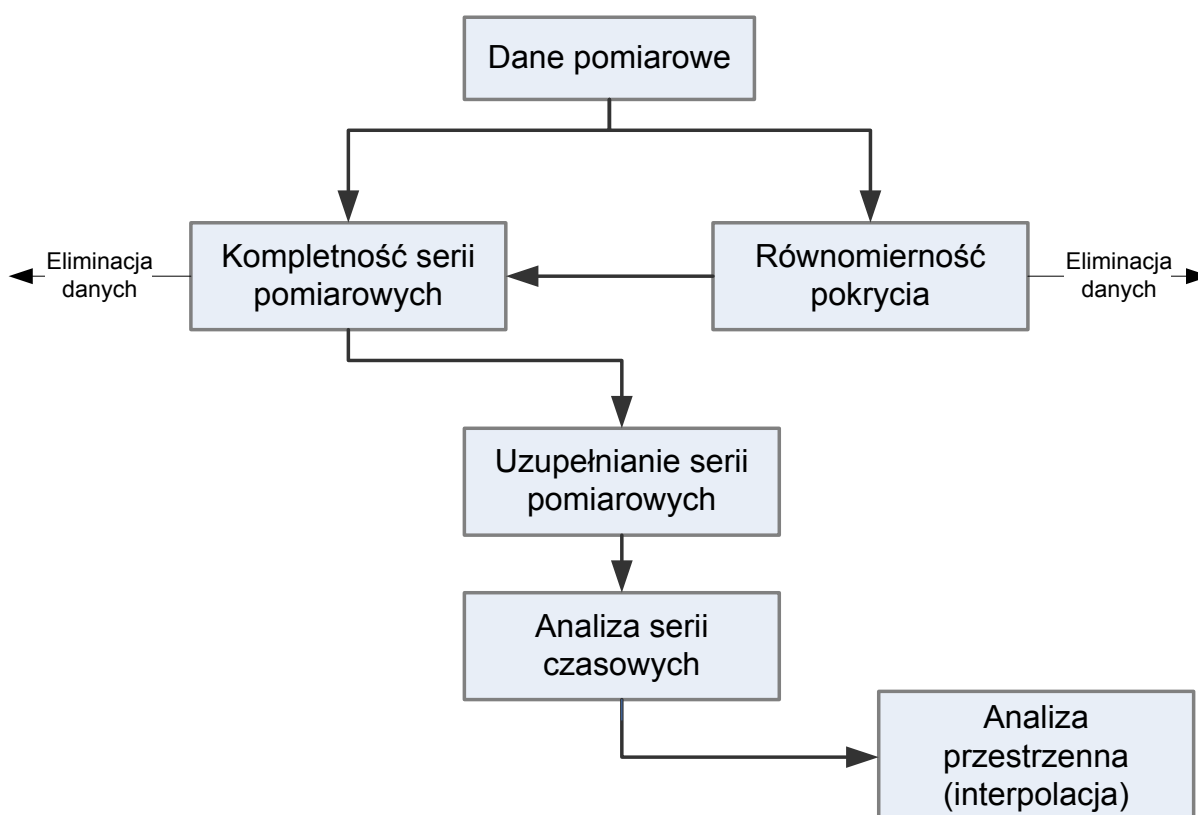

Dodatek: Szczegółowe zagadnienia metodyczne

X.1. Problem zmian ilości i rozkładu przestrzennego danych

Większość opracowań klimatologicznych dotyczących zmienności przestrzennej parametrów klimatu jest wykonywana według ustalonego, tradycyjnego schematu metodycznego: „najpierw agregacja, później interpolacja”. Procedura taka obejmuje zazwyczaj kilka etapów (ryc. 123):



Ryc. 123. Typowy przebieg analizy danych klimatologicznych. Dokładne objaśnienia w tekście.

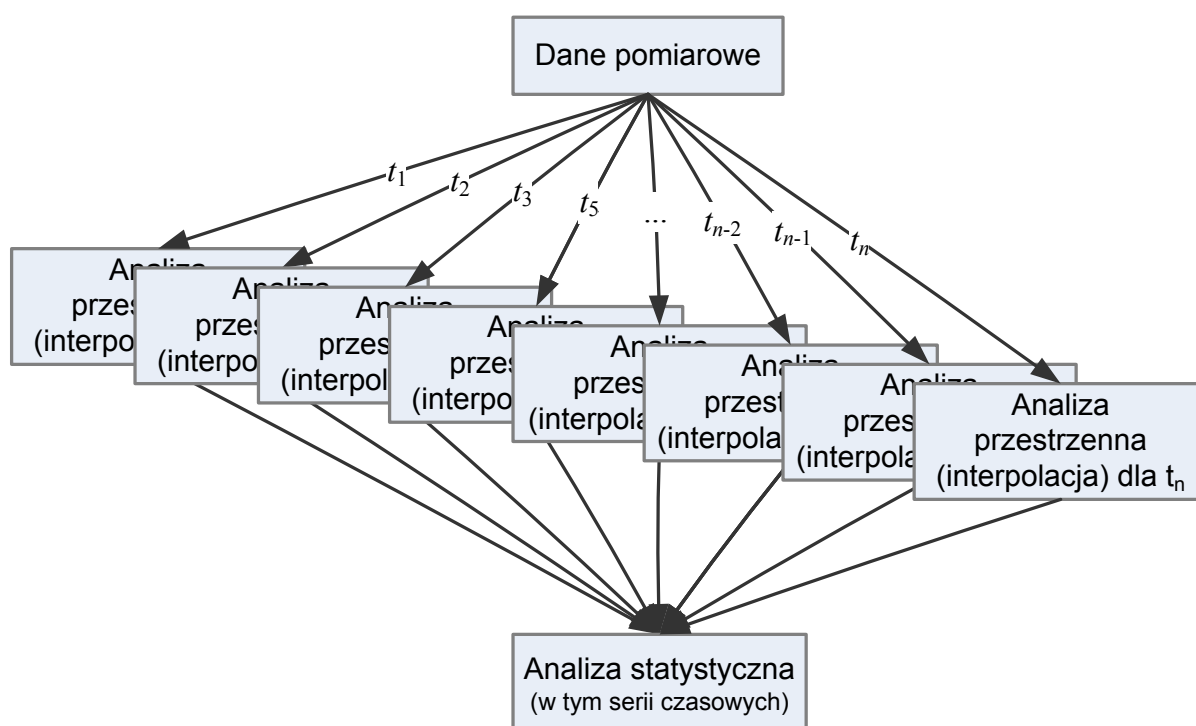
- selekcję punktów pomiarowych ze względu na jednolitość pokrycia analizowanego obszaru. Zazwyczaj, jeśli istnieją skupienia stacji leżących blisko siebie, część z nich eliminuje się w celu uzyskania zagęszczenia typowego dla pozostałego terenu. Zwykle nie testuje się, czy i w jakim stopniu informacje z blisko leżących punktów są nadmiarowe;
- sprawdzanie jednorodności i kompletności ciągów pomiarowych dla poszczególnych stanowisk (punktów). Stacje z danymi budzącymi wątpliwości i/lub z dużymi lukami w ciągach pomiarowych są albo eliminowane z analizowanej bazy, albo z uwagi na „ważność” stanowiska, stosuje się różnorodne techniki statystyczno-matematycznego wyrównywania i uzupełniania danych;
- obliczanie różnorodnych statystyk dla zweryfikowanych punktów pomiarowych;
- interpolację uzyskanych statystyk na cały analizowany obszar.

Opisana wyżej procedura ma wiele zalet, ale także dwie poważne wady. Po pierwsze, w trakcie przygotowania materiału, eliminowana jest część potencjalnie bardzo użytecznych danych z punktów o niekompletnych ciągach pomiarowych i/lub z obszarów o dużym zagęszczeniu stacji. Po drugie agregacja danych dla punktów i interpolacja wyłącznie parametrów statystycznych, jak na przykład średnie okresowe, powoduje, że nieznana pozostaje zmienność czasowa struktury przestrzennej analizowanego parametru. W wypadku wielu parametrów klimatycznych występują powtarzalne układy przestrzenne, związane na przykład z warunkami cyrkulacyjnymi, które zazwyczaj pozostają zamazane lub kompletnie nieczytelne, kiedy analiza ogranicza się do statystyk obliczanych dla miesięcy, sezonów, lat czy wieloleci.

W niniejszym opracowaniu przyjęto odmienną strategię analizy danych: „najpierw interpolacja, później agregacja” (ryc. 124). Jej podstawą jest wykorzystanie wszystkich dostępnych punktowych danych w elementarnym kroku czasowym dla uzyskania obrazów zmienności przestrzennej analizowanego zjawiska. Zagregowane charakterystyki pola zostają określone w drugim etapie, na podstawie nie oryginalnych punktów danych, ale albo metodami GIS-u z rastrowych obrazów etapowych, albo jako statystyki nieprzestrzenne.

