiarowych, wykorzystywane są najczęściej klasyczne powierzchniowe metody interpolacyjne (Mitasova, Hofierka 1993, Rippa 1999, Wechsler 2003, Fornberg, Wright 2004, Ling, Kansa 2004, Larsson, Fornberg 2005, Tay i in. 2005, Fortuna i in. 2006, Raaflaub, Collins 2006, Erdogan 2009, Jankowski 2009). Oprócz wykorzystania powierzchniowych metod interpolacji, do wyznaczenia wartości w punkcie węzłowym można również zastosować metody aproksymacji i interpolacji przy użyciu funkcji jednej zmiennej. W pracy zaprezentowano również takie rozwiązanie. W takim przypadku przestrzeń wokół wyznaczanego węzła siatki dzielona jest na parzystą liczbę sektorów, a punkty pomiarowe w każdym sektorze wyszukiwane są w kolejnych strefach wyznaczonych przez różną wartość promienia poszukiwań R_1 , R_2 , R_3 (rys. 22 a). W każdej strefie odszukiwany jest jeden punkt położony najbliżej wyznaczanego węzła. Następnie punkty pomiarowe, zlokalizowane w sektorach leżących symetrycznie wokół szukanego węzła są rzutowane prostokątnie na płaszczyznę (LH) prostopadłą do płaszczyzny (XY) i przechodzącą przez dwusieczną kąta tworzonego przez naprzeciwległe sektory. Tak rzutowane punkty (stanowiące węzły interpolacji) wykorzystywane są do wyznaczenia wybranej funkcji interpolacyjnej (rys. 22 b). W wyniku rozwiązania układu równań (danej funkcji i prostej prostopadłej do osi L przechodzącej przez szukany węzeł), wyznaczana jest wartość H w szukanym punkcie. Działanie takie wykonywane jest na każdej parze sektorów leżących naprzeciw siebie po obu stronach węzła, a ostateczny wynik obliczonej wartości w punkcie szukanym jest uśredniany. Obliczenia mogą być również wagowane odległościami punktów pomiarowych w każdym sektorze od szukanego węzła.



Rys. 22. Wyznaczenie wartości w punkcie interpolowanym na podstawie punktów pomiarowych (a) przy wykorzystaniu funkcji jednej zmiennej (b)

W przypadku metod globalnych uwzględnienie wpływu wszystkich punktów daje powierzchnię, która jest pozbawiona lokalnych szczegółów morfologicznych i pozwala na ukazanie ogólnych tendencji ukształtowania. Globalna aproksymacja powierzchni pozwoliła wyznaczyć współczynniki funkcji (płaszczyzny lub wielomianu aproksymacyjnego o

automatycznie dobieranym stopniu), które umożliwiają odtworzenie kształtu mierzonych powierzchni, jak obiekty płaskie (boiska, lotniska, płaskie obiekty inżynierskie), czy podstawowych form morfologicznych (wzgórza, niecki, siodła). Zastosowanie aproksymacji globalnej dla całego obiektu pozwala na bardzo wydajne wygenerowanie kompletnej siatki węzłów o dowolnie ustalonej rozdzielczości i efektywne utworzenie na jej podstawie modeli powierzchni. Zaprezentowane rozwiązania pozwalają również na opisanie wielomianem powierzchni bardziej zróżnicowanej morfologicznie. W takim przypadku, w celu osiągnięcia pożądanej dokładności aproksymacji, analizowany obszar jest dzielony na podobszary dostosowane wielkością do kształtów podstawowych (rys. 23). W każdym z podobszarów jest wyznaczany indywidualny wielomian o automatycznie dobieranym stopniu, którego współczynniki są generowane na podstawie lokalnie odnajdywanych punktów pomiarowych. Pozwala to na utworzenie w wydzielonych podobszarach powierzchni aproksymacyjnych, które są najlepiej dopasowane do lokalnych kształtów obiektu pomiarowego (rys. 23 a: Z1 – siodło; Z2 – wzgórze; Z3 – rów; Z4 – niecka). W każdym z podobszarów jest możliwe wydajne wygenerowanie siatki węzłów o dowolnie definiowanej rozdzielczości. Połączenie struktury GRID w jedną całość (po uśrednieniu wartości brzegowych każdego podobszaru) pozwala na wygenerowanie kształtu całego obiektu (kolor brązowy), który nie odbiega od powierzchni wzorcowej (kolor zielony) (rys. 23 b).

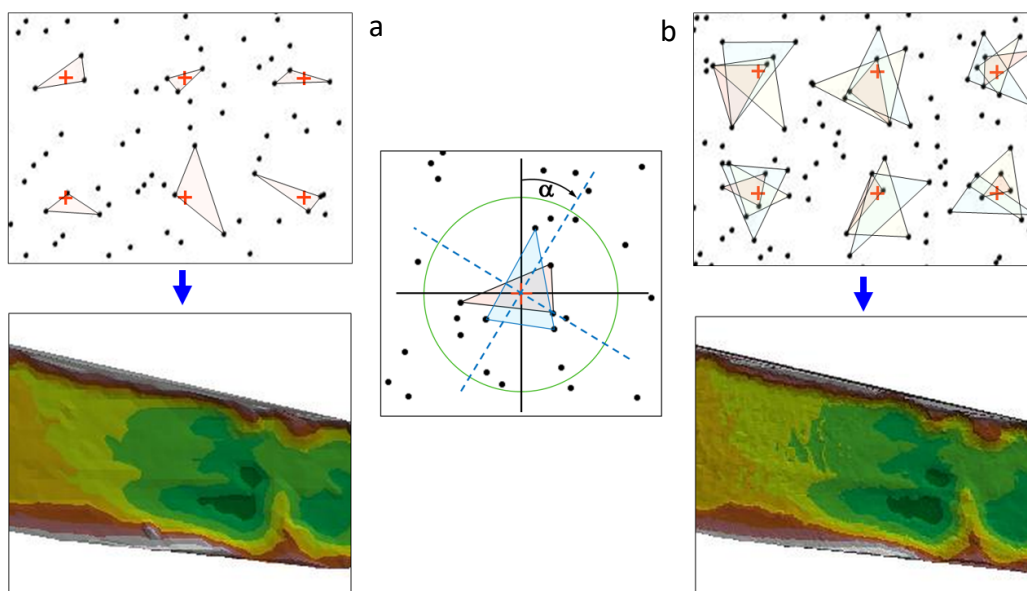


Rys. 23. Wybór podobszarów obiektu do aproksymacji globalnej z wykorzystaniem kształtów podstawowych (a) oraz powierzchnia aproksymacyjna utworzona przez połączenie podobszarów (b)

Zaletą proponowanych metod jest możliwość przechowywania w wynikowej bazie danych jedynie zakresu przestrzeni roboczej oraz współczynników wielomianów aproksymacyjnych danego fragmentu powierzchni. W przypadku podziału powierzchni na wiele fragmentów w wynikowej bazie danych jest przechowywany zestaw współczynników dla kolejnych podobszarów powierzchni sumarycznej. Po przesłaniu tych informacji do innego systemu możliwe jest wydajne wygenerowanie w nim kompletnej struktury GRID o dowolnej rozdzielczości. Struktura wygenerowana przez algorytmy aproksymacji globalnej może być również wykorzystana do odtworzenia ogólnego kształtu mierzonego obiektu i utworzenia

modeli poglądowych, które są wykorzystywane jako wstępny etap planowania właściwej interpolacji lokalnej.

W metodach interpolacji lokalnej jedna funkcja matematyczna jest wykorzystywana wielokrotnie do interpolacji lokalnych podzbiorów danych pomiarowych, które leżą w najbliższym sąsiedztwie kolejno wyznaczanych węzłów. Pozostałe punkty zasobu pomiarowego, położone poza zdefiniowaną przestrzenią poszukiwań, nie mają wpływu na obliczenia. Metody interpolacji lokalnej obejmują zasadniczą większość przypadków generowania struktury GRID na podstawie megadanych zarejestrowanych na obiektach pomiarowych. Pozwalają one na najdokładniejsze odzwierciedlenie zróżnicowanych form morfologicznych analizowanej powierzchni. W wykorzystywanych algorytmach interpolacji metody wyszukiwawcze są zintegrowane z algorytmami wyznaczającymi wartości w interpolowanych punktach. Zaproponowana w pracy metoda pozwala na oddzielenie procesów odpowiadających za lokalizację węzłów i wyszukiwanie punktów pomiarowych według zdefiniowanych kryteriów od procesów właściwej interpolacji, co zapewnia precyzyjne sterowanie wydajnością przetwarzania. Możliwość połączenia pliku z lokalizacją węzłów oraz zbioru z odszukanymi wokół nich punktami pomiarowymi w jeden zestaw danych pozwala na jednorazowe przeskanowanie całego spektrum pomiarowego. Po przekazaniu danych do algorytmu interpolacyjnego możliwe jest wykorzystanie całego potencjału systemu komputerowego wyłącznie do realizacji operacji obliczeniowych wymaganych przez określony algorytm. Jest to szczególnie korzystne w przypadku algorytmów wykorzystujących do obliczeń minimalną liczbę punktów pomiarowych (rys. 24 a). Ze względu na potrzebę przetworzenia dużej ilości danych w pracy zwrócono uwagę na wydajne algorytmy interpolacji lokalnej oraz zaproponowano metodę wielu trójkątów opisujących (MT - Multi Triangles).



Rys. 24. Modele powierzchni utworzone na podstawie punktów pomiarowych zlokalizowanych w trójkątach spełniających warunki interpolacji;
a – metoda klasyczna, b – metoda MT (Multi Triangles)

Po uwzględnieniu opcji wielu trójkątów, na podstawie przekazanych danych, następuje interpolacja jedną z wydajnych metod, które umożliwiają wyznaczenie wartości w węźle w oparciu o trzy punkty pomiarowe. W takiej sytuacji, wartości w każdym węźle wyznaczone są przez ten sam wybrany algorytm interpolacyjny. Obliczenie następuje oddzielnie dla wszystkich odnalezionych układów lokalizacyjnych (punktów odnalezionych w poszczególnych trójkątach), a ostateczna wartość, przypisywana do danej lokalizacji węzła jest uśredniana (rys. 24 b). Pozwala ona na uzyskanie pożądanej dokładności generowanej powierzchni interpolacyjnej z zachowaniem wysokiej wydajności przetwarzania, co ma szczególne znaczenie w wykorzystaniu megadanych przetwarzanych w czasie rzeczywistym.

Każda metoda interpolacji wykorzystuje inny algorytm do wyznaczania wartości w punktach węzłowych GRID i dodatkowo bazuje na różnie zlokalizowanych punktach pomiarowych (różne zagęszczenie, różna morfologia powierzchni), dlatego generowany model interpolacyjny może mieć różną dokładność. Z powodu odmienności morfologicznej różnych powierzchni oraz dużej liczby algorytmów i ich parametrów przetwarzania, liczba kombinacji możliwych do przeanalizowania jest praktycznie nieograniczona. Zastosowanie systemu *SYPMEG* ułatwia dobór najkorzystniejszego algorytmu interpolacyjnego (oraz zdefiniowanie jego parametrów) do przetworzenia różnie zlokalizowanych danych na odmiennych morfologicznie obiektach. Pozwala to, w zależności od potrzeb, na ustalenie kompromisu między jego dokładnością, a wydajnością (tab. 1).

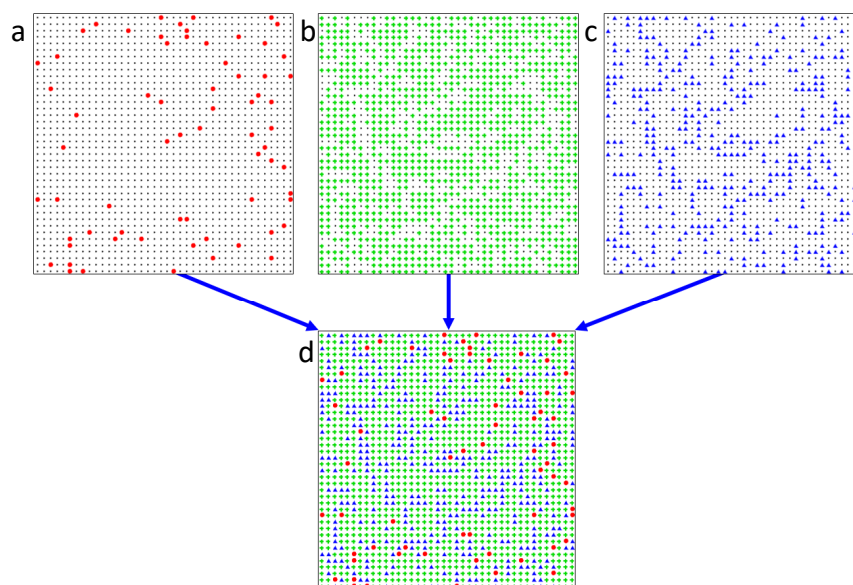
Tabela 1

Nr	Algorytm interpolacyjny	RMS	PI
1	Nearest Neighbor	1,265	1,000
2	Minimum Curvature	0,956	1,121
3	Natural Neighbor	0,832	5,644
4	Shepard	0,742	11,511
5	Kriging (Spherical)	0,935	5,556
6	Kriging (Cubic)	0,930	5,443
7	Kriging (Quadratic)	0,919	5,323
8	Kriging (Gaussian)	0,884	8,889
9	Kriging (Exponential)	0,864	11,114
10	Kriging (Rational Quadratic)	0,852	8,779
11	Kriging (Logarithmic)	0,794	15,212
12	Kriging (Linear)	0,752	13,331
13	RBF (Inverse Multiquadric)	0,853	10,219
14	RBF (Multilog)	0,809	14,002
15	RBF (Multiquadric)	0,727	15,112
16	RBF (Thin Plate Spline)	0,673	29,124
17	Local Polynomial (n=3)	0,783	4,378
18	Local Polynomial (auto max n=18)	0,729	8,458
19	Inverse Distance (n=2)	1,018	2,219
20	Inverse Distance (n=2) (+MT)	0,939	2,984
21	Weighting the Fields of Triangles	0,861	3,644
22	Weighting the Fields of Triangles (+MT)	0,768	4,211
23	Local Plane and Line	0,856	3,852
24	Local Plane and Line (+MT)	0,761	4,514

Zestawienie dokładności i wydajności algorytmów interpolacyjnych

Analiza i ocena wielu czynników wpływających na dokładność modeli interpolacyjnych w wirtualnej przestrzeni roboczej ułatwia wybór najlepszego rozwiązania dla modeli rzeczywistych. Z kolei automatyzacja procesu analiz umożliwia przebadanie wielu parametrów wybranych metod interpolacji oraz porównawcze określenie ich dokładności i wydajności. W opisywanej pracy wykonano analizę porównawczą 24 algorytmów interpolacyjnych, a jej wynik pozwolił na ocenę zarówno ich dokładności, jak i wydajności. Do oceny dokładności wykorzystano współczynnik RMS i diagramy różnicowe wygenerowane dla każdego z algorytmów. W celu porównania wydajności różnych algorytmów zastosowano przekształcenie ilorazowe (PI) względem punktu odniesienia (Kopczewska 2006, Sobczyk 2006). Za punkt odniesienia odpowiadający jednostkowej wartości tego przekształcenia ($PI = 1$) przyjęto czas przetwarzania uzyskany przez najmniej zaawansowany obliczeniowo algorytm najbliższego sąsiada. Zestawienie wielkości współczynnika dopasowania powierzchni oraz jakości modelu wynikającej z diagramu różnicowego i wydajności pracy danego algorytmu umożliwiło wybór najkorzystniejszego rozwiązania.