Aby móc zastosować takie podejście, najpierw należy jednak mieć pewność, że zmiany rozmiarów i struktury sieci pomiarowej zachodzące pomiędzy kolejnymi terminami pomiarów nie mają istotnego wpływu na jakość uzyskiwanych wyników. Przeprowadzone

przez Stacha i Tamulewicz analizy przestrzenne sum miesięcznych opadów dla tego samego wielolecia 1956-1980 (2005a, 2005b, 2005c) wykazały, że zachodzące wówczas względnie niewielkie zmiany ilości i rozmieszczenia danych pomiarowych nie zmieniały w wyraźny sposób dokładności uzyskanego rozkładu przestrzennego zjawiska. Dokładność ta uzależniona była bowiem przede wszystkim od specyfiki samego zjawiska w postaci lokalnej zmienności pola opadów, wyrażonej poprzez statystyki najbliższego sąsiada (por. rozdz. V). Zróżnicowanie globalne sum opadów, charakteryzowane na przykład przez odchylenie standardowe, miało marginalne znaczenie.



Ryc. 124. Przebieg analizy danych klimatologicznych uwzględniającej wszystkie dostępne wyniki pomiarów oraz najpierw ich przestrzenną, a później czasową syntezę. Dokładne objaśnienia w tekście.

W cytowanych powyżej pracach wykonywano jednak ocenę jedynie względem sum miesięcznych. Dane MSDO będące podzbiorem kompletnych ciągów sum dobowych opadów mają wyraźnie odmienny charakter. Nie mając możliwości przeprowadzenia odpowiednich testów na zbiorach sum dobowych opadów, uznano za wystarczające uzasadnienie wyżej postawionej tezy dobrze udokumentowany przykład z literatury. Są nim wyniki analiz przeprowadzonych w trakcie eksperymentu zwanego *Spatial Interpolation Comparison 97* (SIC-97, Dubois i in. 2003). Celem tego przedsięwzięcia było przetestowanie na jednym zbiorze danych opadowych estymacji przestrzennych wykonanych za pomocą różnych metod

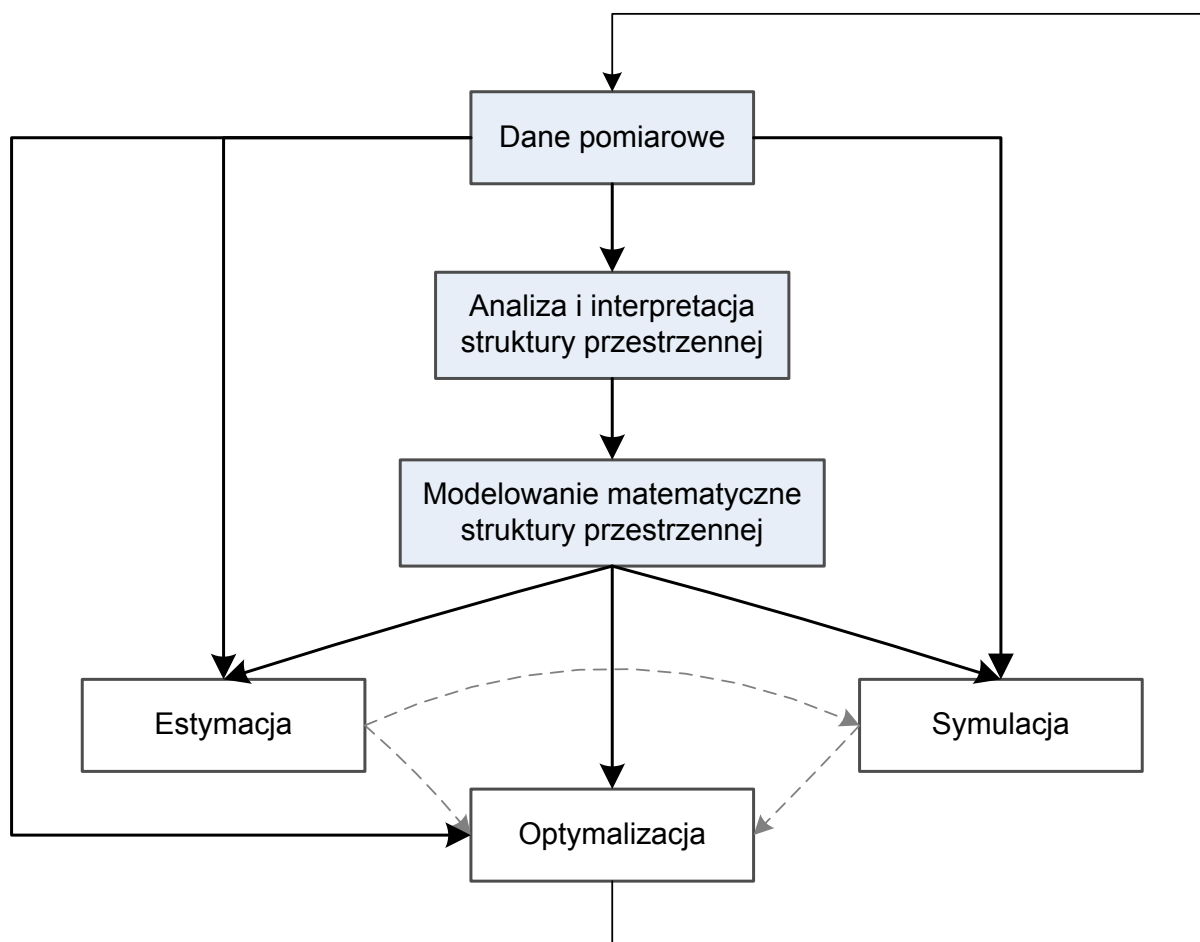
interpolacyjnych pod kątem ich dokładności i przydatności do rutynowych ocen zagrożeń skażeniami (szybkość obliczeń, stopień komplikacji procedury, możliwość automatyzacji itp.). Jako przykładowe dane posłużyły sumy dobowe opadów z 8 maja 1986 roku kiedy nad Szwajcarią przemieszczała się chmura radioaktywnych aerozoli pochodząca z uszkodzonego reaktora elektrowni atomowej w Czarnobylu na Ukrainie.

W ramach SIC-97 Dubois i Shibli (2003) przeprowadzili także ocenę wpływu ilości danych na jakość uzyskiwanych wyników, wyrażoną poprzez pierwiastek średniego błędu kwadratowego (RMSE). Eksperyment polegał na losowaniu z całego zbioru 467 danych podzbiorów od 25 do 300 punktów w kroku co 25. Na podstawie tych danych dokonywano estymacji metodami IDW – odwrotnej wagi odległości (ang. *inverse distance weighting*), MQ – funkcji multikwadratowych (ang. *multiquadratic function*), OK – zwykłego krigingu (ang. *ordinary kriging*), optymalizując parametry obliczeń za pomocą krosvalidacji (Davis 1987, Deutsch, Journel 1998, Goovaerts 1997, Wackernagel 2003). W celu uniknięcia przypadkowych wyników, losowanie i obliczenia dla każdej wielkości podzbioru powtarzano 10 razy. Dla każdej z trzech metod tendencje RMSE były identyczne. Po przekroczeniu 100, kolejne przyrosty ilości wykorzystywanych danych nie miały już istotnego wpływu na wielkość błędu estymacji. Ponieważ optymalizacja parametrów obliczeń metodą krosvalidacji ma wadę polegającą na tym, że te same dane służą do wykonania estymacji i oceny jej jakości, przeprowadzono dodatkowy test. Użyto do niego pozostałe z losowania dane to jest $467 - N$, gdzie N oznacza wielkość wylosowanej próby = 25, 50, 100, ..., 300. Tym razem wartość błędu stabilizowała się po przekroczeniu 275 danych.

Żeby ocenić znaczenie przedstawionych powyżej wyników eksperymentu numerycznego wykonanego w ramach SIC-97, trzeba brać pod uwagę kilka czynników. Po pierwsze. eksperyment był wykonany na jednym, konkretnym zbiorze sum dobowych opadów. Relacje te mogą się zmieniać prawdopodobnie w znacznym zakresie w zależności od genezy i charakteru opadów. Po drugie, Szwajcaria jest znacznie mniejsza od Polski (tylko 41,3 tys. km²), ale w całości górzysto-wyżynna, co bardzo wpływa na zmienność przestrzenną opadów. Porównując zagęszczenie sieci pomiarowej obu krajów i przyjmując 300 stanowisk w Szwajcarii jako poziom zapewniający stabilne wyniki estymacji, można oceniać, że dla Polski oznaczałoby to około 1200-1500 punktów. Jest to, zdaniem autora, szacunek bardzo ostrożny, a rzeczywista wartość – niższa.

X.2. Estymacja, symulacja i optymalizacja

Opis i modelowanie (por. rozdz. III.2) struktury przestrzennej analizowanej cechy rzadko stanowią główny cel prac wykorzystujących metody geostatystyki. Jest to zazwyczaj tylko etap, w którym zidentyfikowane i matematycznie określone reguły zmienności przestrzennej są następnie wykorzystywane do estymacji, symulacji lub optymalizacji (ryc. 125). Takie motywy stanowiły również genezę niniejszej rozprawy, chociaż w dalszym etapie, o czym wspomniano we wprowadzeniu (por. rozdz. I), jej zakres został ograniczony. Poniżej zostaną w skrócie omówione te elementy wybranej metodyki estymacji i symulacji pola prawdopodobieństwa MSDO⁵⁴, które miały największy wpływ na sposób rozwiązania problemu stanowiącego istotę niniejszej rozprawy – analizy i modelowania struktury przestrzennej maksymalnych sum dobowych opadów na terenie Polski.



Ryc. 125. Schemat przebiegu pełnej analizy geostatystycznej. Elementy ujęte w niniejszej rozprawie zaznaczono szarym tłem. Szersze objaśnienia w tekście.

⁵⁴ Wyniki tych prac zostaną opublikowane w osobnej rozprawie.

Sednem geostatystyki jest problem estymacji wartości cechy z w nieopróbowanej lokalizacji \mathbf{u} . Dysponujemy w tym momencie zbiorem n pomiarów cechy z dokonanych w szeregu punktów o znanych współrzędnych \mathbf{u}_α , czyli $z(\mathbf{u}_\alpha)$. Wykazują one wszystkie pewien stopień autokorelacji przestrzennej. W większości metod interpolacyjnych ową nieznaną, szacowaną wartość oblicza się jako liniową kombinację z $n(\mathbf{u}_\alpha)$ sąsiednich wyników pomiarów [30]:

$$\hat{z}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u}_\alpha)} \lambda_\alpha z(\mathbf{u}_\alpha) \quad [30]$$

W geostatystyce nieznaną wartość $z(\mathbf{u})$ jest interpretowana jako realizacja funkcji losowej $FL(\mathbf{u})$, która jest w pełni charakteryzowana przez jej rozkład prawdopodobieństwa $F(\mathbf{u};z) = \text{Prob}\{Z(\mathbf{u}) \leq z\}$. Wagi λ_α uwzględniają odległość lokalizacji danych od miejsca estymacji, ich przestrzenną konfigurację (w tym ich potencjalne skupienia), a przede wszystkim ich przestrzenną strukturę określoną poprzez model uzyskany na podstawie autokowariancji empirycznej [2] (por. rozdz. III.2.2). Uzyskuje się je z rozwiązania układu $(n(\mathbf{u}_\alpha) + 1)$ równań liniowych, zwanych równaniami zwykłego krigingu (ang. *Ordinary Kriging* = **OK**, [31]):

$$\sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta \gamma(\mathbf{u}_\alpha - \mathbf{u}_\beta) - \mu = \gamma(\mathbf{u}_\alpha - \mathbf{u}) \quad \alpha = 1, \dots, n(\mathbf{u})$$

$$\sum_{\beta=1}^{n(\mathbf{u})} \lambda_\beta = 1 \quad [31]$$

gdzie μ to mnożnik Lagrange'a konieczny ze względu na ograniczenie wag (sumowanie się wag do jedności).

Kriging daje w większości wypadków znacznie lepsze efekty niż inne metody estymacji (= interpolacji, Dubois i in. 2003, Dubois, Shibli 2003, Lam 1983, Weber, Englund 1992, 1994, Zimmerman i in. 1999), ponieważ jego algorytm zawiera procedury minimalizacji błędów i zapewnienia ich losowego rozkładu (estymator jest nieobciążony, co oznacza bilansowanie się błędów ujemnych i dodatnich).

W wielu zastosowaniach, zarówno teoretycznych, jak i zwłaszcza praktycznych, równie ważne jak ocena nieznannej, lokalnej (punktowej), wartości analizowanej cechy jest określenie niepewności tej oceny. Poza tym, zwykły kriging, jak wiele innych metod interpolacyjnych, daje wyniki będący optymalną estymacją z punktu widzenia kryterium najmniejszych kwadratów. Często jednakże, dla konkretnych potrzeb, niezbędny jest szacunek nieznannej

wartości dokonany względem innych kryteriów. Dlatego, w obrębie geostatystyki opracowano kilka metod umożliwiających, zamiast określenia jedynie jednej „optymalnej” wartości analizowanej cechy w nieopróbowanej lokalizacji, budowę kompletnych lokalnych modeli jej rozkładu (dystrybuanty). W odniesieniu do danych MSDO takie podejście ma dodatkowy, metodologiczny aspekt. Ze względu na nieciągłość przestrzenną epizodów opadowych i asynchroniczność MSDO, celem estymacji nie może być określenie konkretnej wartości, ale raczej prawdopodobieństwa jej wystąpienia. Efektem końcowym nie jest wówczas typowa mapa rastrowa czy izarytmiczna, ale model pola prawdopodobieństwa, z którego możemy zależnie od potrzeb uzyskać dowolne przekroje (mapy dla określonych progów prawdopodobieństwa) lub bardziej złożone charakterystyki. Do wspomnianych wyżej metod geostatystycznych umożliwiających budowę kompletnych lokalnych modeli rozkładu estymowanej cechy należą kriging multigaussowski (ang. *multigaussian kriging*), kriging dysjunktywny (ang. *disjunctive kriging*) oraz kriging wartości kodowanych (ang. *indicator kriging – IK*), zwany czasem w polskiej literaturze tematu krigingiem wskaźnikowym bądź indyktorowym (Zawadzki 2005). Są one opisane i porównane w kilku najważniejszych podręcznikach teorii geostatystyki (Chilès, Delfiner 1999, Goovaerts 1997, Webster, Oliver 2001), a także w szeregu prac eksperymentalnych (Carr, Deng 1987, Goovaerts 2001, Lajaunie 1990, Lark, Ferguson 2004, Moyeed, Papritz 1999, Papritz and Dubois 1999, Solow 1986). W ramach projektu badawczego, którego częściowym efektem jest niniejsza rozprawa, zdecydowano się zastosować kriging danych kodowanych (IK). Decyzję tę podjęto analizując bilans wad i zalet tej metodyki.

Wady krigingu wartości kodowanych obejmują (Christakos 2000, Deutsch, Journel 1998, Goovaerts 1997, Olea 1999):

- utratę części informacji ze względu na dyskredytyzację ciągłej dystrybuanty empirycznej;
- pracochłonność – konieczność czasochłonnego budowania modelu semiwariancji dla każdej wartości progowej; często występujące trudności w określeniu modeli dla wartości bardzo niskich i bardzo wysokich zmuszają do subiektywnych decyzji, a te rodzą wątpliwości co do optymalności uzyskanych estymacji;
- wykraczanie estymowanych prawdopodobieństw poza dopuszczalny zakres (0, 1) oraz błędy w ich relacjach porządkowych;
- arbitralnie przyjmowaną metodę interpolacji/ekstrapolacji uzyskanej warunkowej dystrybuanty.

Zaletami krigingu wartości kodowanych są:

- potwierdzona w dziesiątkach zastosowań i testów metodycznych skuteczność;
- brak trudnych do weryfikacji założeń dotyczących rozkładu statystycznego populacji (metoda nieparametryczna);
- żadna z alternatywnych metod nie jest wyraźnie lepsza;
- łatwość ich używania; alternatywne metody są bardziej skomplikowane – bardziej „podatne” na błędy metodyczne,
- łatwa możliwość uwzględnienia danych uzupełniających („twardych” i „miękkich”);
- powszechna dostępność oprogramowania (Deutsch, Journel 1998, Mao, Journel 1998, Pardo-Igúzquiza, Dowd 2005, Richmond 2002); na bazie publicznie dostępnego kodu komputerowego tradycyjnych metod estymacji geostatystycznej (SK, OK) łatwo można samemu napisać program komputerowy realizujący obliczenia IK.

Istotą metody IK jest najpierw budowa globalnego modelu niepewności opartego na dystrybucie empirycznej wyliczonej ze wszystkich danych pomiarowych, a następnie jego lokalna modyfikacja na podstawie wektora danych kodowanych z sąsiadujących z lokalizacją punktu estymacji stanowisk pomiarowych.