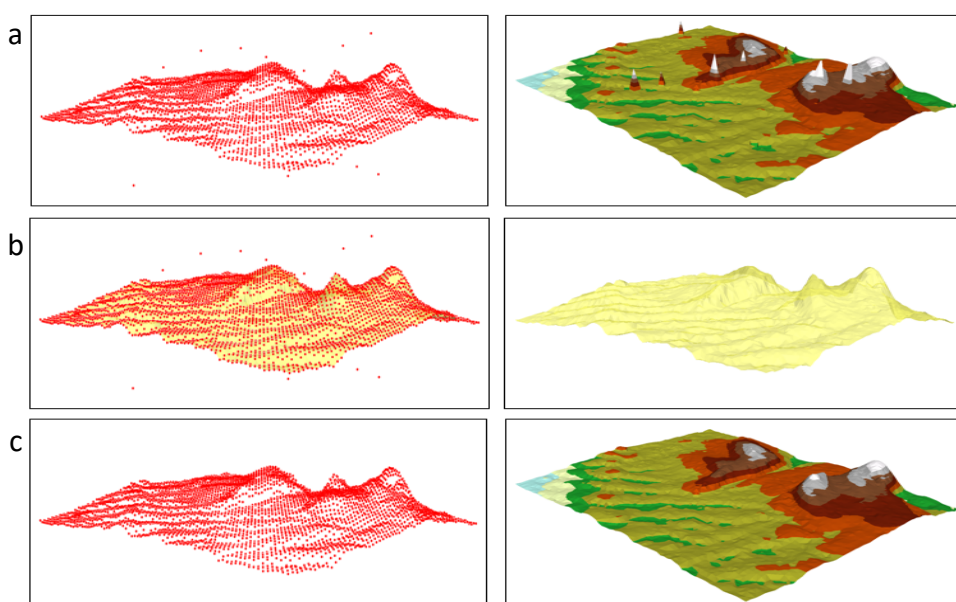
W pracy zaproponowano również metodę wykorzystującą kombinację algorytmów interpolacyjnych do wygenerowania struktury GRID z określoną dokładnością. Pozwala ona osiągnąć kompromis między szybkością wydajnych algorytmów interpolacyjnych a jakością zaawansowanych metod obliczeniowych. Ukształtowanie obiektu pomiarowego, zagęszczenie punktów pomiarowych czy rozdzielczość tworzonej siatki mogą być różne, dlatego trudno wskazać jeden uniwersalny algorytm interpolacyjny, który pozwalałby jednoznacznie najlepiej wygenerować na całym obszarze najdokładniejszą strukturę GRID. Dzięki modułowej konstrukcji systemu *SYPMEG* jest możliwe skonstruowanie kombinacji wybranych jednostek wykonawczych i złożenie z nich sekwencji przetwarzania pracującej automatycznie w określonych cyklach roboczych. Każdy z użytych w kombinacji algorytmów oblicza wartości na przyporządkowanych do niego węzłach, a wyniki łączone są w kompletną strukturę GRID (rys. 25).



Rys. 25. Połączenie różnych grup węzłów (a, b, c) wyznaczonych przez trzy odmienne algorytmy w kompletną strukturę GRID (d)

Prezentowane rozwiązanie pozwala na stopniowe wyznaczanie wartości przez wybrane do kombinacji algorytmy (np. według wydajności lub dokładności obliczeń). Jest to realizowane w kolejnych etapach przetwarzania wyłącznie na wybranych grupach węzłów, co pozwala wykorzystać potencjał obliczeniowy i wydajność poszczególnych metod, z jednoczesnym zachowaniem ustalonej dokładności modelu. W kolejnych cyklach procesu coraz dokładniejsze (i jednocześnie coraz mniej wydajne) algorytmy wykonują obliczenia na coraz mniejszych zbiorach danych, ponieważ interpolacja nie obejmuje już węzłów wyznaczonych poprzednimi metodami. Zastosowana metoda pozwala na stopniowe skracanie czasu niezbędnego do przeprowadzenia obliczeń w każdym cyklu, z zachowaniem wymaganej dokładności generowanego modelu.

Ostatni etap drugiej fazy procesu przetwarzania uwzględniał eliminację niepożądanych węzłów, których wartości mogą być błędnie wygenerowane w poprzednich etapach. Na tym etapie ze zbiorów były usuwane zdublowane informacje, co jest następstwem sekwencyjnego rejestrowania i przetwarzania pomiarów. Informacje się dublują także podczas wykorzystywania pomiarów zarejestrowanych w różnych epokach pomiarowych na tym samym obszarze lub podczas określania niezależnych sektorów przetwarzania i uzgadniania styków obszarów pomiarowych. Dzięki możliwości ustalenia przedziałów dokładności wybranych współrzędnych punktów położonych blisko siebie, zastosowane procedury pozwalają na wyeliminowanie redundancji powodującej niepożądane zwiększenie objętości zasobów. Dodatkowo umożliwiają uśrednienie wartości w zdublowanych węzłach, które obliczono z różnych zasobów pomiarowych dla tej samej lokalizacji przestrzennej. W zaprezentowanej metodyce istnieje także możliwość eliminacji nieprawidłowo wyznaczonych węzłów, które tworzą nieciągłości w generowanej powierzchni modelu interpolacyjnego (rys. 26).



Rys. 26. Eliminacja węzłów oddalonych od ogólnej powierzchni przy wykorzystaniu lokalnych funkcji aproksymacyjnych; a – węzły i DTM przed eliminacją, b – powierzchnia aproksymacyjna złożona z wielu funkcji lokalnych, c – węzły i DTM po eliminacji

Dzięki możliwościom systemu *SYPMEG* zadanie to jest realizowane w sposób zautomatyzowany dla zasobów przechowywanych w wielu zbiorach, z wykorzystaniem globalnych i lokalnych powierzchni aproksymacyjnych. Parametry pozwalające wybrać węzły do eliminacji mogą być definiowane globalnie (dla całego zasobu) lub indywidualnie (dla pojedynczych zbiorów danych), co pozwala na skontrolowanie zarówno opracowanych zbiorów końcowych, jak i wszystkich wyników pośrednich. Etap eliminacji niepożądanych węzłów kończy drugą fazę przetwarzania megadanych. Ostatecznie dane wynikowe operacji obliczeniowych drugiej fazy stanowią dane źródłowe dla trzeciej fazy procesu.

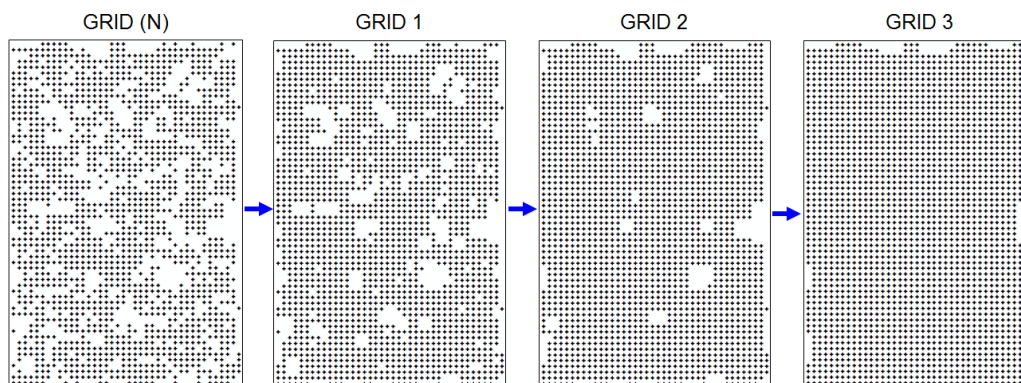
4.2.8.3. Przetwarzanie struktury GRID

W trzeciej fazie procesu przetwarzania megadanych były przetwarzane wygenerowane struktury GRID. Pozwoliło to na kontynuację przetwarzania megadanych przez administrowanie gotową strukturą siatki węzłów w celu polepszenia jej jakości, wykonania kompresji, konwersji, archiwizacji oraz wykorzystania jej podczas przetwarzania w kolejnych epokach czasowych.

W pierwszym etapie nastąpiło uporządkowanie wygenerowanego zasobu wynikowego, które pozwoliło na znacznie przyspieszenie i uproszczenie dostępu do informacji oraz umożliwiło wydajniejsze wykorzystanie jej do analiz. Podczas przetwarzania struktury GRID istotny jest dobór odpowiedniego algorytmu sortowania do charakterystyki sortowanej struktury danych (Knuth 1998, Grębosz 1999, Sysło 2002, Banachowski i in. 2006, Jankowscy 2009). W przypadku przetwarzania siatki węzłów wydajność sortowania można dodatkowo podnieść, wykorzystując charakterystyczne cechy tej struktury danych i odpowiednio modyfikując algorytmy sortujące. Z zastosowaniem indeksacji kolumn i wierszy proces sortowania przebiega szybciej, ponieważ raz zaindeksowane rekordy nie są ponownie brane pod uwagę i całe grupy danych (kolumn lub wierszy) mogą być wykluczane z wielokrotnego przetwarzania. Dodatkowo współrzędne siatki węzłów mogą być reprezentowane za pomocą liczb całkowitych, co pozwala zastosować wydajne algorytmy sortowania pozycyjnego. Podczas doboru algorytmu sortującego można również wykorzystać częściowe posortowanie danych na wcześniejszych etapach przetwarzania. W prezentowanej metodyce omówiono możliwości wykorzystania określonych rodzajów algorytmów sortujących do przetworzenia struktur danych o różnej charakterystyce.

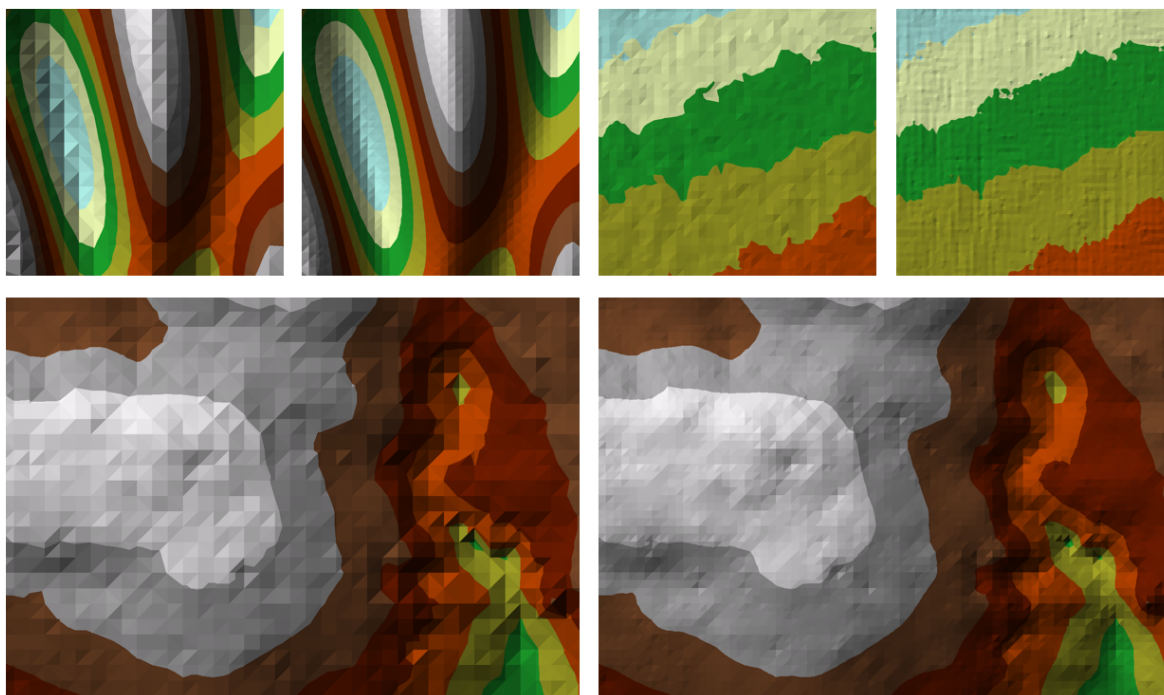
W kolejnym etapie przeanalizowano możliwości uzupełniania niekompletnej struktury GRID przez wyznaczenie brakujących węzłów wewnętrznych. Kompletna struktura GRID przyczynia się do podniesienia dokładności modelowania powierzchni oraz do polepszenia jakości analiz przestrzennych. W pracy zaproponowano dwie metody uzupełniania brakujących węzłów, tj. iteracyjną i jednoprzebiegową oraz przeanalizowano ich dokładność. Dzięki zastosowaniu w systemie *SYPMEG* przetwarzania wsadowego oraz możliwości pracy iteracyjnej, obie metody pozwalają na przetworzenie wielu zbiorów w sposób zautomatyzowany, co ogranicza udział operatora systemu do ustalenia początkowych wartości parametrów przetwarzania. Zastosowanie metody iteracyjnej umożliwia sukcesywne uzupełnianie struktury GRID w kolejnych cyklach przetwarzania, oddzielnie dla węzłów zlokalizowanych wzdłuż osi X i Y. Pozwala to na zdefiniowanie indywidualnych parametrów

przetwarzania w każdym cyklu oraz wygenerowanie ściśle określonych grup węzłów uzupełniających strukturę w każdym z etapów przetwarzania (rys. 27).