Na podstawie rozkładów prawdopodobieństwa uzyskanych metodą krigingu wartości kodowanych obliczane są także dwie charakterystyki ogólne: średnia wartość szacowanej dystrybuanty, zwana średnią oczekiwaną (ang. *E-type mean* [32], ryc. 128) oraz wariancja warunkowa:

$$\begin{aligned} [z(\mathbf{u})]_E^* &= \int_{-\infty}^{+\infty} z dF(\mathbf{u}; z|(n)) \\ &\approx \sum_{k=1}^{K+1} z_k' [F(\mathbf{u}; z_k|(n)) - F(\mathbf{u}; z_{k-1}|(n))] \end{aligned} \quad [32]$$

gdzie z_k , $k = 1, \dots, K$ to K przyjętych wartości progowych, a $z_0 = z_{min}$, $z_{K+1} = z_{max}$ to minimum i maksimum zakresu cechy z , które są parametrami podawanymi przez osobę wykonującą obliczenia (operatora). Średnia warunkowa każdej klasy $(z_{k-1}, z_k]$ określona jako z_k' jest uzyskiwana na podstawie procedury interpolacji, której rodzaj i parametry są również specyfikowane przez operatora.

Wszystkie metody estymacji oparte na regresji liniowej liczonej metodą najmniejszych kwadratów mają jednak istotną wadę (tab. 28). Jest nią zmienne lokalnie „wygładzanie” rozkładu statystycznego estymowanej cechy. W kategoriach ilościowych wada metod interpolacyjnych polega na tym, że wygenerowany za ich pomocą model (powierzchnia opisana izoliniami), ma inną charakterystykę statystyczną (rozkład zmiennej/histogram) i przestrzenną (semiwariogram) niż oryginalna próbka, na podstawie której powstał. Aby ten problem choć częściowo pokonać, na gruncie geostatystyki opracowano kilka metod symulacji przestrzennej, które honorując dane pomiarowe, umożliwiają generowanie szeregu równie prawdopodobnych obrazów (modeli), w pełni zachowujących charakterystyki rozkładu statystycznego próbki i strukturę przestrzenną zjawiska (Deutsch, Journel 1998, Goovaerts 1997, tab. 28). Daje to duże możliwości ilościowej oceny niepewności dokonanych szacunków. Geostatystyczne metody estymacji dają optymalne oceny (wraz z szacunkiem ich niepewności) wartości analizowanego parametru w punkcie, metody symulacji dla obszarów (zlewni rzecznych, regionów fizycznogeograficznych, jednostek administracyjnych itp.). Istotą symulacji geostatystycznych jest zastosowanie metodyki Monte Carlo (próbkiowania rozkładu statystycznego), ograniczonej poprzez uwzględnienie danych pomiarowych i modelu struktury przestrzennej zjawiska. Z tego względu w nazwach poszczególnych procedur stosowany jest określnik „symulacja warunkowa” (ang. *conditional simulation*).

Tabela. 28. Cechy i zastosowania estymacji oraz symulacji geostatystycznych.

| | Estymacja (kriging) | Symulacja warunkowa |
|-----------------------|---|--|
| Efekt obliczeń | jeden model „deterministyczny” | wiele realizacji (modeli „probabilistycznych”) |
| Właściwości | honoruje dane, minimalizuje wariancję błędu | honoruje dane, histogram, wariogram, gęstość spektralną i in. |
| Obraz | łagodny (gładki), zwłaszcza wobec „chaotycznego” modelu wariogramu | bardzo zróżnicowany, zwłaszcza wobec „chaotycznego” modelu wariogramu |
| Dane wynikowe | tendencja do tworzenia powierzchni trendu z dala od danych; możliwość lokalizacji punktów pomiarowych | taka sama zmienność obrazu w każdej części; niemożność odgadnięcia lokalizacji punktów pomiarowych |
| Zastosowania | tworzenie map izarytmicznych; estymacja punktowa | modelowanie niejednorodności; szacowanie niepewności; estymacja obszarowa |

Różnice między geostatystycznym estymacjami a symulacjami w kontekście analizowanych danych przedstawiono na przykładzie rocznych MSDO z roku 1976 (tab. 29, ryc. 126-128). Tabela 29 i rycina 126 pokazują relacje między statystykami opisowymi i

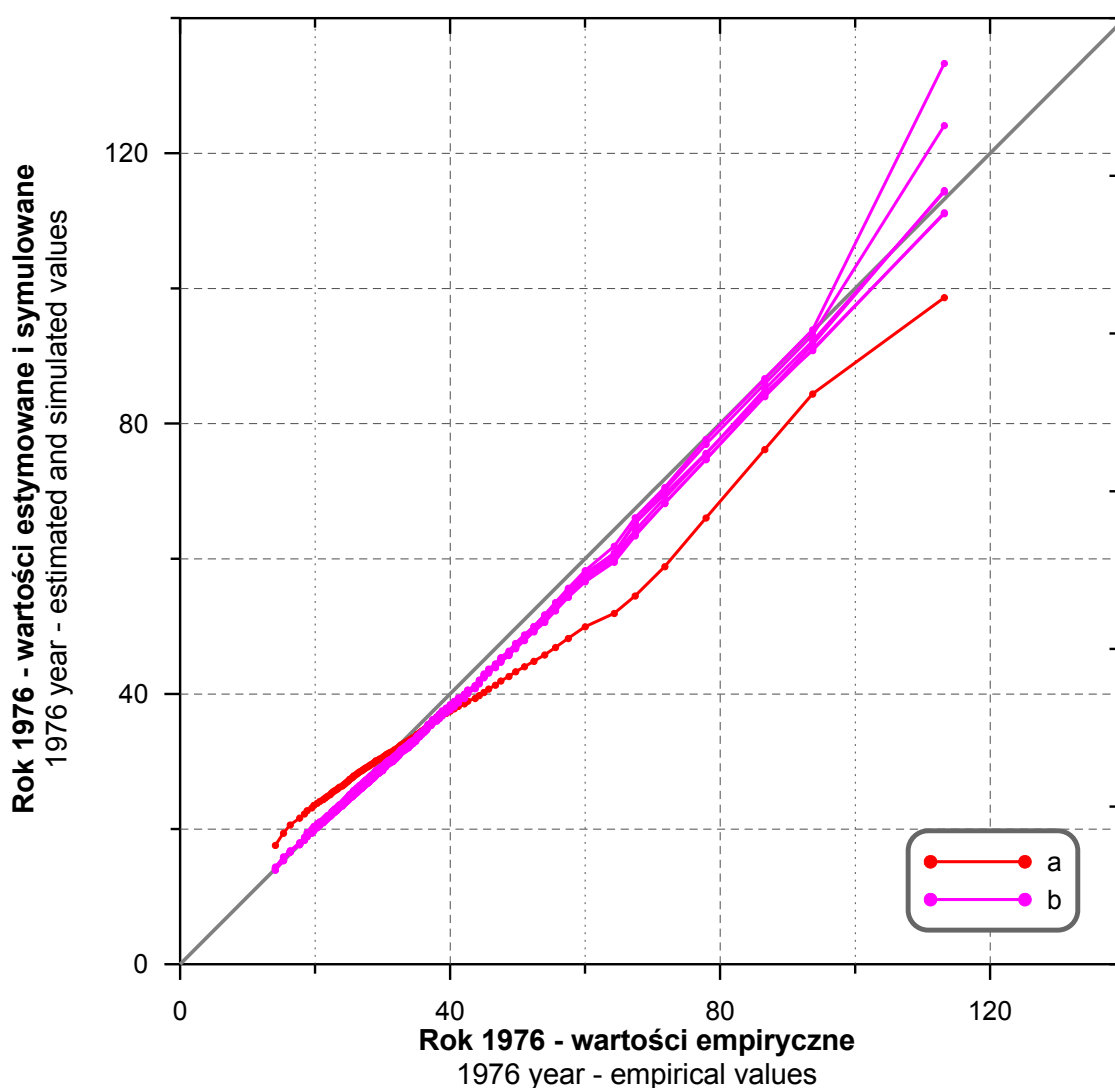
rozkładem danych empirycznych a danych modeli estymowanych i symulowanych. Zbiór estymacji ma znacząco mniejszy zakres niż oryginalne dane pomiarowe (wyższe minimum i niższe maksimum) i mniejszą zmienność reprezentowaną przez odchylenie standardowe. Porównanie kwantyli rozkładu (ryc. 126) wskazuje, że estymacje generalnie zawyżają wartości mniejsze od średniej i zaniżają od niej większe. Brak dobrego odwzorowania charakterystyk zmienności cechy jest szczególnie dokuczliwy, jeśli, tak jak to jest w tym przypadku, analizuje się zmienność przestrzenną wartości ekstremalnych, a nie przeciętnych. Rozkład wartości symulowanych jest praktycznie tych wad pozbawiony. Nie odwzorowują one jednak całkowicie wiernie odpowiednich charakterystyk statystycznych danych pomiarowych, lecz jest to cecha raczej pożądana⁵⁵. Dane pomiarowe to przecież niewielka próba z całej populacji⁵⁶, i są także obciążone błędami⁵⁷, a więc wyliczone z nich statystyki stanowią jedynie przybliżenie. Zbyt ściśle ich honorowanie mogłoby prowadzić do fałszywego wyobrażenia o pewności i podobieństwie symulowanych wartości. Równie ważne są różnice modeli estymowanych i symulowanych w ujęciu przestrzennym (ryc. 127 i 128). Semiwariorogram uzyskany z próby danych estymowanych różni się zasadniczo od wyjściowego semiwariorogramu empirycznego – brakuje nuggetu, początkowy odcinek ma charakter paraboliczny, semiwariancje wyliczone dla każdego odstępów są o połowę mniejsze. Estymowany obraz jest zatem sztucznie wygładzony, tworząc mylne wrażenie ciągłości przestrzennej analizowanej cechy. Daje to co prawda bardziej klarowny, zgeneralizowany obraz zróżnicowania geograficznego, ale należy pamiętać, że owa generalizacja nie jest jednolita. Zależy ona bowiem od lokalnej konfiguracji danych: wygładzanie jest minimalne w bezpośrednim sąsiedztwie lokalizacji pomiarów, a zwiększa się wraz ze wzrostem odległości punktu estymacji od stanowisk pomiarowych. Estymacje krigingowe są zatem bardziej zróżnicowane na gęściej opróbowanych obszarach, mogą tworzyć pozorne struktury przestrzenne. Wygładzone mapy interpolowane nie powinny być zatem używane w tych zastosowaniach, gdzie końcowy wynik jest czuły na obecność wartości ekstremalnych i na charakter ich ciągłości przestrzennej. Obrazy symulowane są tych wad w znacznym stopniu pozbawione (ryc. 128). Semiwariorogramy danych symulowanych oscylują wokół semiwariorogramu wyliczonego z pomiarów, a początkowa nieciągłość (nugget) jest wiernie odwzorowana (ryc. 127). Każdy obraz symulowany jest unikatowy, ale honorowanie danych

⁵⁵ Odstępstwa dystrybuanty i semiwariorogramu danych symulowanych od wartości wyliczonych z oryginalnych pomiarów nazywane są fluktuacjami ergodycznymi (ang. *ergodic fluctuations*, Deutsch, Journal 1998, Goovaerts 1997)

⁵⁶ Patrz przypis 16 (s. 40).

⁵⁷ Patrz podrozdział V.2 (s. 65).

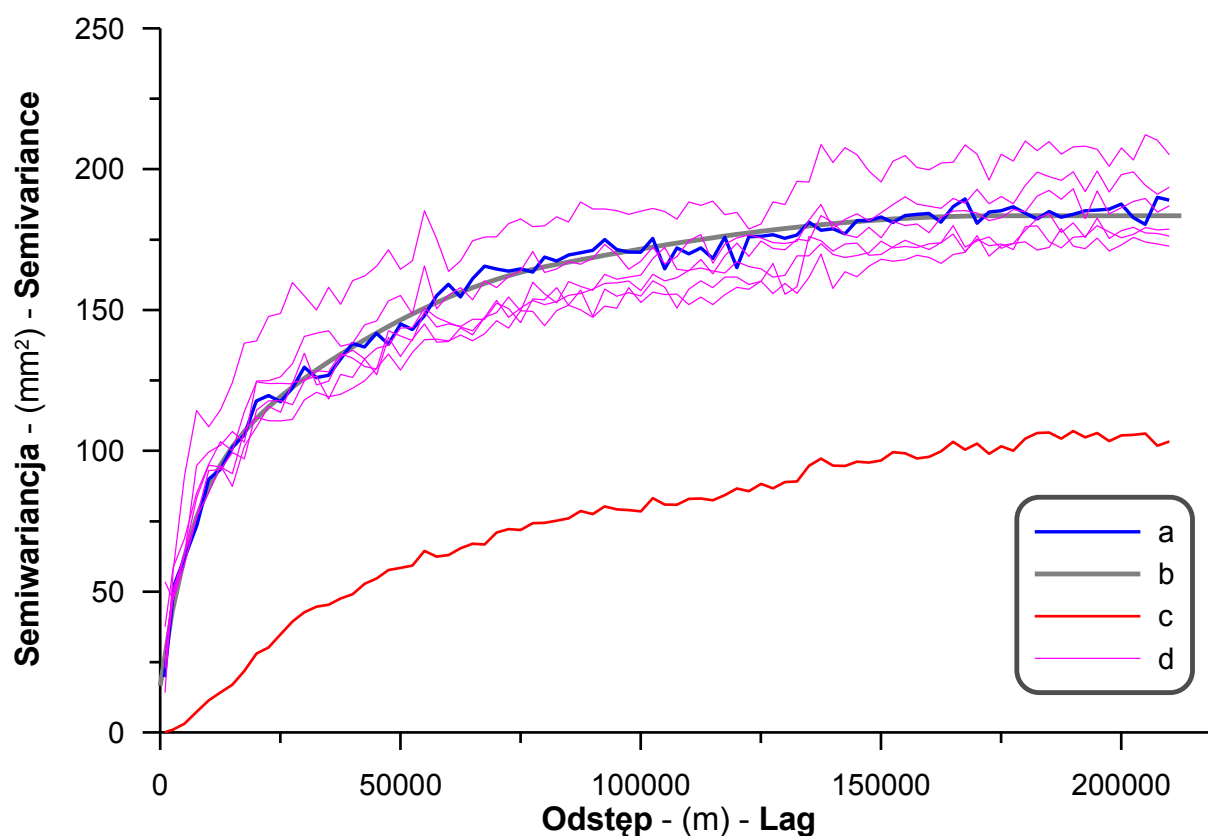
pomiarowych powoduje, że główne cechy zmienności przestrzennej są zachowane. Na mapach widoczna jest „ziarnistość” wynikająca z bezpośredniego sąsiedztwa węzłów siatki o silnie zróżnicowanych wartościach – co może być interpretowane jako nieciągłości opadów wynikające z ich ograniczonego zasięgu przestrzennego oraz asynchroniczności danych MSDO. To duża zaleta, ale należy pamiętać, że owa właściwość symulowanych map MSDO wynika z użycia jednego, globalnego modelu struktury przestrzennej, który przy tak dużym obszarze i tak genetycznie zmiennym zjawisku nie zawsze jest adekwatny lokalnie. W miejscach występowania i dominacji szczególnie rozległych opadów frontalnych o łagodnej zmienności przestrzennej, symulowany obraz może być sztucznie zróżnicowany. Rozwiązaniem tego problemu mogłoby być zastosowanie lokalnych modeli, co w przypadku estymacji rzeczywiście ma już miejsce (Haas 1990, Minasny i in. 2005, Whelan i in. 2001).



Ryc. 126. Porównanie kwantyli (0,001, 0,005, 0,01, 0,02, 0,03, ..., 0,97, 0,98, 0,99, 0,995, 0,999) rozkładu danych empirycznych rocznych MSDO z roku 1976 z (a) wynikami jego estymacji (średnia oczekiwana IK: *E-type mean*) i (b) symulacji (trzy pierwsze realizacje) w siatce 1 na 1 km.

Tabela. 29. Porównanie statystyk oryginalnego zbioru danych pomiarowych rocznych MSDO z roku 1976 oraz wyników jego estymacji (średnia oczekiwana IK: *E-type mean*) i symulacji (trzy pierwsze realizacje) w siatce 1 na 1 km.