Rys. 27. Etapy uzupełniania struktury GRID w kolejnych iteracjach

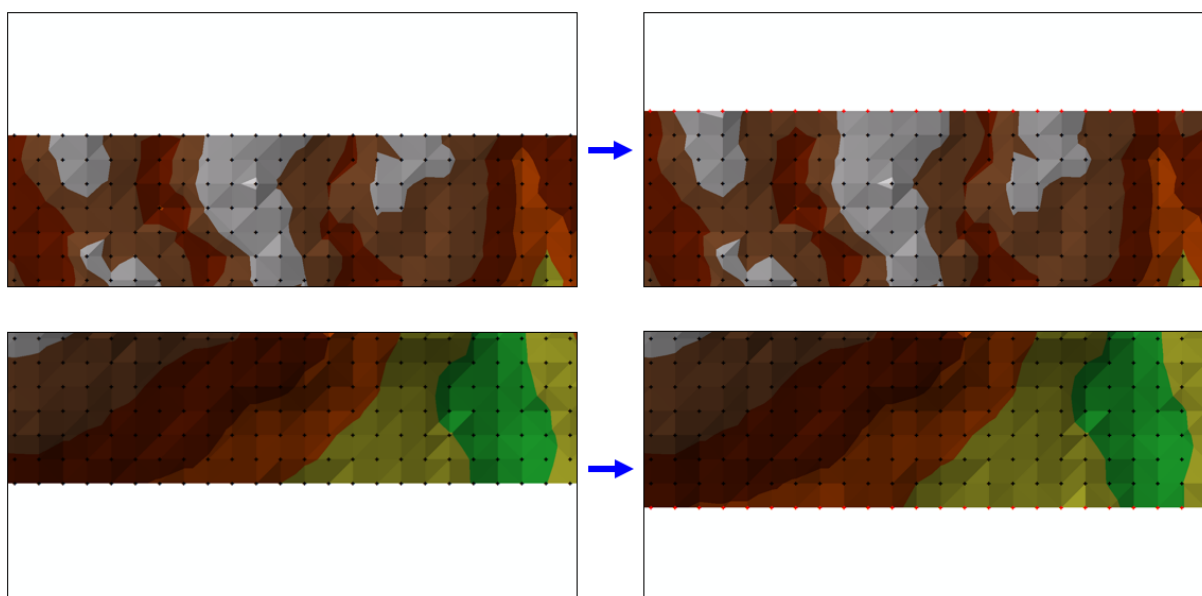
Zastosowanie procesu iteracyjnego pozwala na osiągnięcie lepszych rezultatów niż wygenerowanie brakujących węzłów w jednym przebiegu. Kontrolowanie dopuszczalnej uzupełnianej przerwy w strukturze w czasie wykonywania kolejnych iteracji pozwala na podniesienie dokładności wygenerowanego modelu końcowego w stosunku do modelu utworzonego jednorazowo. Z kolei metoda jednorazowa pozwala na bardzo wydajne wygenerowanie brakujących węzłów dzięki wykorzystaniu aproksymacji analizowanej powierzchni za pomocą jednej funkcji wyznaczonej globalnie lub szeregu funkcji lokalnych. W przypadku wykorzystania funkcji lokalnych kształt przetwarzanej powierzchni może być dowolny. Zaprezentowana metodyka pozwala również na zwiększenie rozdzielczości już istniejącej siatki węzłów (rys. 28).



Rys. 28. Efekt zagęszczania struktury GRID

Proces taki może posłużyć do wygładzania generowanych powierzchni tworzonych na bazie struktury GRID i przyczynić się do wyeliminowania jej uskoków oraz zakłóceń w przebiegu warstw. Ponadto brakujące węzły można wyznaczyć lub zwiększyć ich rozdzielczość (zagęszczenie) w sytuacji braku dostępu do zasobów pomiarowych. W przypadku omawianych metod uzupełnianie niewielkich ubytków w strukturze GRID jest bardziej opłacalne ekonomicznie od przeprowadzenia nowej interpolacji dla kompletnych zasobów pomiarowych. Wygenerowane dodatkowe węzły uzupełniają występujące braki w pierwotnej strukturze i pozwalają na podniesienie jakości generowanego modelu. Uzyskanie kompletnej siatki węzłów o zdefiniowanej rozdzielczości umożliwia także prowadzenie analiz zmian modelu powierzchni w kolejnych epokach czasowych w dowolnym punkcie analizowanej przestrzeni.

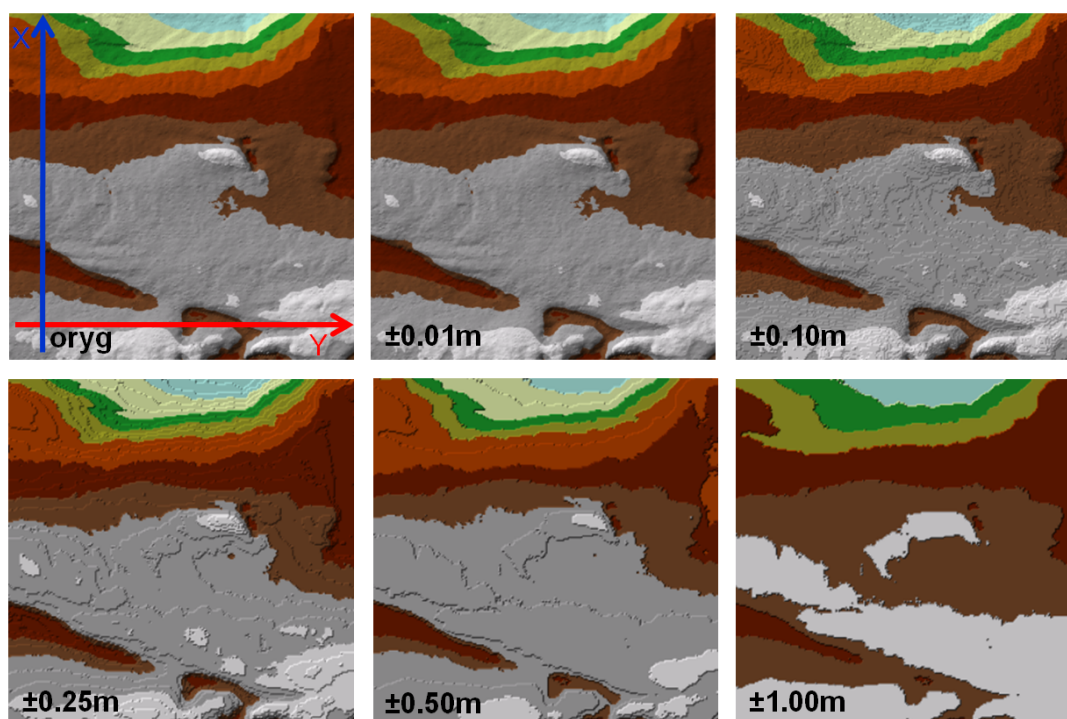
W trzecim etapie przeanalizowano dokładność wyznaczenia węzłów brzegowych generowanych na krawędziach modeli powierzchni, na których brakuje odpowiednio rozmieszczonych punktów pomiarowych lub wcześniej wyznaczonych węzłów. Jednostronna lokalizacja punktów wykorzystywanych do wyznaczenia wartości w węzłach brzegowych z reguły przyczynia się do ekstrapolacji wartości, co sprawia, że szukane wielkości są wyznaczone z mniejszą dokładnością lub są eliminowane ze zbioru wynikowego. Wyeliminowanie punktów położonych na granicach mierzonych obszarów powoduje częściową utratę informacji oraz utrudnia dokładne spasowanie opracowywanych obszarów. Zaprezentowane w pracy rozwiązania pozwalają na wyznaczenie węzłów brzegowych, a przeprowadzone analizy umożliwiają wybór najlepszego wariantu. Zaproponowany proces obliczenia wartości w każdym węźle brzegowym zachodzi niezależnie dla każdego z końców współliniowej struktury węzłów, co pozwala na wygenerowanie wszystkich węzłów brzegowych spełniających przyjęte założenia dokładnościowe. W efekcie następuje poszerzenie styków obszarów roboczych, co umożliwia ich dokładniejsze spasowanie (rys. 29).



Rys. 29. Efekt wyznaczenia węzłów brzegowych na różnych modelach interpolacyjnych

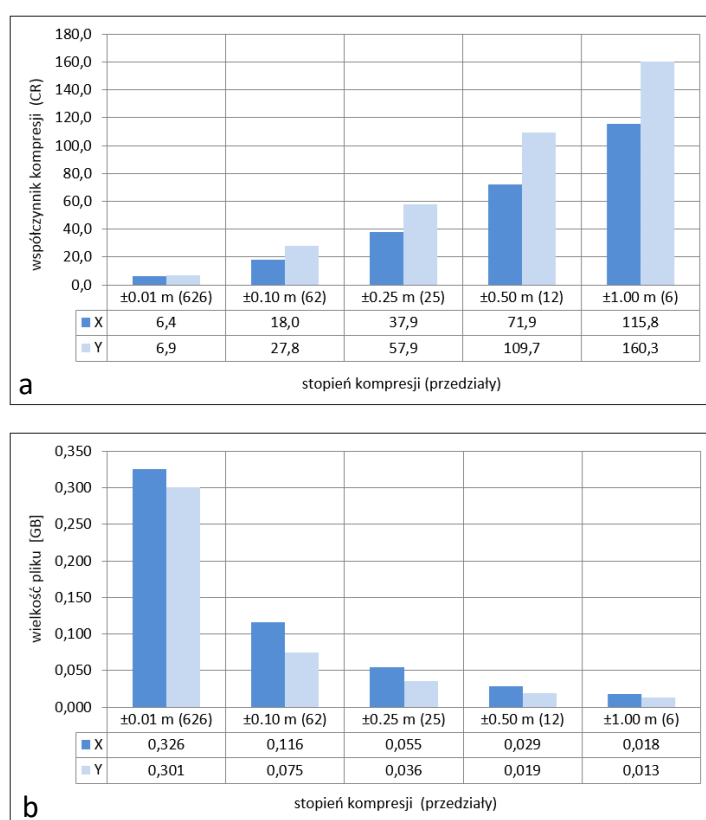
Wykorzystanie liniowej lokalizacji węzłów (w wierszach i w kolumnach) pozwoliło na wyznaczenie pierwszego i ostatniego węzła we wszystkich współliniowych strukturach punktów siatki GRID. W systemie *SYPMEG* możliwe jest także wykorzystanie dowolnych algorytmów interpolacyjnych (zastosowanych globalnie i lokalnie) do wyznaczenia struktury węzłów brzegowych oraz przeanalizowanie jakości generowanych przez nie wyników. Dzięki przeprowadzonym analizom możliwe było również ustalenie, w jakim stopniu dokładność wewnętrznej struktury GRID wpływa na jakość generowanych węzłów brzegowych. System *SYPMEG* ułatwia także dobór algorytmów interpolacyjnych do określonej morfologii powierzchni oraz ustalenie, które z nich pozwalają na osiągnięcie najlepszych wyników, mimo ekstrapolacji wartości w węzłach brzegowych.

Czwarty etap obejmował analizę kompresji zbiorów GRID z wykorzystaniem regulowanej straty dokładności (Rissanen, Landgon 1990, Howard, I Vitter 1994, Sayood 2002, Przelaskowski 2005). Zastosowanie kompresji pozwala na ograniczenie objętości zbiorów, co jest szczególnie istotne w przypadku megadanych, ponieważ przyspiesza proces ich przekazywania między systemami komputerowymi oraz ogranicza koszty ich magazynowania. W procesie kompresji następuje redukcja wielkości zasobów, polegająca na zmianie sposobu organizacji zapisu danych w nich przechowywanych. W zaproponowanej metodyce jest ona wykonywana w taki sposób, aby w procesie dekompresji zapewnić odtworzenie struktury GRID z pożądaną dokładnością (rys. 30). Wykorzystanie iteracyjnej pracy systemu *SYPMEG* pozwala na regulację stopnia kompresji i dostosowanie go do dokładności pozyskanych i archiwizowanych informacji.



Rys. 30. Powierzchnie DTM wygenerowane dla różnych stopni kompresji

Dodatkowo moduły kompresji i dekompresji zaprojektowano w taki sposób, aby można było przetworzyć dowolną (kompletną lub niekompletną) strukturę GRID. Proces kompresji jest realizowany dzięki pogrupowaniu struktury węzłów w określone przedziały, których rozpiętość pozwala na zdefiniowanie stopnia kompresji zasobów. Wielkość przedziałów decydująca o stopniu kompresji może być dobrana do dokładności pozyskania punktów pomiarowych oraz dokładności wygenerowania punktów węzłowych w procesie interpolacji. Pozwala to na ustalenie takiego stopnia kompresji, który nie wpływa znacząco (lub zupełnie nie ma wpływu) na odtworzenie danych w procesie dekompresji, umożliwiając jednocześnie znaczną redukcję archiwizowanych zasobów. Zastosowana kompresja z regulowaną dokładnością modelu pozwala na wygenerowanie zbiorów wynikowych, dla których wielkość stopnia kompresji decydującej o objętości pliku wynikowego można dowolnie regulować (rys. 31).

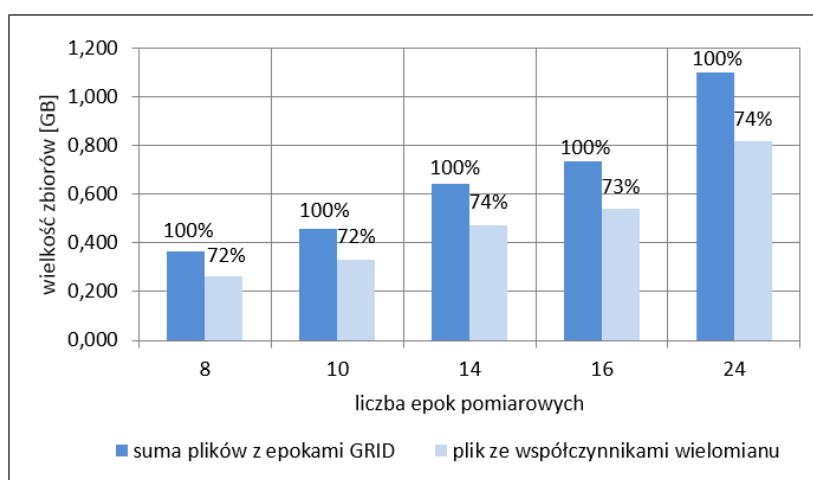


Rys. 31. Współczynnik kompresji (a) i wielkości zbiorów danych (b) dla różnych stopni kompresji

Oprócz wygenerowania modeli docelowych, pozwala to również na wydajne utworzenie modeli poglądowych, które można wykorzystać do przedstawienia ogólnych tendencji morfologicznych analizowanego obiektu. Dodatkowa możliwość wyboru sposobu sortowania kompresowanej struktury GRID według współrzędnej x lub y pozwala na dostosowanie porządku kompresji do ukształtowania morfologicznego powierzchni, co wpływa na zwiększenie współczynnika kompresji (rys. 31: X – kolor niebieski; Y – kolor czerwony). Cykliczna praca systemu *SYPMEG* pozwala na zbadanie różnych stopni kompresji dla

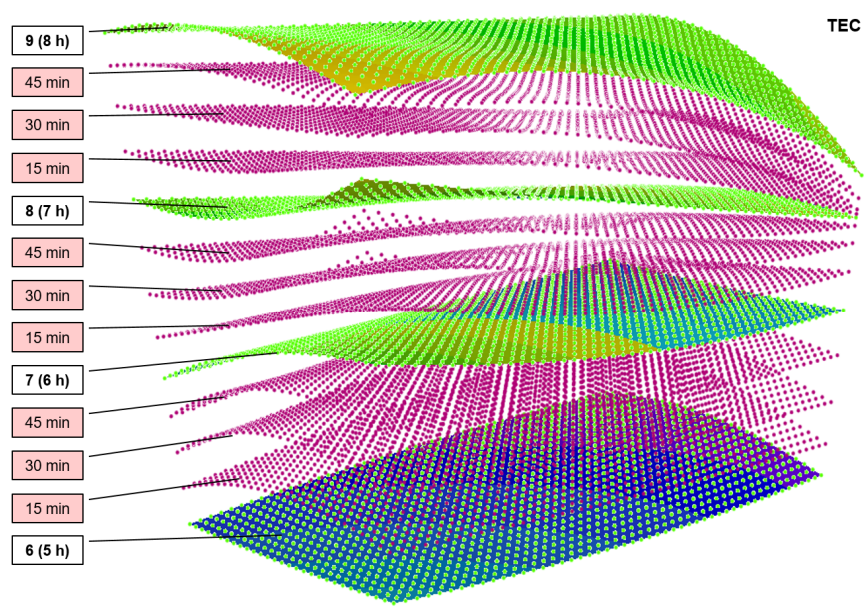
analizowanego obiektu oraz umożliwia wybór takiego skompresowanego modelu, który spełnia określone wymagania dokładnościowe. Przeprowadzone analizy pozwoliły określić progową wartość stopnia bezstratnej kompresji, której zastosowanie pozwala na osiągnięcie wymaganej jakości powierzchni generowanego modelu.