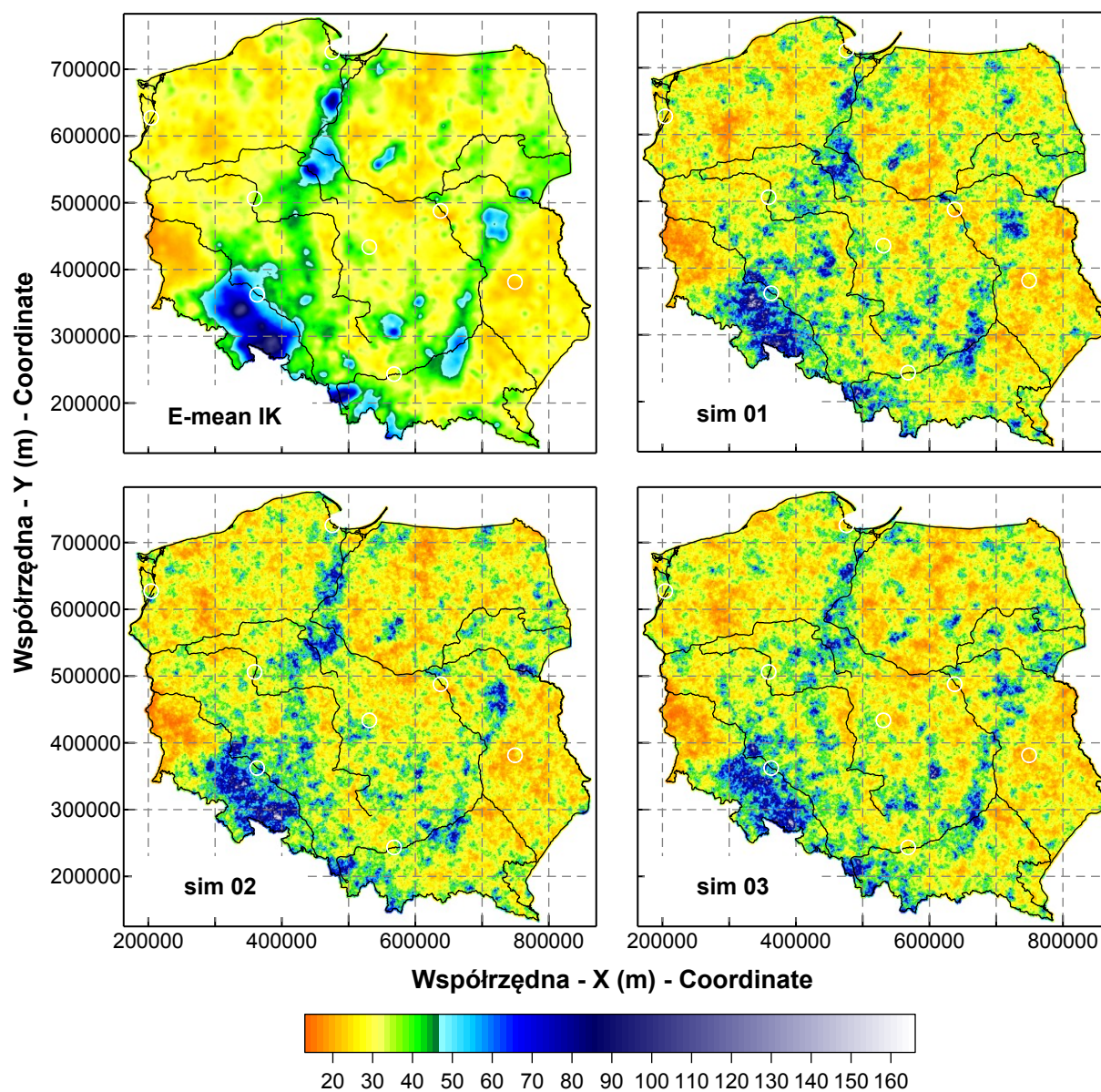
| Parametr | Parameter | 1976 dane/data | 1976 est. E-mean | 1976 sim. 01 | 1976 sim. 02 | 1976 sim. 03 |
|--------------|--------------------|----------------|------------------|--------------|--------------|--------------|
| Liczebność | Count | 2678 | 319114 | 319114 | 319114 | 319114 |
| Minimum | Minimum | 13 | 16,5 | 13,0 | 13,0 | 13,0 |
| 1 kwartyl | First quartile | 25,1 | 27,4 | 24,6 | 24,6 | 24,6 |
| Mediana | Median | 30,8 | 31,2 | 29,6 | 29,6 | 29,7 |
| 3 kwartyl | Third quartile | 39,4 | 37,1 | 37,4 | 37,2 | 37,6 |
| Maksimum | Maximum | 166,1 | 109,0 | 166,1 | 166,1 | 166,1 |
| Średnia | Mean | 34,61 | 33,71 | 33,44 | 33,31 | 33,40 |
| Odch. stand. | Standard deviation | 14,389 | 10,199 | 14,125 | 13,917 | 13,720 |
| Skośność | Skewness | 2,040 | 2,363 | 2,406 | 2,381 | 2,190 |
| Kurtoza | Kurtosis | 9,918 | 11,592 | 12,539 | 12,416 | 10,880 |



Ryc. 127. Porównanie semiwariogramu empirycznego (a) i jego modelu (b) danych pomiarowych rocznych MSDO z roku 1976 z semiwariogramami empirycznymi (c) estymacji (średnia oczekiwana IK: *E-type mean*) i (d) symulacji (trzy pierwsze realizacje). Semiwariogramy (c) i (d) wyliczono dla 2700 danych pobranych losowo z siatki o rozdzielczości 1 na 1 km (319114 węzłów).

Podsumowując, mimo że najbardziej dokładne szacunki wartości punktowych uzyskuje się metodami estymacji, do ilustracji rozważań na temat zmienności struktury przestrzennej w

ujęciu obszarowym najlepiej nadają się obrazy symulowane. Z tego względu były one wykorzystywane w rozdziale VI niniejszej rozprawy.



Ryc. 128. Estymowana (średnia oczekiwana IK: *E-type mean*) i symulowana w siatce 1 na 1 km (trzy pierwsze realizacje) zmienność pola rocznych MSDO wyrażonych w mm w roku 1976.

Ostatnim, choć niemniej ważnym, zastosowaniem geostatystyki jest optymalizacja próbkowania. Intuicyjnie wiadomo, że stopień pożądanego zagęszczenia pomiarów powinien być funkcją zmienności czasowej i/lub przestrzennej mierzonego zjawiska. Tam i wtedy gdzie/kiedy zmienność jest niewielka, pomiarów może być mało, gdy zaś coś zmienia się szybko i na niewielkim obszarze, powinniśmy sieć pomiarową zagęścić i zwiększyć ich frekwencję. Geostatystyka oferuje szerokie spektrum narzędzi umożliwiającym optymalizację

próbkowania przy zakładanym stosunku kosztów do efektów i danym zbiorze lokalizacji potencjalnych obiektów badań (Pardo-Igúzquiza, Dowd 2005, van Groeningen i in. 1997, 1999).