W piątym etapie zaproponowano wykorzystanie wygenerowanej struktury GRID do utworzenia zbiorów wynikowych pozwalających na przeprowadzenie analiz zmian modeli w czasie. W przypadku megadanych gromadzenie informacji w wielu epokach pomiarowych wymaga wykorzystania wyjątkowo dużych zasobów systemowych, a do ich przetworzenia jest niezbędne zastosowanie wydajnych algorytmów obliczeniowych. Dzięki zaproponowanej metodyce proces archiwizacji dużej ilości danych pomierzonych w różnych epokach czasowych oraz wykonywanie analiz z ich wykorzystaniem może być w znacznym stopniu zautomatyzowany i uproszczony. Wykorzystanie w tym celu struktury GRID jest szczególnie korzystne, ponieważ pozwala na ograniczenie ilości danych biorących udział w przetwarzaniu oraz uporządkowanie przechowywanych zasobów. Ustalona w każdej epoce pomiarowej jednakowa lokalizacja węzłów struktury GRID pozwala na porównanie wartości zmieniających się w czasie w tym samym miejscu (w wybranym punkcie węzłowym lub grupie punktów) zidentyfikowanym na mierzonym obiekcie. Dane z każdej epoki pomiarowej są wykorzystywane do interpolacji struktury GRID, która za każdym razem jest uzupełniana do kompletnej siatki węzłów. Pozwala to na uzyskanie kompletu danych w każdym węźle w każdej epoce. Umożliwia to z kolei ustalenie określonych interwałów czasowych w każdym miejscu analizowanej przestrzeni. Dzięki temu możliwe jest wyselekcjonowanie zmieniających się wartości w czasie w każdym węźle. Zastosowanie funkcji pozwalającej na opisanie zmian w czasie w każdym węźle wieloma współczynnikami umożliwia przechowanie wszystkich informacji o zmieniającym się modelu w jednym pliku. Uporządkowanie zapisu z wielu epok pomiarowych i umieszczenie ich w jednym zbiorze pozwala na zmniejszenie objętości przechowywanych danych (rys. 32), co umożliwia wydajniejszą archiwizację, zwiększenie szybkości wymiany informacji oraz uproszczenie administrowania zasobami.



Rys. 32. Porównanie wielkości zbiorów zawierających dane z epokami GRID z plikiem przechowującym współczynniki wielomianu interpolacyjnego

Przechowywanie informacji o zmianie wartości w poszczególnych węzłach w postaci współczynników funkcji umożliwia również odtworzenie pożądanych wielkości w dowolnym momencie czasowym oraz bezstratne wygenerowanie na ich podstawie określonej powierzchni modelu, przypisanej do dowolnie wybranej epoki. Dysponując danymi pozwalającymi utworzyć szereg czasowy dla każdego węzła struktury GRID, można przeanalizować zmiany zachodzące w mierzonej przestrzeni w wybranych punktach obiektu i je porównać. Dzięki możliwości definiowania interwału czasowego jest również możliwe wygenerowanie struktury GRID dla epok pośrednich na podstawie węzłów utworzonych w sąsiadujących epokach pomiarowych. Zbiór węzłów wygenerowanych w ten sposób dla określonej epoki pośredniej tworzy pośrednią odtworzeniową strukturę GRID, umożliwiającą wygenerowanie pośredniej powierzchni na określony moment czasowy. Przykład pokazany na rysunku 33 demonstruje odtworzenie pośrednich epok pomiarowych na przykładzie pomiaru całkowitej zawartości elektronów (ang. TEC - Total Electron Content), która jest miarą wpływu jonosfery na sygnał nadawany przez satelitę.



Rys. 33. Tworzenie pośrednich struktur GRID na podstawie węzłów wygenerowanych w 4 epokach pomiarowych

Zaprezentowana metodyka pozwala na odtwarzanie danych w wybranych momentach czasowych bez konieczności przechowywania całego zasobu pomiarowego oraz umożliwia ograniczenie ilości informacji niezbędnych do odtworzenia kompletnej struktury GRID w dowolnym momencie. Przyczynia się to do łatwiejszej archiwizacji danych pomierzonych w wielu epokach oraz do wydajniejszego przesyłania informacji między aplikacjami generującymi modele powierzchni.

W ostatnim etapie przetwarzania megadanych uwzględniono ustalenie struktury i formatu zapisu danych przed przekazaniem ich do zewnętrznego systemu obsługi lub do bazy SIS. W zależności od wymagań stawianych przez końcowego użytkownika danych struktura informacji przechowywanej w pliku oraz jego ostateczny format może być dowolnie określona.

Przekazywane zbiory powinny być również uporządkowane zgodnie z założeniami ich wykorzystania w systemie docelowym. Ze względu na dużą objętość zasobów należy ustalić ilość informacji przechowywanej w pojedynczym zbiorze, która jednorazowo może być przetworzona przez system docelowy, tj. określić maksymalną akceptowaną przez system zewnętrzną liczbę rekordów przechowywanych w pojedynczym pliku. Na tym etapie jest także wykonywana ostateczna kontrola i eliminacja błędnie zapisanych lub niepożądanych rekordów struktury GRID. Wykorzystanie cyklicznej pracy systemu *SYPMEG* pozwala na automatyczne przetworzenie dowolnej ilości zbiorów i podzielenie tworzonych zasobów na określone porcje informacji. Uniwersalny sposób zapisu danych stosowany w systemie (jednopoziomowa baza trójpolowych rekordów XYH) pozwala na ich wyeksportowanie do dowolnej aplikacji zewnętrznej. Zasób wynikowy może być również przekonwertowany zgodnie z wymogami wykorzystywanej aplikacji. Oprócz współrzędnych lokalizujących każdy węzeł struktury GRID w trójwymiarowej przestrzeni, do każdego rekordu zasobu wynikowego mogą być dołączone dane, np.: wymagany w bazie danych numer rekordu, wspólny dla określonej grupy danych numer obiektu pomiarowego, informacje o numerze sesji pomiarowej lub wymagany w celu klasyfikacji danych identyfikator urządzenia pomiarowego. Pozwala to pogrupować dane, sklasyfikować je oraz zautomatyzować ich przypisanie do wybranych analiz w systemie zewnętrznym.

4.2.9. Podsumowanie etapów badań i uzyskane wyniki

W prezentowanej pracy, posiłkując się licznymi modelami teoretycznymi i praktycznymi powierzchniami, przeprowadzono szereg analiz przy wykorzystaniu autorskich algorytmów przetwarzania danych. Realizację proponowanych w metodyce działań podzielono na trzy główne etapy: przetwarzanie punktów pomiarowych, generowanie struktury GRID i przetwarzanie struktury GRID. Każdy z etapów podzielony został na szereg zadań funkcjonalnych, które poddane zostały kontroli dokładności. W pierwszym etapie zaproponowano sposób wstępnego przygotowania danych pomiarowych oraz przeanalizowano parametry lokalizacyjne pomierzonych punktów. Dokonano analizy rozmieszczenia punktów pomiarowych pod kątem ich wykorzystania w algorytmach interpolacji oraz zaproponowano sposoby filtracji niepożądanych punktów. W drugim etapie przebadano różne konfiguracje kluczowych parametrów lokalizacyjnych struktury GRID oraz przeanalizowano wpływ parametrów konstrukcyjnych siatki węzłów na dokładność i jakość tworzonych modeli powierzchni. Zaproponowano metodę automatycznego doboru rozdzielczości generowanej siatki węzłów do morfologii badanej powierzchni. Poddano ocenie rozmieszczenie i metody doboru punktów pomiarowych do określonej metody interpolacji. Zbadano możliwości wykorzystania globalnej aproksymacji powierzchni do generowania podstawowych kształtów morfologicznych. Zaproponowano wykorzystanie aproksymacji powierzchni do wydajnego przekazywania danych pomiędzy aplikacjami SIS oraz możliwość odtwarzania powierzchni o dowolnie ustalonej rozdzielczości GRID. Zaproponowano metodę interpolacji wykorzystującą wydajne algorytmy obliczeniowe oraz przeanalizowano dokładność i wydajność algorytmów wykorzystywanych do lokalnej interpolacji wartości w węźle. Zaproponowano sposoby zwiększenia wydajności przy wykorzystaniu kombinacji

algorytmów interpolacyjnych oraz metody eliminacji niepożądanych węzłów. W trzecim etapie zaproponowano metody przetwarzania i poprawiania jakości siatki węzłów po jej wygenerowaniu. Poddano analizie metody sortowania węzłów oraz przedstawiono metody uzupełniania wewnętrznej struktury GRID. Zaproponowano metody generowania węzłów brzegowych oraz przedstawiono sposoby archiwizacji struktury GRID i jej kompresji z wykorzystaniem regulowanego stopnia dokładności. Przeanalizowano możliwości wykorzystania struktury GRID w celu przygotowania informacji do analiz w czasie (dla danych przestrzennych zarejestrowanych w wielu epokach pomiarowych) oraz zaproponowano metodę archiwizacji i odtwarzania powierzchni obiektu w dowolnej epoce pomiarowej (w dowolnym momencie czasu). W ostatnim etapie ustalono sposoby wymiany informacji pomiędzy proponowanym systemem, a aplikacjami SIS.

Zaproponowana metodyka postępowania pozwoliła na uporządkowanie kolejności etapów przetwarzania danych podczas całego procesu generowania DTM oraz umożliwiła ustalenie kluczowych parametrów przetwarzania na każdym z nich. W celu dokonania kompleksowej analizy poszczególnych etapów, cały proces podzielono na szereg elementów składowych, z których każdy odpowiada za kolejne fazy przetwarzania danych. Mogą one tworzyć sekwencje przetwarzania lub być wykorzystywane jako niezależne moduły. Założenia funkcjonalne metodyki i systemu przewidują także wykorzystanie automatycznego przetwarzania wsadowego wielu zbiorów danych jak również cyklicznych procesów iteracyjnych. Dzięki takiemu rozwiązaniu, oraz ustawicznej kontroli dokładności na poszczególnych etapach przetwarzania, możliwe jest określenie parametrów pracy poszczególnych elementów systemu, pozwalających na uzyskanie numerycznego modelu terenu charakteryzującego się pożądaną dokładnością i jakością. Kompleksowa analiza dokładności wyników kolejnych iteracji, umożliwia odpowiedni dobór parametrów przetwarzania na poszczególnych etapach oraz pozwala na ocenę ich wpływu na jakość generowanych powierzchni. Badania prezentowane w pracy obejmują także analizy przy wykorzystaniu danych źródłowych o swobodnie definiowanej charakterystyce rozmieszczenia przestrzennego. Jednym z założeń prezentowanej metodyki i opracowanego systemu jest możliwość tworzenia wirtualnych przestrzeni roboczych i dokonywanie w oparciu o nie analiz numerycznych pod kątem weryfikacji dokładności poszczególnych faz procesu. Takie rozwiązanie umożliwia dobór odpowiednich algorytmów przetwarzania danych oraz ustalenie optymalnych parametrów ich pracy dla swobodnie definiowanej lokalizacji punktów pomiarowych, przy różnorodnym ukształtowaniu morfologicznym badanej powierzchni. Analiza i dobór parametrów przetwarzania zmierza ostatecznie do wygenerowania kompletnej struktury GRID pozwalającej na uzyskanie numerycznego modelu terenu, charakteryzującego się minimalną liczbą punktów tworzących i maksymalną dokładnością. Możliwość pracy na wirtualnych przestrzeniach roboczych pozwala na określenie pożądaných parametrów przetwarzania danego algorytmu w dowolnych układach przestrzennych (różna lokalizacja punktów pomiarowych, różna morfologia powierzchni, różne parametry konstrukcyjne GRID), co w efekcie umożliwia zdefiniowanie warunków pracy podczas przetwarzania rzeczywistych danych pomiarowych. Ze względu na to, że ilość danych

możliwych do uzyskania współczesnymi metodami pomiarowymi jest bardzo duża i wciąż rośnie lawinowo, w badaniach skoncentrowano również uwagę na maksymalnym zautomatyzowaniu ich przetwarzania przy jednoczesnym ograniczeniu czynności podejmowanych przez operatora systemu. Przeprowadzone analizy pozwoliły na wypracowanie kryteriów umożliwiających przyspieszenie i zautomatyzowanie procesu przetwarzania danych pozyskiwanych masowo oraz ustalenie optymalnych elementów konstrukcyjnych struktury GRID. Umożliwiły one dobór parametrów przetwarzania w różnych fazach procesu, co pozwoliło na uzyskanie pożądanej dokładności i jakości generowanego DTM, przy jednoczesnej optymalizacji wielkości zbiorów tworzących model. Wypracowane kryteria pozwalają na zwiększenie wydajności przetwarzania danych masowych oraz poprawę ich dynamicznego wykorzystania w systemach informacji przestrzennej.