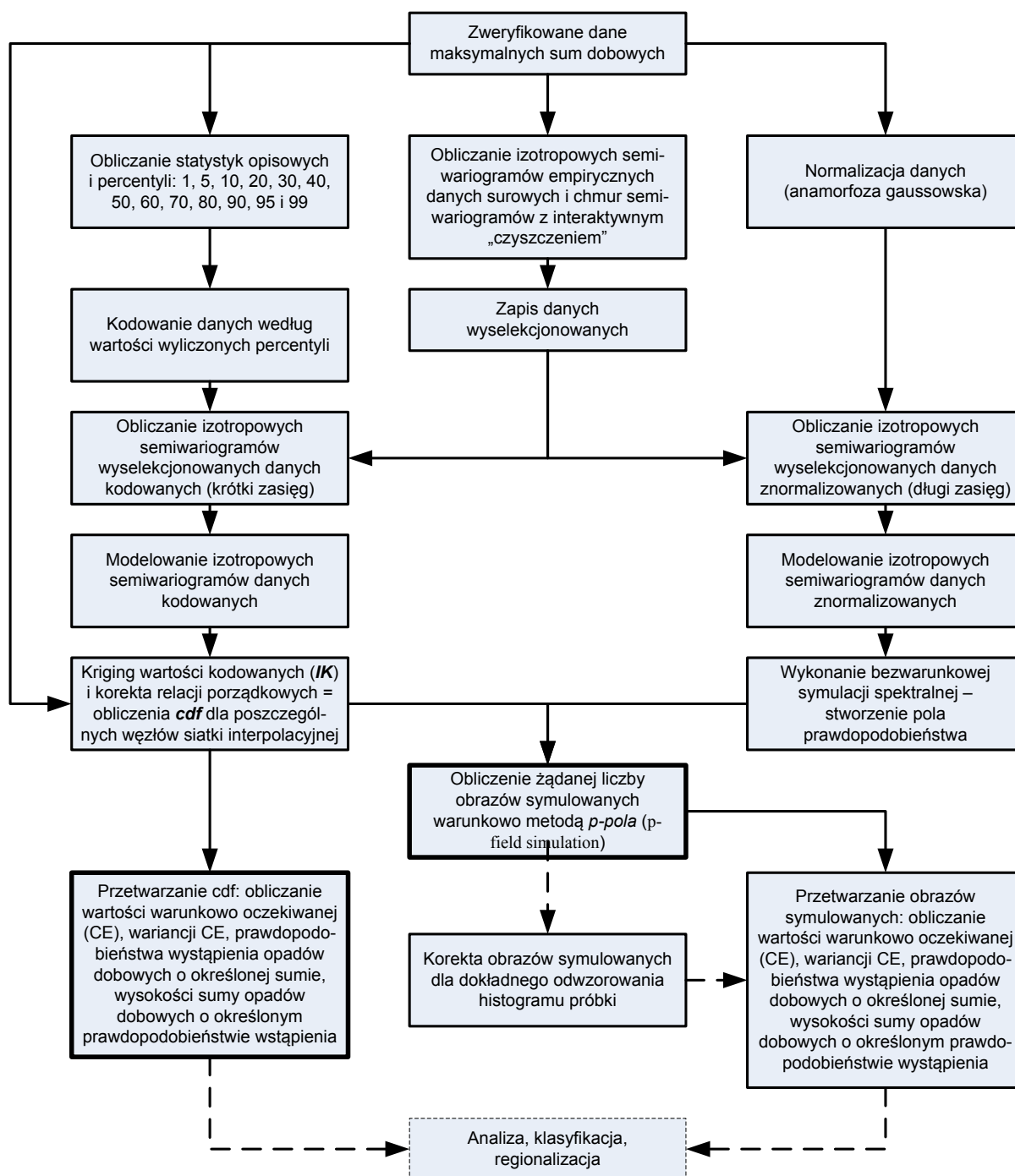
Najczęściej stosowane metody symulacji przestrzennej, tak zwane „sekwencyjne” (ang. *sequential simulation*), mają istotną słabość, która dyskwalifikowała je do zastosowania w niniejszej pracy. Nie można bowiem, albo z dużymi „problemami”, z ich pomocą w wiarygodny sposób odwzorowywać struktur wielkoskalowych symulowanego zjawiska (Deutsch, Journel 1998, Goovaerts 1997). MSDO są zaś efektem kilku procesów działających w różnych skalach przestrzennych: opadów konwekcyjnych, frontalnych i orograficznych. Symulacje metodami sekwencyjnymi są także dość czasochłonne, co przy 325 zbiorach danych jest bardzo istotne (300 zbiorów miesięcznych i 25 rocznych MSD z lat 1956-1980). Zdecydowano się zatem skorzystać z algorytmu symulacji pola prawdopodobieństwa (ang. *p-field simulation*, Srivastava 1992, Froidevaux 1993), który wyżej wymienionych wad jest pozbawiony. Inicjalne pole prawdopodobieństwa o strukturze przestrzennej zgodnej z modelem semiwariogramu analizowanego zbioru danych generowano za pomocą bardzo szybkiej, bezwarunkowej symulacji spektralnej (Yao 1998a, 1998b). Wartości wszystkich węzłów siatki symulacyjnej były następnie jednocześnie losowane z odpowiednich warunkowych kumulacyjnych funkcji rozkładu prawdopodobieństwa (ang. *conditional cumulative distribution function – ccdf*), uzyskanych za pomocą metody krigingu wartości kodowanych.

Kompletny schemat wykonywanych obliczeń struktury przestrzennej, estymacji i symulacji MSDO przedstawiono na rycinie 129. Zaznaczono na nim, jaka część całego projektu badawczego została wykorzystana w niniejszej rozprawie. Normalizację zmiennych oraz obliczenia i modelowanie semiwariancji wykonano za pomocą programu ISATIS (Bleinès i in. 2006), natomiast pozostałą część procedury – korzystając z programu IKSIM (Ying 2000).

Decyzje o zastosowanej metodyce estymacji i symulacji pola prawdopodobieństwa MSDO pociągnęły za sobą istotne konsekwencje w zakresie sposobu analizy i modelowania ich struktury przestrzennej.

W szczególności, użycie do estymacji pola prawdopodobieństwa MSDO krigingu danych kodowanych spowodowało konieczność kodowania binarnych danych względem takiej ilości progów, która w zadawalająco dokładny sposób pozwoli modelować lokalne dystrybuanty. Z

danych binarnych dla każdej wartości progowej obliczane były semiwariogramy empiryczne, do których następnie dopasowywano modele. Ta ścieżka metodyczna, choć skomplikowana i czasochłonna, dała unikatową możliwość analizy zróżnicowania struktury przestrzennej MSDO w różnych klasach ich wysokości.



Ryc. 129. Schemat całej procedury analizy i modelowania struktury przestrzennej oraz estymacji i symulacji pola prawdopodobieństwa maksymalnych opadów dobowych dla pojedynczego zbioru danych miesięcznych lub rocznych. Operacje opcjonalne zaznaczono linią przerywaną. Kroki uwzględnione w niniejszej pracy zaznaczono szarym tłem.

Inicjalne bezwarunkowe pole prawdopodobieństwa, używane przy symulacji metodą *p-pola*, było generowane w przestrzeni danych znormalizowanych. Wymagało to opracowania modeli konwersji poszczególnych podzbiorów MSDO do postaci standardowego rozkładu normalnego, a następnie wyliczenie z nich empirycznych miar struktury przestrzennej i opracowanie ich matematycznych modeli. Jak to szczegółowo opisano w dodatku X.4, to pozornie kolejne „utrudnienie” metodyczne miało dwie istotne zalety. Po pierwsze umożliwiało bezpośrednio porównywanie semiwariogramów poszczególnych zbiorów MSDO, a po drugie – na etapie modelowania redukowało destrukcyjny wpływ danych anomalnych, nie zmieniając w żaden sposób autentycznych prawidłowości autokorelacji przestrzennej analizowanych danych.

X.3. Anizotropia struktury przestrzennej

Podobieństwo / zróżnicowanie zmienia się w przestrzeni nie tylko wraz z odległością, ale także najczęściej również z kierunkiem. Właściwość ta jest nazywana anizotropią struktury przestrzennej (Goovaerts 1997, Gringarten, Deutsch 2001, Olea 2006). Wyróżnia się dwie jej odmiany:

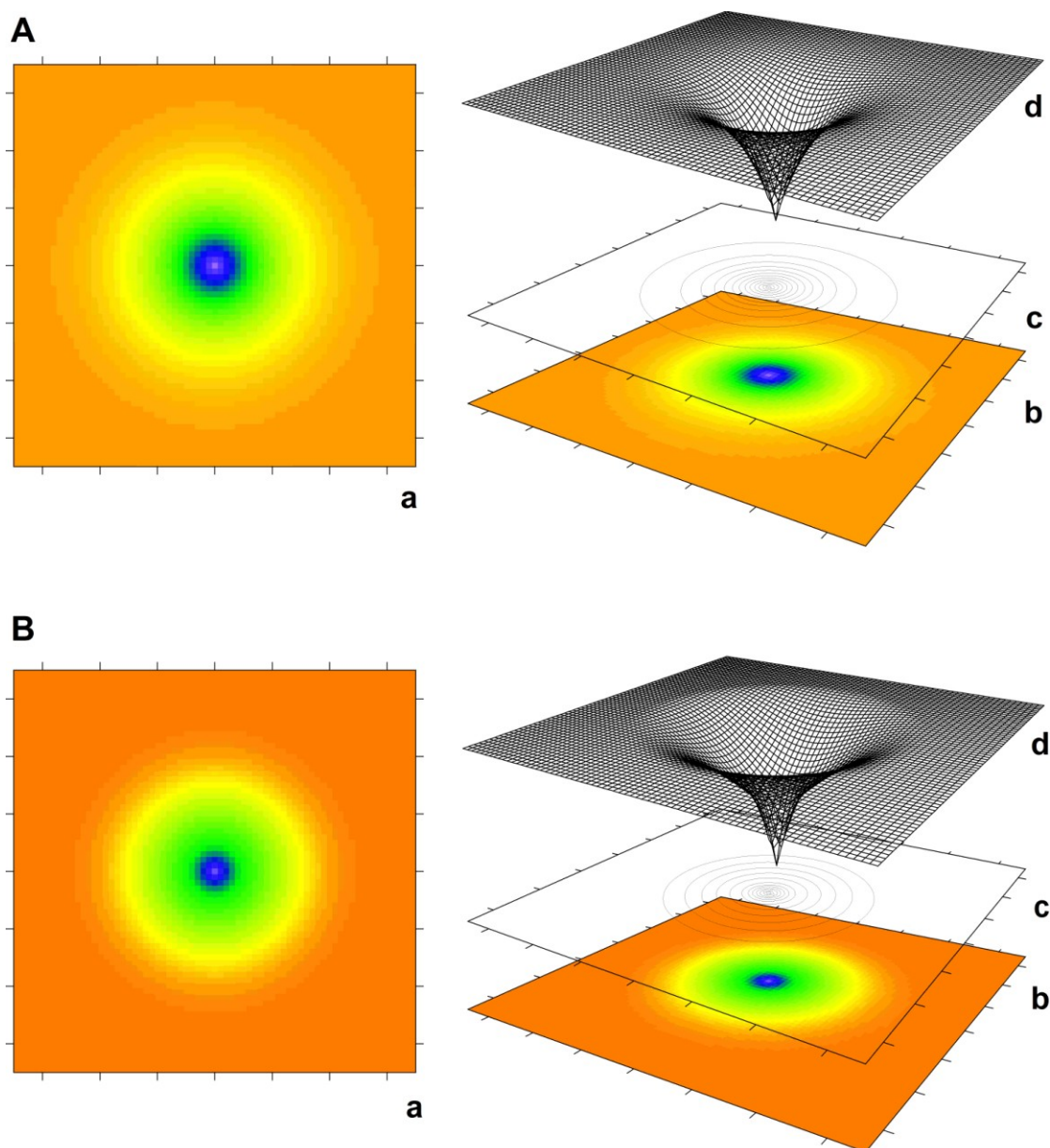
- geometryczną, kiedy semiwariogramy kierunkowe mają taki sam kształt i wariancję progową, a różnią się zasięgiem; róża zasięgów dla poszczególnych kierunków tworzy kształt elipsy,
- oraz strefową, kiedy semiwariogramy dla poszczególnych kierunków różnią się wariancją progową.