4.2.10. Możliwości praktycznego wykorzystania wyników badań

Ostatecznie prezentowane rozwiązanie umożliwia w praktyce wygenerowanie dowolnej wirtualnej przestrzeni roboczej, zbliżonej do przestrzeni rzeczywistej, a następnie wykorzystanie jej właściwości do kontroli dokładności kolejnych faz postępowania oraz dokonania doboru właściwych sekwencji przetwarzania w całym procesie. Zastosowanie struktury GRID na poszczególnych etapach przetwarzania umożliwia uporządkowanie informacji oraz zastosowanie podobnych mechanizmów kontroli pozwalających na porównanie dokładności kolejnych faz procesu. Dzięki temu zoptymalizowane zostają wszystkie etapy przetwarzania poczynając od analizy danych pomiarowych poprzez generowanie GRID aż po wykorzystanie gotowej struktury do różnych zadań. Pozwala to na zapewnienie zarówno dużej dynamiki przetwarzania danych jak i osiągnięcie wysokiego stopnia automatyzacji zadań w kolejnych fazach procesu. Zachowanie spójności na kolejnych etapach przetwarzania zapewnia dostarczenie kompletu danych do wszechstronnych analiz. Z kolei dzięki możliwości kontroli dokładności generowanych modeli na każdym etapie, możliwe staje się wypracowanie założeń zmierzających do uzyskania jak najkorzystniejszego rozwiązania analizowanych problemów. W oparciu o weryfikowalne pod względem dokładnościowym i ekonomicznym etapy, możliwe jest ustalenie wydajnych sposobów postępowania podczas przetwarzania megadanych w procesie tworzenia DTM, zapewniających zarówno wysoką jakość tworzonych modeli jak i znaczną automatyzację całego procesu. Możliwość precyzyjnej kontroli modeli wirtualnych na każdym etapie procesu umożliwia opracowanie spójnych metod postępowania, które mogą zostać wykorzystane do przetwarzania informacji pochodzącej z rzeczywistych pomiarów.

Zastosowanie prezentowanej metodyki oraz rozwiązań algorytmicznych do przetwarzania megadanych w celu utworzenia numerycznego modelu terenu jest tylko jedną z wielu możliwości ich wykorzystania. Uniwersalne podejście do przetwarzania struktury GRID umożliwia wykorzystanie omawianej metodyki oraz systemu do analizowania różnego rodzaju pomiarów, a określone założenia funkcjonalne umożliwiają jego zastosowanie w wielu dziedzinach badawczych. Ze względu na to, że duża ilość informacji, określających różne cechy mierzonych obiektów, rejestrowana jest w podobny sposób (lokalizacja obiektu plus jego cecha), możliwe jest wykorzystanie potencjału prezentowanej metodyki do wszechstronnej

analizy różnych przypadków tego typu. Można ją wykorzystać wszędzie tam, gdzie dane pomiarowe (w postaci punktów rozproszonych) zorganizowane mogą zostać w strukturę trójpolowych rekordów, zawierających informację o opisywanym lub mierzonym obiekcie w postaci współrzędnych (x,y) oraz dowolnego współczynnika z przyporządkowanego do jego poszczególnych punktów. Gdy współczynnik z jest równy wysokości punktu h można wygenerować DTM, jednak gdy z jest określoną cechą obiektu w danym punkcie (x,y) można wygenerować powierzchnię numeryczną stanowiącą reprezentację danych cech w modelowanej przestrzeni. Ze względu na to, że prezentowane rozwiązania stosują algorytmy redukcji ilości danych wykorzystywanych do generowania powierzchni, możliwe jest uzyskanie pożądanej organizacji przestrzennej dla różnego rodzaju informacji przy zachowaniu ustalonej dokładności wynikowej tworzonego modelu. Dzięki uniwersalnej metodzie przetwarzania możliwe jest zastosowanie prezentowanej metodyki do wygenerowania danych wykorzystywanych przy budowie numerycznego modelu powierzchni przez dowolną aplikację SIS dla różnych zjawisk zlokalizowanych w przestrzeni. Pozwala to na implementację metodyki i systemu w wielu dziedzinach naukowych i gospodarczych, gdzie analizie poddawana jest informacja o przestrzennej lokalizacji dowolnych parametrów opisujących badane zjawisko.

Bibliografia

- Amies A., Sluiman H., Tong Q.G., Liu G.N. 2012. *Infrastructure as a Service Cloud Concepts. Developing and Hosting Applications on the Cloud*, IBM Press.
- Andrew S., Tanenbaum A.S. 2010. *Systemy operacyjne*. Wyd. Helion, Gliwice.
- Andrienko N., Andrienko G. 2006. *Exploratory Analysis of Spatial and Temporal Data A Systematic Approach*. Springer, Berlin.
- Anselin L., Syabri I., Kho Y. 2006. *An introduction to spatial data analysis*. GeoDa, Geographical Analysis, 38(1): 5-22.
- Arctur D., Zeiler M. 2004. *Designing Geodatabases. Case Studies in GIS Data Modeling*. ESRI Press, Redlands, CA.
- Axelsson P. 2000. *DEM generation from laser scanner data using adaptive TIN models*. International Archives of Photogrammetry and Remote Sensing, 33(B4/1): 110–117.
- Ballas D., Kingston R., Stillwell J. 2004. *Using a spatial microsimulation decision support system of policy scenario analysis*. Recent Advances in Design and Decision Support Systems in Architecture and Urban Planning, Kluwer, Dordrecht.
- Banachowski L., Diks K., Rytter W. 2006. *Algorytmy i struktury danych*. WNT, Warszawa.
- Banerjee, S., Carlin B., Gelfand A. 2004. *Hierarchical Modeling and Analysis for Spatial Data*. Boca Raton, Chapman and Hall-CRC.
- Beaudoin J.D., Hughes Clarke J.E., Bartlett J.E. 2004. *Application of surface sound speed measurements in postprocessing for multi-sector multibeam echosounders*. International Hydrographic Review, 5(3):26–31.
- Bédard Y., Rivest S., Proulx M.-J. 2007. *Spatial On-Line Analytical Processing (SOLAP). Concepts, Architectures and Solutions from a Geomatics Engineering Perspective*. W: *Data Warehouses and OLAP. Concepts, Architectures and Solutions*. Chap. 13. Red. R. Wrembel, C. Koncilia. IRM Press (Idea Group), London, UK.

- Bielecka E. 2006. *Systemy informacji geograficznej. Teoria i zastosowania*. Wydawnictwo PJWSTK, Warszawa.
- Blythe J. 2005. *Task Scheduling Strategies for Workflow-based Applications in Grids*. IEEE International Symposium on Cluster Computing and the Grid, IEEE Press.
- Bose R., Frew J. 2005. *Lineage retrieval for scientific data processing: a survey*. ACM Computing Surveys, 37: 1-28.
- Bretar F., Chauve A., Bailly J.S., Mallet C., Jacome A. 2009. *Terrain surfaces and 3-D landcover classification from small footprint full-waveform lidar data: application to badlands*. Hydrology and Earth System Sciences, 13: 1531-1545.
- Bu Y., Howe B., Balazinska M., Ernst M.D. 2010. *HaLoop: efficient iterative data processing on large clusters*. Proc. VLDB Endow., 3:285-296.
- Carrier C.G., Povel O. 2003. *Characterising data mining software*. Intelligent Data Analysis, 7:181-192.
- Cavalli M., Mathys N., Jacome A. 2008. *Quality and vertical accuracy assessment of airborne LiDAR data: a case study on the Draix basins*. Geophysical Research Abstracts, EGU General Assembly, 10.
- Chang K. 2007. *Introduction to Geographic Information Systems*. 4th edition. McGraw-Hill, New York.
- Chen C.F., Li Y.Y., Dai H.L. 2011. *An application of Coons patch to generate grid-based digital elevation models*. International Journal of Applied Earth Observation and Geoinformation, 13(1):830-837.
- Cheney J., Chiticariu L., Tan W.C. 2009. *Provenance in databases. Why, how, and where*. Foundations and Trends in Databases, 1(4):379-474.
- Chrisman N.R. 2003. *Exploring Geographical Information Systems*. Wiley, Hoboken, NJ.
- Chrobak D. 2003. *Fortran praktyka programowania*. Mikom, Warszawa.
- Connolly T., Begg C. 2004. *Systemy baz danych*. Wyd. RM, Warszawa.
- Cressie N., Johannesson G. 2008. *Fixed rank kriging for very large spatial data sets*. Journal of the Royal Statistical Society, Series B, Statistical Methodology, 70(1): 209-226.
- Dean J., Ghemawat S. 2008. *MapReduce: Simplified data processing on large clusters*. Commun. ACM, 51(1):107-113.
- Deng Y. 2007. *New trends in digital terrain analysis: landform definition, representation, and classification*. Progress in Physical Geography, 31(4): 405-419.
- Dudek W. 2006. *Bazy danych SQL. Teoria i praktyka*. Helion, Gliwice.
- Erdogan S. 2009. *A comparison of interpolation methods for producing digital elevation models at the field scale*. Earth Surface Processes and Landforms, 34(3):366-376.
- Fahringer T. 2005. *ASKALON: a tool set for cluster and Grid computing. Concurrency and Computation. Practice and Experience*, Wiley Inter Science, 17:143-169.
- Finley A., Sang H., Banerjee S., Gelfand A. 2009. *Improving the performance of predictive process modeling for large datasets*. Computational Statistics and Data Analysis, 53(8): 2873-2884.
- Fischer M.M. 2006. *Spatial Analysis and GeoComputation*. Springer, Berlin.

- Fornberg B., Wright G. 2004. *Stable computation of multiquadric interpolants for all values of the shape parameter*. Computers Math. Applic, 48(5/6): 853-867.
- Fortuna Z., Macukow B., Wąsowski J. 2006. *Metody numeryczne*. WNT, Warszawa.
- Goodchild M.F. 2009. *Geographic information systems and science: today and tomorrow*. Annals of GIS, 15(1): 3-9.
- Grębosz J. 1999. *Symfonia C++*. Oficyna Kallimach, Kraków.
- Griffith D. 2004. *Faster maximum likelihood estimation of very large spatial autoregressive models: An extension of the Smirnov-Anselin result*. Journal of Statistical Computation and Simulation, 74(12): 855-866.
- Grządziel A. 2004. *Pomiary batymetryczne – dawniej i dziś*. Przegląd Morski, 4.
- Guo D., Gahegan M., MacEachren A.M., Zhou B. 2005. *Multivariate analysis and geovisualization with an integrated geographic knowledge discovery approach*. Cartography and Geographic Information Science, 32(2): 113-132.
- Ha S.H., Bae S.M., Park S.C. 2006. *Digital content recommender on the Internet*. IEEE Intelligent Systems, 21: 70-77.
- Hamilton T., Beaudoin J. 2010. *Modelling uncertainty caused by internal waves on the accuracy of MBES*. International Hydrographic Review, 11.
- Harris R., Sleight P., Webber R. 2005. *Geodemographies, GIS and Neighbourhood Targeting*. Wiley, Chichester, UK.
- Hastie T., Tibshirani R., Friedman J. 2009. *The Elements of Statistical Learning*. Data Mining, Inference, and Prediction, Springer Publishing Company, New York.
- Higdon D. 2002. *Space and space-time modeling using process convolutions*. W: *Quantitative Methods for Current Environmental Issues*. Red. C.W. Anderson, V. Barnett, P.C. Chatwin, A.H. El-Shaarawi. Springer-Verlag, London, UK.
- Hindman B., Konwinski A., Zaharia M., Stoica I. 2009. *A common substrate for cluster computing*. Proceedings of USENIX HotCloud'09: Workshop on Hot Topics in Cloud Computing.
- Howard P.G., Vitter J.S. 1994. *Arithmetic Coding for Data Compression*. Proceedings of the IEEE, 82(6): 857-865.
- Huang Y., Pei J., Xiong H. 2006. *Mining colocation patterns with rare events from spatial data sets*. Geoinformatica, 10(3): 239-260.
- Jankowsky J.M. 2009. *Przegląd metod i algorytmów numerycznych*. WNT, Warszawa.
- Johannesson G., Cressie N., Huang H. C. 2007. *Dynamic multiresolution spatial models*. Environmental and Ecological Statistics, 14: 5-25.
- Józwiak J., Podgórski J. 2000. *Statystyka od podstaw*. PWE, Warszawa.
- Keim D.A., Panse C., Sips M. 2004. *Visual data mining in large geospatial point sets*. IEEE Computer Graphics and Applications, 24(5): 36-44.
- Knuth D. 1998. *The Art of Computer Programming*. Vol. 3. *Sorting and Searching*. Addison-Wesley, Reading, Massachusetts.

- Kobler A., Pfeifer N., Ogrinc P., Todorovski L., Ostir K., Dzeroski S. 2007. *Repetitive interpolation: a robust algorithm for DTM generation from aerial laser scanner data in forested terrain*. Remote Sensing of Environment, 108: 9-23.
- Kopczewska K. 2006. *Ekometria i statystyka przestrzenna*. CeDeWu, Warszawa.
- Kott R.K., Walczak K. 1991. Programowanie w języku Fortran 77. Wydawnictwa Naukowo-Techniczne, Warszawa.
- Kraus K., Pfeifer N. 2001. *Advanced DTM generation from LIDAR data*. International Archives of Photogrammetry and Remote Sensing, XXXIV(3/W4): 23-30.
- Kwiatkowska M. 2007. *Systemy wspomagania decyzji. Jak korzystać z wiedzy i informacji*. Mikom, Warszawa.
- Larsson E., Fornberg B. 2005. *Theoretical and computational aspects of multivariate interpolation with increasing flat basis functions*. Computers Math. Applic, 49(1): 103-130.
- Li Z., Zhu Q., Gold C. 2005. *Digital Terrain Modeling: Principles and Methodology*. CRC Press, Boca Raton, FL.
- Lihong W., Wenhai X., Wenbo W. 2013. *Survey of Seafloor Targets with Varied Sizes by Multi Beam Sonar in Different Depth Water*. Applied Mechanics and Materials, 263-266: 909-914.
- Ling L., Kansa E.J. 2004. *Preconditioning for radial basis functions with domain decomposition methods*. Mathematical and Computer Modelling, 40(13): 1413-1427.
- Liu X. 2008. *Airborne LiDAR for DEM generation: some critical issues*. Progress in Physical Geography, 32: 31-49.
- Liu H., Yang X. 2012. *Software reuse in the emerging cloud computing era*. Hershey, PA: Information Science Reference, 9: 204-227.
- Longley P.A., Batty M. 2003. *Advanced Spatial Analysis*. The CASA Book of G/S. Redlands, CA: ESRI Press.
- Longley P.A., Goodchild M.F., Maguire D.W., Rhind D.W. 2005. *Geographical Information Systems: Principles, Techniques, Management and Applications*. Wiley, Hoboken, NJ.
- Longley P.A., Goodchild M.F., Maguire D.J., Rhind D.W. 2011. *Geographical information systems and science*. Third Edition. Wiley, Hoboken, NJ.
- Makowski A. 2005. *Systemy informacji topograficznej kraju*. Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa.
- Mallet C., Bretar F. 2009. *Full-waveform topographic Lidar*. State-of-the-Art., ISPRS Journal of Photogrammetry and Remote Sensing, 64: 1-16.
- Masser L. 2005. *Spatial Data Infrastructure: An Introduction*. ESRI Press, Redlands, CA.
- Maune D.F. 2001. *Digital Elevation Model Technologies and Applications*. The DEM Users Manual, American Society for Photogrammetry and Remote Sensing, Bethesda, MD.
- Miller H., Han J. 2009. *Geographic data mining and knowledge discovery: An overview*. W.: *Geographic data mining and knowledge discovery*. Red. H. Miller, J. Han. CRC Press, Taylor and Francis Group.
- Mitasova H., Hofierka J. 1993. *Interpolation by regularized spline with tension. Application to terrain modeling and surface geometry analysis*. Mathematical Geology, 25(6): 657-669.

- Nisan N., Schocken S. 2008. *Elementy systemów komputerowych*. WNT, Warszawa.
- Nyerges T. 2001. *Geographic Information Systems for Group Decision Making*. Taylor and Francis, London.
- O'Sullivan D., Unwin D.J. 2003. *Geographic Information Analysis*. Wiley, Hoboken, NJ.
- Paine D.P., Kiser J.D. 2003. *Aerial Photography and Image Interpretation*. Wiley, Hoboken, NJ.
- Paradysz J. 2005. *Statystyka*. Wyd. AE, Poznań.
- Pelikant A. 2009. *Bazy danych*. Helion, Gliwice.
- Peng Z.H., Tsou M.H. 2003. *Internet GIS. Distributed Geographic Information Services for the Internet and Wireless Networks*. Wiley, Hoboken, NJ.
- Pigozzi B.W. 2004. *A hierarchy of spatial marginality through spatial filtering*. Professional Geographer, 56(4): 460–470.
- Press W.H., Metcalf M., Teukolsky S.A. 2014. *Numerical Recipes in FORTRAN 90*. Cambridge University Press.
- Proulx M.J., Bédard Y. 2008. *Fundamental Characteristics of Spatial OLAP Technologies as Selection Criteria*. Location Intelligence April 29, Santa Clara, USA.
- Przelaskowski A. 2005. *Kompresja danych: podstawy, metody bezstratne, kodery obrazów*. BTC, Warszawa.
- Raaflaub L.D., Collins M.J. 2006. *The effect of error in gridded digital elevation models on the estimation of topographic parameters*. Environmental Modelling and Software, 21(5): 710–732.
- Reutebuch S.E., Andersen H.E., McGaughey R.J. 2005. *Light detection and ranging (LIDAR). An emerging tool for multiple resource inventory*. Journal of Forestry, 103(6): 286-292. .
- Rippa S. 1999. *An algorithm for selecting a good value for the parameter c in radial basis function interpolation*. Advances in Computational Mathematics, 11(2-3):193-210.
- Rissanen J., Landgon G.G. 1990. *A General Approach to Data Compression: Binary Arithmetic Codes*. IBM Research Report RJ 7443, IBM Almaden Research Center, San Jose, CA, USA.
- Rivest S., Bédard Y., Proulx M.J., Nadeau M., Hubert F., Pastor J. 2005. *Solap technology: merging business intelligence with geospatial technology for interactive spatio-temporal exploration and analysis of data*. ISPRS Journal of Photogrammetry & Remote Sensing, 60(1): 17-33.
- Rościszewski K., Witt R. 2003. *Nauka Fortran 90/95*. Wydawnictwo UJ, Kraków.
- Sayood K. 2002. *Kompresja danych wprowadzenie*. RM, Warszawa.
- Schabenberger O., Gotway C.A. 2005. *Statistical Methods for Spatial Data Analysis*. Chapman and Hall, CRC: Boca Raton, FL.
- Shi W., Zheng S., Tian Y. 2009. *Adaptive mapped least squares SVM-based smooth fitting method for DSM generation of LIDAR data*. International Journal of Remote Sensing, 30(21): 5669-5683.
- Soares T. 2006. *Deductive database. Implementatio parallelism and applications*, ICLP, Springer-Verlag, Berlin, Heidelberg.
- Sobczyk M. 2006. *Statystyka*. Wydawnictwo Naukowe PWN, Warszawa.

- Sobieski W. 2008. *GNU Fortran z elementami wizualizacji danych*. Wyd. UWM, Olsztyn.
- Stein M. 2008. *A modeling approach for large spatial datasets*. Journal of the Korean Statistical Society, 37(1): 3-10.
- Stroud J.R., Müller P., Sansó B. 2001. *Dynamic models for spatiotemporal data*. Journal of the Royal Statistical Society, Ser. B, Statistical Methodology, 63: 673-689.
- Suchecky B. 2010. *Ekonometria przestrzenna*. C.H. Beck, Warszawa.
- Sysło M.M. 2002. *Algorytmy*. WSiP, Warszawa.
- Takefusa A. 2001. *Bricks: A Performance Evaluation System for Scheduling Algorithms on the Grids*. JSPS Workshop on Applied Information Technology for Science.
- Tay L.T., Sagar B.S.D., Chuah H.T. 2005. *Analysis of geophysical networks derived from multiscale digital elevation models: a morphological approach*. IEEE Geoscience and Remote Sensing Letters, 2(4): 399-403.
- Thomas J.J., Cook K.A. 2005. *Illuminating the path: The research and development agenda for visual analytics*. IEEE Computer Society, Los Alamitos, CA.
- Travostino F. 2006. *Grid Networks*. J. Wiley and Sons, New York.
- Trykozko A. 1999. *Ćwiczenia z języka Fortran*. Mikom, Warszawa.
- Tzeng S., Huang H.-C., Cressie N. 2005. *A fast, optimal spatial-prediction method for massive datasets*. Journal of the American Statistical Association, 100: 1343-1357.
- Van Niel T.G., McVicar T.R., Li L., Gallant J.C., Yang Q. 2008. *The impact of misregistration on SRTM and DEM image differences*. Remote Sensing of Environment, 112(5): 2430-2442.
- Vázquez J.P., Pascual J.P. 2008. *Automated spot heights generalization in trail maps*. International Journal of Geographical Information Science, 22(1): 91-110.
- Voorsluys W., Broberg J., Buyya R. 2011. *Introduction to Cloud Computing. W: Principles and Paradigms. Cloud Computing*. Red. R. Buyya, J. Broberg, A. Goscinski. Wiley Press, New York.
- Wack R., Wimmer A. 2002. *Digital terrain models from airborne laser scanner data – a grid based approach*. International Archives of Photogrammetry and Remote Sensing, 34(3B):293- 296.
- Wechsler S.P. 2003. *Perceptions of digital elevation model uncertainty by DEM users*. URISA Journal 15(2): 57–64.
- Wilson J.P. 2012. *Digital terrain modeling*. Geomorphology, 137: 107-121.
- Worboys M.F., Duckham M. 2004. *GIS: A Computing Perspective*. CRC Press, Boca Raton, FL.
- Yan J., Thill J.-C. 2009. *Visual data mining in spatial interaction analysis with self-organizing maps*. Environment and Planning, B, 36: 466-486.
- Zhang X., Pazner M. 2004. *The icon imagemap technique for multivariate geospatial data visualization*. Approach and software system, Cartography and Geographic Information Science, 31 (1): 29-41.
- Zhou Q., Chen Y. 2011. *Generalization of DEM for terrain analysis using a compound method*. ISPRS Journal of Photogrammetry and Remote Sensing, 66(1): 38-45.

5. Omówienie pozostałych osiągnięć naukowo-badawczych

5.1. Charakterystyka wybranych prac naukowo-badawczych

W roku 1991 ukończyłem studia na Wydziale Geodezji i Urzędzeń Rolnych Akademii Rolniczo-Technicznej w Olsztynie uzyskując tytuł zawodowy magistra inżyniera w zakresie geodezji urzędzeń rolnych. Tematem mojej pracy magisterskiej było opracowanie komputerowego systemu edukacyjnego „*Wcięcia geodezyjne*”. Posiadając pięcioletnie doświadczenie w programowaniu komputerów (programowaniem zajmuję się od 1986 r.) utworzyłem system, w skład którego wchodziło 12 programów komputerowych pozwalających na zaprezentowanie wiedzy z wybranej dziedziny geodezji, sprawdzenie jej podczas testu, a następnie umożliwienie dokonania obliczeń praktycznych. Programy te w pierwszych latach mojej pracy w Instytucie Geodezji i Fotogrametrii (od 1991 r) były rozwijane, wdrażane i wykorzystywane w kształceniu i egzaminowaniu studentów. W początkowym okresie mojej pracy w Instytucie Geodezji i Fotogrametrii tematyka moich badań koncentrowała się na opracowywaniu metodyki pozyskiwania informacji z map nawigacyjnych i hydrograficznych w procesie digitalizacji i skanowania z wektoryzacją. Informacje pozyskiwane w ten sposób wykorzystywane były do tworzenia map numerycznych m.in. batymetrii dna morskiego. Zasoby rastrowe i wektorowe tworzone na bazie takich map zasilaty generowaną na ich podstawie mapę numeryczną. W tym okresie prowadziłem również badania w ramach projektów badawczych: KBN Nr 9-0727-91-01 „*Ocena zmian środowiska na podstawie wyników przetwarzania komputerowej informacji pozyskanych z dawnych map*” oraz KBN Nr 9T 12E 027 08 „*Systemy pozyskiwania i przetwarzania informacji graficznych o terenie i środowisku*” (załącznik 5, rozdz. 4, pkt. 1-2). W ramach tych projektów powstało kilka publikacji, które dodatkowo zostały zaprezentowane w referatach wygłoszonych na konferencjach naukowych (zał. 5, rozdz. 6, pkt. 1-6). Rezultatem prac nad projektami badawczymi było również utworzenie szeregu programów komputerowych wspomagających analizy badawcze (np. *HPCAD* - system przygotowania obrazów rastrowych do wektoryzacji, *WEKT-EKO* – system komputerowego przetwarzania i analizy informacji graficznych czy *KRESL* – program sterujący kreśleniem na ploterze wybranych elementów mapy numerycznej) (zał. 5, rozdz. 14, pkt. 3-6). W tym okresie prowadziłem również prace nad automatyzacją i oceną dokładności procesów przetwarzania graficzno-numerycznego. W celu przeprowadzenia badań utworzyłem pakiet programów wchodzący w skład systemu *RIDA (Raster Images Deformations Analysis - Analiza deformacji obrazów rastrowych)*, który przeznaczony jest do rejestracji, oceny i prezentacji zniekształceń powstających w procesie digitalizacji i skanowania z wektoryzacją (zał. 5, rozdz. 14, pkt. 8-9, 11). Wyniki tych badań zaprezentowałem na konferencjach naukowych (zał. 5, rozdz. 6, pkt. 7-10).

Naszkicowana poniżej tematyka moich publikacji odzwierciedla główne etapy prac badawczych prowadzonych po uzyskaniu stopnia naukowego doktora. Publikacje charakteryzujące etapy i zakres badań w dalszym ciągu tego rozdziału przedstawiane będą w postaci [N], gdzie N jest kolejną liczbą naturalną. Wykaz omawianych publikacji znajduje się w rozdziale 5.2. Pozostałe publikacje zamieszczone zostały w wykazie publikacji stanowiącym załącznik 5.

W publikacji [1] przedstawiłem główne założenia i metodykę badań możliwych do realizacji przy wykorzystaniu pakietu RIDA. Pakiet pozwala na zautomatyzowanie, a zatem również znaczne przyspieszenie, procesu kontroli urządzeń służących pozyskiwaniu informacji z materiałów analogowych. System korzysta z danych analogowych, które mogą być zawarte w specjalnie opracowanym modelu testowym lub mogą być pozyskiwane z dowolnych, przewidywanych do testowania, materiałów graficznych. Przy jego pomocy można wygenerować dane służące badaniu dokładności informacji pozyskiwanych drogą digitalizacji, ocenie błędów osobowych poszczególnych obserwatorów oraz określeniu stałego modelu błędów digimetru. Pozwala on również przygotować dane służące badaniu dokładności informacji pozyskiwanych w procesie skanowania i wektoryzacji, określenie zniekształceń obrazu rastrowego powodowanych przez skaner oraz efektywności zastosowanych algorytmów kalibracji rastra oferowanych przez specjalistyczne systemy oprogramowania. W rezultacie na podstawie informacji pochodzących z analiz materiałów analogowych oraz pomiaru modelu wzorcowego możliwe jest dokonanie transformacji współrzędnych punktów kontrolnych oraz wprowadzenie do nich poprawek w oparciu o zbiór korekcyjny. System *RIDA* był sukcesywnie rozwijany i wykorzystywany w dalszej pracy naukowo-badawczej.

W kolejnym etapie badań opracowałem metodykę kontroli skanerów stosowanych do pozyskiwania informacji graficznych w procesie tworzenia map numerycznych. Rezultaty badań zaprezentowane zostały w publikacji [2]. Zaproponowana metodyka kontroli skanerów wielkoformatowych pozwala na określenie charakteru zniekształceń wnoszonych przez skaner, obszarów ich występowania, wielkości tych zniekształceń i ich znaku w stosunku do osi układu współrzędnych skanowanego arkusza. Z przeprowadzonych badań wynika, że proces skanowania wprowadza zniekształcenia rastra, których charakter jest ściśle związany z cechami konstrukcyjnymi skanerów (mechanicznymi, elektronicznymi i optycznymi). Kalibracja softwarowa skanerów w oparciu o dostarczany z nimi wzorec nie zawsze eliminuje niedoskonałość tych urządzeń w stopniu wystarczającym dla potrzeb opracowań geodezyjnych. Opracowałem metodę, w której oprócz klasycznej kalibracji obrazów rastrowych, błędy wnoszone przez skanery w procesie pozyskiwania informacji z materiałów analogowych mogą być także korygowane w oparciu o pomiary kontrolne dokonywane dla każdego urządzenia indywidualnie.

W publikacji [3] przedstawiłem założenia graficznej analizy błędów występujących w procesie przetwarzania graficzno-numerycznego i numeryczno-graficznego. Badania i analizy obejmowały zarówno zniekształcenia wnoszone do modelu numerycznego przez skanery i digimetry jak również błędy powodowane przez źle skalibrowany obraz rastrowy. Badałem także precyzję ploterów pisakowych i rastrowych oraz dokładność tworzonych przez nie materiałów analogowych. Analizami objąłem również sam materiał analogowy, który również często ulega deformacji tracąc właściwą geometrię obrazu. Zaproponowana analiza graficzna pozwala w prosty i czytelny sposób skontrolować i zinterpretować zniekształcenia powstające w różnych etapach procesu tworzenia mapy numerycznej. Można ją stosować jako uzupełnienie klasycznych metod analitycznych lub jako niezależną kontrolę na poszczególnych etapach procesu produkcyjnego. Dane uzyskane z procesu kontroli graficznej umożliwiły

wprowadzenie odpowiednich korekt do procesu produkcyjnego podnosząc tym samym dokładność przetwarzania informacji.

Kolejny problem badawczy podjęty przeze mnie w publikacji [4] dotyczył opracowania metod oceny i minimalizacji składowych zniekształceń obrazów rastrowych w systemach przetwarzania informacji przestrzennych. W zaprojektowanym schemacie przetwarzania graficzno-numerycznego pokazałem zasady wyodrębniania składowych błędów w celu rozszerzenia możliwości ich efektywnej minimalizacji. Po wyselekcjonowaniu składowych zniekształceń wnoszonych przez skanery następowała ich eliminacja przy pomocy modelu korekcyjnego. W celu wyeliminowania składowych zniekształceń wnoszonych przez skanery do tworzonych przez nie plików bitmapowych, podjąłem próbę ustalenia stałego modelu kompozycji błędów i wykorzystania go do redukcji zniekształceń geometrii generowanych obrazów rastrowych. Następnie opracowałem modele korekcyjne badanych typów skanerów i przy ich pomocy wyeliminowałem składowe systematyczne zniekształcenia obrazów rastrowych. Zastosowanie takiego rozwiązania wydatnie usprawniło korekcję geometrii obrazów rastrowych, podnosząc tym samym dokładność pozyskiwanych informacji. Określenie badanych zniekształceń, pozwoliło na ustalenie obszarów odpowiedniej lokalizacji punktów dostosowania na danym materiale analogowym, co z kolei warunkuje prawidłową korektę obrazu rastrowego metodami softwarowymi.

Artykuł [5] kończy grupę publikacji związaną z przetwarzaniem graficzno-numerycznym i numeryczno-graficznym. Przedstawiłem w nim zagadnienia dotyczące analizy i prognozy wielkości zniekształceń obrazów rastrowych z wykorzystaniem szeregów czasowych. W czasie badań stwierdziłem, że systematyczny charakter zniekształceń wnoszonych przez skanery oraz możliwość ich okresowej redukcji w drodze kalibracji urządzeń pozwala na opracowanie założeń dotyczących dokładności pozyskiwania informacji w czasie. Analizując dane ze skaningu twierdziłem, że proces rozkalibrowywania się skanera ma charakter okresowy. Ponadto zniekształcenia na poszczególnych punktach, w kolejnych cyklach wykazują wyraźne podobieństwo. Podczas przeprowadzonych analiz opisałem przebieg oscylacji zniekształceń przy pomocy szeregu czasowego. W badaniach wykorzystałem modele addytywne i autoregresyjne, które po zdefiniowaniu prawidłowego okresu oscylacji, pozwalają wyznaczyć prognozę zmian na podstawie szeregu dokonanych pomiarów wartości rzeczywistych. Sporządzenie prognoz z wykorzystaniem szeregów czasowych umożliwiło ustalenie okresów prawidłowej pracy skanerów między kalibracjami, w których zniekształcenia powodowane przez te urządzenia można uznać za dopuszczalne w tworzonym modelu numerycznym.

Kolejna grupa publikacji zawiera wyniki badań dotyczących przetwarzania pomiarów rejestrowanych masowo przez urządzenia o działaniu ciągłym oraz wykorzystania struktury GRID w różnych dziedzinach badawczych. Wszystkie badania poprzedzane były opracowaniem autorskich aplikacji użytkowych, które posłużyły do przeprowadzania analiz rozpatrywanych problemów badawczych. W ramach prowadzonych badań od 2001 do 2014 powstały kolejne moduły pakietu *GRID* (zał. 5, rozdz. 14, pkt. 13-27), który w roku 2015 został przekształcony w system *SYPMEG* (**S**ystem of **P**rocessing of **M**egadata with using the **GRID** structure - System przetwarzania megadanych z wykorzystaniem struktury *GRID*) (zał. 5, rozdz. 14, pkt. 28). W

tym okresie prowadziłem również badania w ramach projektów badawczych: KBN Nr 9T 12C OS2 13 „*Opracowanie założeń systemu trójwymiarowej morskiej mapy numerycznej do zastosowań specjalnych*” oraz KBN Nr 0 T00A 012 21 „*Opracowanie projektu systemu dynamicznego pozycjonowania okrętu metodami porównawczymi*” (załącznik 5, rozdz. 4, pkt. 3-4). Okres ten skutkuje opracowaniem szeregu publikacji, których rezultaty były przedstawiane w formie referatów na międzynarodowych i krajowych konferencjach naukowych (zał. 5, rozdz. 6, pkt. 11-29).

W artykule [6] przedstawiliśmy wybrane aspekty przetwarzania wyników pomiaru ukształtowania dna rejestrowanych przez urządzenia o działaniu ciągłym. Artykuł zawiera koncepcję oraz ogólne założenia przetwarzania zbiorów obserwacji, wykorzystywanych do generowania numerycznego modelu dna, będących wynikiem pomiaru sondą wielowiązkową MBES (Multi Beam Echo Sounder). Analizowano w nim sposoby wstępnego przetwarzania danych pomiarowych, które posłużyć mogą do budowy modelu numerycznego powierzchni. Po eliminacji błędów grubych wyniki pomiarów wykorzystywane były do określenia charakterystyki ilościowej i jakościowej ukształtowania dna, a następnie zastosowane do wyznaczania parametrów przetwarzania danych w kolejnych etapach. Zaproponowano sposoby dostosowania wielkości zbiorów i ich wewnętrznej struktury do wymagań systemów przestrzennych. Wykorzystano do tego generowanie regularnej siatki trójkątów równobocznych, regularnej siatki kwadratów oraz liniowe modele profilowe eliminacji punktów. Zaproponowano przykłady profilowej generalizacji ilościowej polegające na proporcjonalnym zmniejszeniu liczby punktów opisujących dany obiekt. Opracowano również metody pozwalające na wierne odwzorowanie kształtu profilu, wzdłuż którego dokonywana jest analiza przy jednoczesnej maksymalnej redukcji liczby punktów. Proponowane metody umożliwiają profilową redukcję liczby punktów przy jednoczesnym zachowaniu ustalonej dokładności, pozwalając jednocześnie na zachowanie lokalnych różnicowań rzeźby odtwarzanej powierzchni.

Z kolei w pracy [7] zaprezentowałem analizę wpływu lokalizacji punktów pomiarowych pozyskanych przez MBES na wyniki generowane przez wybrane algorytmy interpolacyjne. W badaniach wykazałem, że liczba interpolowanych węzłów pozostawionych w bazie zależy głównie od możliwości wyszukania wokół nich punktów pomiarowych, które będą spełniały odpowiednie wymogi lokalizacyjne dla danej metody obliczeniowej. Przy dużym zagęszczeniu punktów pomiarowych zwiększa się prawdopodobieństwo ich prawidłowego położenia wokół tworzonego węzła, co umożliwia zastosowanie wydajnych algorytmów obliczeniowych. Stwierdziłem, że pozwalają one na wyznaczenie wartości w węźle na podstawie minimalnej liczby, najbliższych położonych punktów pomiarowych przy jednoczesnym podniesieniu jakości generowanego modelu powierzchni. Stwierdziłem także, że wykorzystanie dowolnie rozmieszczonych punktów pomiarowych do obliczenia wartości w węźle może prowadzić do nieprawidłowych rozwiązań. Z tego powodu szczególnie ważne jest ustalenie ich właściwej lokalizacji przez prawidłowe zdefiniowanie przestrzeni poszukiwań. W wyniku przeprowadzonych analiz zaproponowałem sposoby lokalizacji punktów pomiarowych wybieranych do wyznaczenia węzłów. Poprawne ustalenie parametrów przetwarzania

przyczyniło się do zwiększenia odnalezionych prawidłowych układów węzeł-punkty pomiarowe, a tym samym do podniesienia dokładności generowanej powierzchni numerycznej.

Kolejny problem badawczy dotyczył generowania numerycznego modelu terenu na podstawie wyników pomiaru naziemnym skanerem laserowym TLS (Terrestrial Laser Scanning). W publikacji [8] przedstawiliśmy analizę i rozwiązanie problemów wynikających ze specyfiki profilowej rejestracji punktów dokonywanej przez pomiar LiDAR (Light Detection And Ranging). Cechą charakterystyczną tak realizowanych pomiarów jest to, że im wyznaczane punkty są bardziej oddalone od instrumentu pomiarowego, tym wzajemna odległość profili jest większa. Na podstawie analiz stwierdzono, że wygenerowanie modelu DTM (Digital Terrain Model) przy zastosowaniu struktury TIN (Triangulated Irregular Network) na podstawie tak pozyskanych danych jest niekorzystne, ponieważ szczególnie trójkąty o niewielkich rozmiarach i o kształtach znacznie odbiegających od regularnych tworzą lokalne zniekształcenia. W badaniach stwierdzono, że interpolacja punktów pośrednich może w takiej sytuacji być obciążona większymi błędami, a układ trójkątów szczególnie wzdłuż profili pomiarowych, ma istotny wpływ na przebieg warstw. W wyniku analiz zaproponowano metody pozwalające na redukcję zniekształceń modelu powierzchni przez wykorzystanie algorytmów interpolacyjnych. Pozwalają one na uzyskanie ciągłości modelu przez kompensację zniekształceń pomimo liniowego układu punktów pomiarowych, który jest charakterystyczny przy skaningu laserowym.

Kolejny etap badań obejmował ocenę elementów konstrukcyjnych obiektów budowlanych z wykorzystaniem numerycznego modelowania powierzchni. W publikacji [9] przeanalizowaliśmy możliwości zastosowania struktury GRID do oceny odchylenia powierzchni ściany budynku od płaszczyzny pionowej. Budynek został pomierzony w technologii TLS, a wyniki pomiaru opracowano stosując specjalnie opracowane założenia konstrukcyjne struktury jednorodnych węzłów. W wyniku analiz stwierdzono, że lokalizacja punktów pomiarowych na badanym obiekcie zależy nie tylko od rozdzielczości skanowania ale również wynika z niedokładności spowodowanej różnym kątem odbicia promienia lasera od zróżnicowanej faktury mierzonej powierzchni. Dodatkowo punkty pomiarowe nie są współliniowe co wynika z technologii zastosowanego pomiaru. Skutkuje to utworzeniem niejednorodnej powierzchni modelu przestrzennego charakteryzującej się niewielkimi losowymi zniekształceniami, które przyczyniają się do nierównomiernych i przypadkowych odchylenia położenia punktów pomiarowych od płaszczyzny w kierunku każdej ze współrzędnych przestrzennych. Podczas analiz stwierdziliśmy, że precyzyjna analiza kształtu mierzonej powierzchni w założonych punktach teoretycznych jest utrudniona ze względu na nierównomierne rozmieszczenie mierzonych punktów w przestrzennym układzie odniesienia. Ustalono, że zróżnicowanie położenia poszczególnych punktów prowadzi do nieregularności linii dzielących przedziały klasowe i uniemożliwia dokładne ich ustalenie. W pracy zaproponowano rozwiązanie, które wykorzystując strukturę GRID pozwoliło na wyeliminowanie większości problemów. Regularna struktura węzłów umożliwiła utworzenie modelu, który aproksymuje kształt powierzchni ściany budynku i pozwala na precyzyjne

prorowadzenie analiz. Wykazano, że zaproponowany sposób pozwala skompensować nieregularności modelu TIN przy jednoczesnym zachowaniu kształtu badanej powierzchni. Ponadto zaproponowano rozwiązanie wykorzystujące naturalne uszeregowanie węzłów w rzędy poziome i pionowe do analiz badanej powierzchni przy pomocy przekrojów. Dodatkowo dzięki regulacji rozdzielczości siatki węzłów przekroje mogą być zlokalizowane wzdłuż punktów teoretycznych w dowolnym miejscu obiektu.

Równolegle prowadziłem prace dotyczące walidacji technik interpolacyjnych podczas wyznaczania całkowitej zawartości elektronów (TEC - Total Electron Content) w jonosferze. Wyniki tych analiz przedstawiliśmy w artykule [10]. Przeanalizowano w nim techniki interpolacji oparte na dopasowaniu danych dwuwymiarowych wielomianem lokalnym i globalnym oraz dopasowaniu powierzchni lokalnych, a także interpolacji zależnej od odległości punktów pomiarowych. Do modelowania jonosferycznego wykorzystano dane GPS o podwójnej częstotliwości z polskiego systemu GBAS (ASG-EUPOS). Za pomocą kilku technik interpolacyjnych, utworzono regularną siatkę GRID pionowych wartości TEC o rozdzielczości przestrzennej 0,25 stopnia dla szerokości i długości geograficznej. Ta siatka została następnie wykorzystana do tworzenia lokalnych map TEC z 5-minutową rozdzielczością czasową, a także do tworzenia korekt opóźnień jonosferycznych do pozycjonowania GPS. Wyniki testów numerycznych wykazały, że możliwe jest wygenerowanie siatki korekty jonosferycznej do precyzyjnego pozycjonowania GNSS, opierając się jedynie na lokalnych danych dotyczących sieci odniesienia (ASG-EUPOS). Przedstawione testy wykazują, że technika interpolacji płaszczyzną lokalną oferuje najbardziej obiecujące wyniki, z dokładnością korekty nieco lepszą niż ta, oferowana przez mapy CODE.

W kolejnym artykule [11] przedstawiłem analizy dotyczące sektorowej lokalizacji punktów pomiarowych na dokładność generowanych powierzchni interpolacyjnych. Zastosowana metodyka badań pozwoliła na dokonanie interpretacji wyników i wybranie optymalnych parametrów interpolacji dla poszczególnych algorytmów przy zastosowaniu zróżnicowanych morfologicznie modeli powierzchni. Podczas tworzenia numerycznych modeli terenu przy pomocy struktury GRID istotną rolę odgrywa prawidłowe wyselekcjonowanie punktów pomiarowych biorących udział w obliczeniach. Selekcji takiej można dokonać przy pomocy sektorów. Z przeanalizowanych przykładów wynika, że optymalnym rozwiązaniem jest ustalenie czterech sektorów poszukiwań dla większości przebadanych algorytmów interpolacyjnych, niezależnie od ukształtowania morfologicznego powierzchni. Podobna sytuacja ma również miejsce po wyeliminowaniu wpływu węzłów brzegowych i dopasowaniu wielkości siatki bazowej rozmieszczenia węzłów do morfologii danej powierzchni przez wybór podobszaru w dowolnym miejscu obiektu. Wykazałem, że podczas prowadzenia badań zmierzających do ustalenia dokładności interpolacji należy uwzględnić wpływ efektów brzegowych. Dokładność interpolacji przy zastosowaniu optymalnej liczby sektorów poszukiwań została potwierdzona w praktyce podczas interpolacji na rzeczywistym numerycznym modelu terenu. Zastosowanie optymalnej liczby czterech sektorów poszukiwań dla każdego węzła pozwoliło na uzyskanie pożądanej jakości DTM przy jednoczesnym zoptymalizowaniu czasu pracy niezbędnego do ich opracowania.

Kolejnym z analizowanych aspektów było wykorzystanie map numerycznych w planowaniu, projektowaniu i geodezyjnej obsłudze inwestycji. Wyniki analiz przedstawiliśmy w artykule [12]. Zaprezentowane przykłady wykazały, że zaprojektowanie sprawnego i efektywnego systemu wymaga podjęcia decyzji dotyczących formy, konstrukcji, zawartości i sposobów tworzenia mapy numerycznej. Analizie poddano efektywność systemów informacji przestrzennych, która w znacznym stopniu zależy od zaprojektowanych konstrukcji baz danych oraz organizacji dostępu do zgromadzonych w nich informacji. W szczególności dotyczy to systemów o charakterze dynamicznym, do których należą systemy rejestracji, gromadzenia, prezentacji i oceny zmian przestrzennego układu środowiska oraz badania przemieszczeń i odkształceń. W analizowanych przykładach przedstawiono podstawowe problemy związane z opracowaniem modernizacji dróg w układach trójwymiarowych. Wykazano, że wartość użytkowa systemu jest zachowana jedynie w przypadku, gdy wprowadzenie jakichkolwiek zmian w dowolnym elemencie powoduje uaktualnienie wszystkich pozostałych, a więc przy zachowaniu związków funkcjonalnych i geometrycznych. W szczególności dotyczy to zdefiniowania układu współrzędnych opracowania, ustalenia jednostek kątowych i liniowych oraz zaimportowania zbiorów punktu będących podstawą geometryczną projektu do systemu przetwarzania danych. Kolejne analizy obejmowały opracowanie szczegółowego projektu podziału działek z układami komunikacyjnymi oraz siecią uzbrojenia terenu, utworzenie DTM na podstawie danych wysokościowych zawartych w mapie zasadniczej, konwersję sieci uzbrojenia terenu do układu trójwymiarowego oraz zaprojektowanie systemu odwadniającego. Należy podkreślić, że struktura mapy numerycznej i elementów projektu umożliwia utworzenie funkcjonalnych powiązań aktualizowanych automatycznie. Stwarza to warunki i możliwości opracowania wielu wariantów i wybór najlepszego z uwzględnieniem wcześniej założonych kryteriów. W pracy przeanalizowano również wykorzystanie mapy numerycznej w ocenie stanu obiektów inżynierskich. Jako przykład ilustrujący możliwości wykorzystania mapy numerycznej w badaniu deformacji obiektów inżynierskich wybrano most, który został pomierzony technologią TLS. Na podstawie pomiarów opracowano modele typu GRID, które stały się podstawą przeprowadzenia analiz zmian kształtu płyty mostu w poszczególnych okresach pomiarowych.

Kolejny problem badawczy dotyczył rejestracji, przetwarzania i analiz przemieszczeń pionowych obiektów inżynierskich. W pracy [13] dokonaliśmy analizy przetwarzania wyników pomiarów przemieszczeń pionowych na przykładzie elewatora zbożowego. W celu przeanalizowania przemieszczeń zarejestrowanych na reperach kontrolnych wykorzystano specjalnie opracowaną aplikację autorską, która pozwoliła na zastosowanie aproksymacji wielomianowej oraz aproksymacji płaszczyzną w oparciu o wybrane punkty pomiarowe zlokalizowane na konstrukcji. W pracy przeanalizowano wyznaczenie odkształceń konstrukcji oraz parametrów pracy obiektu w przestrzennym układzie odniesienia. Przeprowadzone analizy pozwoliły zlokalizować miejsca, w których konstrukcja pracuje najmocniej, czyli rejony gdzie występują największe naprężenia generowane przez obciążenia obiektu. Ponadto przeanalizowano wpływ poszczególnych punktów pomiarowych na deformację punktów sąsiednich. Przeprowadzone badania pozwoliły wyznaczyć odchylenia reperów kontrolnych

od płaszczyzny poziomej, którą teoretycznie powinny tworzyć oraz porównać różnice wysokości pomiędzy poszczególnymi punktami kontrolnymi, a węzłami wygenerowanymi na podstawie płaszczyzny aproksymacyjnej. Wykonane analizy pozwoliły stwierdzić, w której epoce pomiarowej (w jakim okresie) współpłaszczyznowość reperów była najlepsza. Dodatkowo, w oparciu o pomiary dokonane w kilku epokach pomiarowych dokonano analiz zmian konstrukcji w czasie zarówno na punktach pomiarowych jak i interpolacyjnych węzłach struktury GRID. Dzięki strukturze GRID można było zarejestrować tendencję zmian wysokości na wybranych punktach teoretycznych utworzonych przez punkty węzłowe o tej samej lokalizacji w poszczególnych epokach pomiarowych. Umożliwiło to ustalenie rozkładu odkształceń konstrukcji w poszczególnych miejscach fundamentu elewatora, gdzie deformacja obiektu postępowała najszybciej. Ostatecznie w wyniku przeprowadzonych badań zaproponowano rozwiązanie, które pozwoliło na opracowanie prognozy umożliwiającej przewidywanie skutków odkształceń i ich wpływu na stabilność konstrukcji w czasie.

W następnym, rozpatrywanym w pracy [14], zagadnieniu przeanalizowałem dobór wybranych parametrów interpolacji w zależności od lokalizacji punktów pomiarowych. W czasie analiz zwróciłem szczególną uwagę na prawidłowe zdefiniowanie rozdzielczości węzłów w stosunku do zagęszczenia punktów pomiarowych. Ustaliłem, że o dokładności modelu powierzchni nie decyduje wyłącznie zagęszczenie punktów tworzących model. Istotnym czynnikiem umożliwiającym podniesienie jakości DTM jest także odpowiedni dobór lokalizacji punktów pomiarowych biorących udział w wyznaczaniu węzła GRID, pozwalający na wyeliminowanie z procesu obliczeniowego przypadków ekstrapolacji. Podczas analiz stwierdziłem, że odszukanie wymaganej liczby punktów mogących utworzyć figurę opisującą węzeł następuje dzięki odpowiedniemu dopasowaniu promienia poszukiwań do rozdzielczości struktury GRID oraz zagęszczenia punktów pomiarowych. Ustaliłem także, że wybór jednakowej wielkości promienia poszukiwań dla różnego zagęszczenia punktów pomiarowych jest nieekonomiczny ze względu na długi czas przetwarzania zbiorów. Przy dużym zagęszczeniu punktów pomiarowych promień poszukiwań można w sposób kontrolowany zmniejszyć bez straty jakości generowanego modelu, uzyskując dzięki temu większą szybkość przetwarzania zbiorów. W przypadku małego zagęszczenia punktów pomiarowych zbyt mały promień poszukiwań prowadzi do pogorszenia jakości generowanego modelu powierzchni. Jednocześnie zbyt duże zwiększanie wielkości promienia powoduje utratę jakości modelu przez uwidocznienie skokowych zniekształceń jego powierzchni. W czasie badań stwierdziłem, że podczas ustalania promienia poszukiwań należy uwzględniać jednocześnie rozdzielczość struktury GRID oraz zagęszczenie i odpowiednią lokalizację punktów pomiarowych. Zaproponowana metoda doboru tych parametrów pozwala na wygenerowanie dokładnego numerycznego modelu terenu przy zachowaniu krótkiego czasu przetwarzania danych.

5.2. Wybrane publikacje

- [1] Gościewski D. 1998. *Pakiet RIDA v.2.0 – oprogramowanie wspomagające ocenę dokładności pozyskiwania informacji graficznych zasilających mapę numeryczną*. Gaździcki J., Musiał E. (red.), Systemy informacji przestrzennej: VIII Konferencja Naukowo-

- Techniczna. Polskie Towarzystwo Informacji Przestrzennej, Warszawa, Tom 2, s. 99-109. p-ISBN: 83-87660-24-8.
- [2] Gościewski D. 1998. *Metodyka kontroli skanerów stosowanych do pozyskiwania informacji graficznych w procesie tworzenia map numerycznych*. Gaździcki J., Musiał E. (red.), Systemy informacji przestrzennej: VIII Konferencja Naukowo-Techniczna, Warszawa. Polskie Towarzystwo Informacji Przestrzennej, Tom 2, s. 89-98. p-ISBN: 83-87660-24-8.
- [3] Gościewski D. 1999. *Graficzna analiza błędów występujących w procesie przetwarzania graficzno-numerycznego i numeryczno-graficznego. Graphic analysis of errors occurring in graphic-numerical and numerical-graphic processing*. Biuletyn Naukowy ART, Olsztyn, (6) s. 235-248. p-ISSN: 1505-4705.
- [4] Gościewski D. 2000. *Metody oceny i minimalizacji składowych zniekształceń obrazów rastrowych w systemach przetwarzania informacji przestrzennych. Estimation and reduction methods of raster distortions in spatial data processing systems*. Prace Naukowe Instytutu Górnictwa Politechniki Wrocławskiej, Wrocław, 90 (27), s. 143-154. p-ISSN: 0324-9670.
- [5] Gościewski D. 2002. *Analiza i prognoza wielkości zniekształceń obrazów rastrowych z wykorzystaniem szeregów czasowych. Analysing and predicting the magnitude of distortions of raster images using time series*. Zeszyty Naukowe Akademii Rolniczej we Wrocławiu, Geodezja i Urządzenia Rolne, Wrocław, 20 (452), s. 117-130. p-ISSN: 0867-7964.
- [6] Bojarowski K., Gościewski D., Wolak B. 2002. *Technologia przetwarzania wyników pomiaru ukształtowania dna rejestrowanych przez urządzenia o działaniu ciągłym. Technology of processing measurement results related to the sea bottom formation continually registered by recording device*. Zeszyty Naukowe Wyższej Szkoły Morskiej w Szczecinie, Szczecin, (65) s. 53-64. p-ISSN: 0209-2069.
- [7] Gościewski D. 2005. *Influence of measuring point location on selection of interpolation algorithms*. Cygas D., Froehner K.D. (edited), The 6th International Conference on Environmental Engineering. Vilnius Gediminas Technical University Press Technika, Vilnius, Lithuania, Vol. 2, s. 867-871. p-ISBN: 9-986-05850-3.
- [8] Gościewski D., Bojarowski K. 2008. *Generowanie numerycznego modelu terenu na podstawie wyników pomiaru skanerem laserowym. Digital terrain model generation based on laser scanner measurement results*. Czasopismo Techniczne, Środowisko, Kraków, (2-Ś), s. 37-45. p-ISSN: 0011-4561.
- [9] Gościewski D., Bojarowski K. 2012. *Ocena elementów konstrukcyjnych obiektów budowlanych z wykorzystaniem numerycznego modelowania powierzchni. Constructional elements evaluation of boulding objects using digital model of surface*. Zeszyty Naukowe Politechniki Rzeszowskiej, Budownictwo i Inżynieria Środowiska, Rzeszów, 59 (283), s. 51-60. p-ISSN: 0209-2646.
- [10] Krypiak-Gregorczyk A., Wielgosz P., Gościewski D., Paziewski J. 2013. *Validation of interpolation techniques for ionospheric regional total electron content*

- mapping*. Acta Geodynamica et Geomaterialia, 10 (3), s. 275-283. p-ISSN: 1214-9705, DOI: 10.13168/AGG.2013.0027.
- [11] Gościewski D. 2013. *The effect of the distribution of measurement points around the node on the accuracy of interpolation of the digital terrain model*. Springer, Journal of Geographical Systems, 15 (4), s. 513-535. p-ISSN: 1435-5930. DOI: 10.1007/s10109-012-0176-x.
- [12] Bojarowski K., Gościewski D. 2013. *Numerical maps in the designing and registration of engineering objects*. Žróbek R., Kereković D. (ed.), GIS and its implementations. The modern Geodesy, Cadastre and Cartography. Croatian Information Technology Society, GIS Forum, Zagreb, Croatia, s. 40-45. p-ISBN: 978-953-6129-35-5.
- [13] Gościewski D., Bojarowski K. 2013. *Recording, processing, analysis and visualization of vertical displacements*. Žróbek R., Kereković D. (ed.), GIS and its implementations. The new GIS solutions. Croatian Information Technology Society, GIS Forum, Zagreb, Croatia, s. 192-202. p-ISBN: 978-953-6129-35-5
- [14] Gościewski D. 2013. *Selection of interpolation parameters depending on the location of measurement points*. Taylor & Francis, GIScience & Remote Sensing, 50 (5), s. 515-526. p-ISSN: 1548-1603. DOI: 10.1080/15481603.2013.827369.

6. Perspektywy badawcze

Zaproponowana przeze mnie metodyka przetwarzania danych (rozdz. 4.2) oraz analizy przedstawione w artykułach dotyczących głównych kierunków naukowo-badawczych (rozdz. 5.1) mogą zostać wykorzystane do analizowania różnego rodzaju pomiarów pozyskiwanych w różnych dziedzinach badawczych. Obecnie trwają prace nad zastosowaniem przedstawionej metodyki w analizie i przetwarzaniu danych pozyskanych z rynku nieruchomości. W recenzji znajdują się dwie kolejne publikacje. Pierwsza z nich pt. *An evaluation of changes on urban real estate markets with the use of GRID structures* (współautorzy: Gerus-Gościewska M., Szczepańska A.) omawia wykorzystanie struktury GRID do analiz miejskich rynków nieruchomości. Powierzchnia interpolacyjna typu GRID może zostać wykorzystana do analiz zmian wartości w dowolnym punkcie badanego obszaru oraz do dokonania porównań z powierzchniami wygenerowanymi w innych epokach pomiarowych. W strukturach typu GRID jest to możliwe dzięki ustaleniu takiej samej wielkości pola podstawowego dla wszystkich epok pomiarowych. Pozwala to na porównanie ze sobą dowolnych fragmentów (reprezentowanych przez punkty węzłowe) wygenerowanych powierzchni interpolacyjnych, pomimo różnej liczby i lokalizacji punktów pomiarowych oraz ich różnego lokalnego zagęszczenia. Druga z publikacji pt. *Use of the GRID structure to reconstruct and forecast the value of real estate in selected measurement epochs* (współautorzy: Gerus-Gościewska M., Szczepańska A.) prezentuje rozwiązania pozwalające na wyodrębnienie danych umożliwiających porównanie cen nieruchomości zarejestrowanych w różnych epokach, odtwarzanie przebiegu zmian w czasie. Dysponując danymi pozwalającymi utworzyć szereg czasowy przypisany do każdego węzła struktury GRID można przeanalizować zmiany zachodzące w mierzonej przestrzeni w wybranych punktach obiektu i porównać je ze sobą. Dzięki ujednoliceniu struktur w każdej z epok możliwe jest wyselekcjonowanie zmieniających się wartości w czasie w każdym węźle

(zlokalizowanym w tym samym miejscu w poszczególnych epokach) oraz na dokonanie aproksymacji tych zmian przy wykorzystaniu wielomianu o dowolnie ustalonym stopniu. Przechowywanie informacji o zmianie wartości w poszczególnych węzłach w postaci współczynników wielomianu umożliwia wygenerowanie na ich podstawie powierzchni modelu przypisanej do dowolnie wybranej epoki oraz odtworzenie szukanych wielkości w dowolnym momencie czasowym z założoną dokładnością.

Badania nad wykorzystaniem struktury GRID oraz jej własnościami teoretycznymi i innymi zastosowaniami praktycznymi będą kontynuowane we współpracy z jednostkami naukowymi, co wyrażam przekonanie, przyczyni się do powstania szeregu kolejnych publikacji naukowych.

Dariusz Gościński