Często spotykane są także kombinacje obu typów.

Charakter anizotropii identyfikuje się na podstawie tak zwanej powierzchni semiwariogramu (ang. *semivariogram surface*, Deutsch, Journel 1998, Pannatier 1996, ryc. 130 i 131). Oblicza się ją dla ustalonej siatki przedziałów odległości i kierunku – najczęściej w układzie współrzędnych prostokątnych⁵⁸. Ponieważ $\gamma(\mathbf{h}) = \gamma(-\mathbf{h})$, powierzchnia taka jest w każdym kierunku symetryczna względem początku układu. Jeśli analizowana cecha wykazuje identyczny spadek podobieństwa w każdym kierunku (struktura izotropowa) i ustaloną wariancję progową, to geometrycznym tego obrazem jest płaszczyzna o kształcie symetrycznego „lejka” z kołnierzem (ryc. 130A). Jego przekrój w każdej płaszczyźnie równoległej do osi odległości jest okręgiem. Kiedy izotropowy semiwariogram ma charakter

⁵⁸ Kierunek definiuje się poprzez podanie wielkości współrzędnych X i Y (tangens kąta).

zagnieżdżony (złożony), wówczas odzwierciedla się to, oczywiście w kształcie lejka, w postaci segmentów o różnym nachyleniu (ryc. 130B).

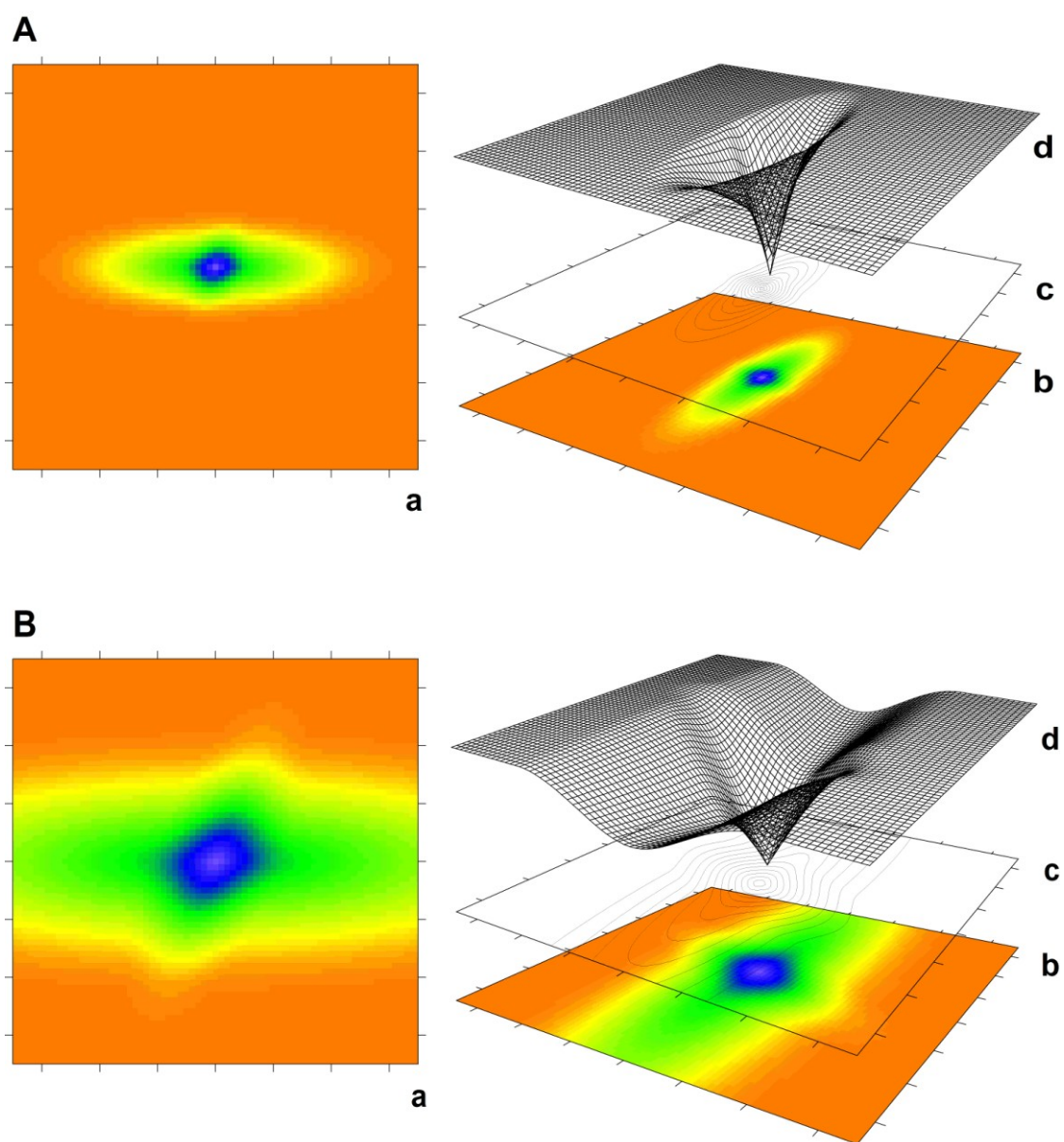


Ryc. 130. Schemat powierzchni modeli semiwariancji prostych (A) i złożonych (B) przy izotropowej autokorelacji danych: a – mapa rastrowa w projekcji płaskiej, b, c, d – obraz w rzucie perspektywnym, odpowiednio: mapa rastrowa, izoliniowa i siatkowy model trójwymiarowy.

Przy anizotropii geometrycznej „lejek” jest spłaszczony – ma w każdej płaszczyźnie równoległej do osi odległości przekrój elipsy. Każdy element⁵⁹ anizotropowego semiwariogramu złożonego może mieć inny kierunek i stosunek anizotropii – elipsy będą się wtedy różnić orientacją i mimośrodem (ryc. 131A). Kiedy semiwariogramy kierunkowe

⁵⁹ Poza składową nuggetową.

różnią się wariancją progową, lecz mają ten sam zasięg (anizotropia strefowa), wtedy kołnierz lejka niepodobieństwa nie stanowi płaszczyzny równoległej od osi odległości, lecz tworzy sfalowaną powierzchnię o kształcie siodła, a jego przekroje są okręgami. Najbardziej skomplikowany kształt występuje przy złożonej anizotropii geometryczno-strefowej – lejek ma segmenty różniące się kierunkiem i/lub stopniem spłaszczenia oraz sfalowany kołnierz (ryc. 131B). Dla uproszczenia powierzchnię semiwariogramu przedstawia się najczęściej na płaszczyźnie w postaci mapy izoliniowej lub pikselowej (rastrowej). W tej drugiej kolorem lub natężeniem szarości obrazuje się wartość semiwariancji dla danego przedziału odległości wzdłuż osi X i Y.



Ryc. 131. Schemat powierzchni złożonych modeli semiwariancji anizotropowych przy geometrycznym (A) i geometryczno-strefowym (B) typie anizotropii: a – mapa rastrowa w projekcji płaskiej, b, c, d – obraz w rzucie perspektywicznym, odpowiednio: mapa rastrowa, izoliniowa i siatkowy model trójwymiarowy.

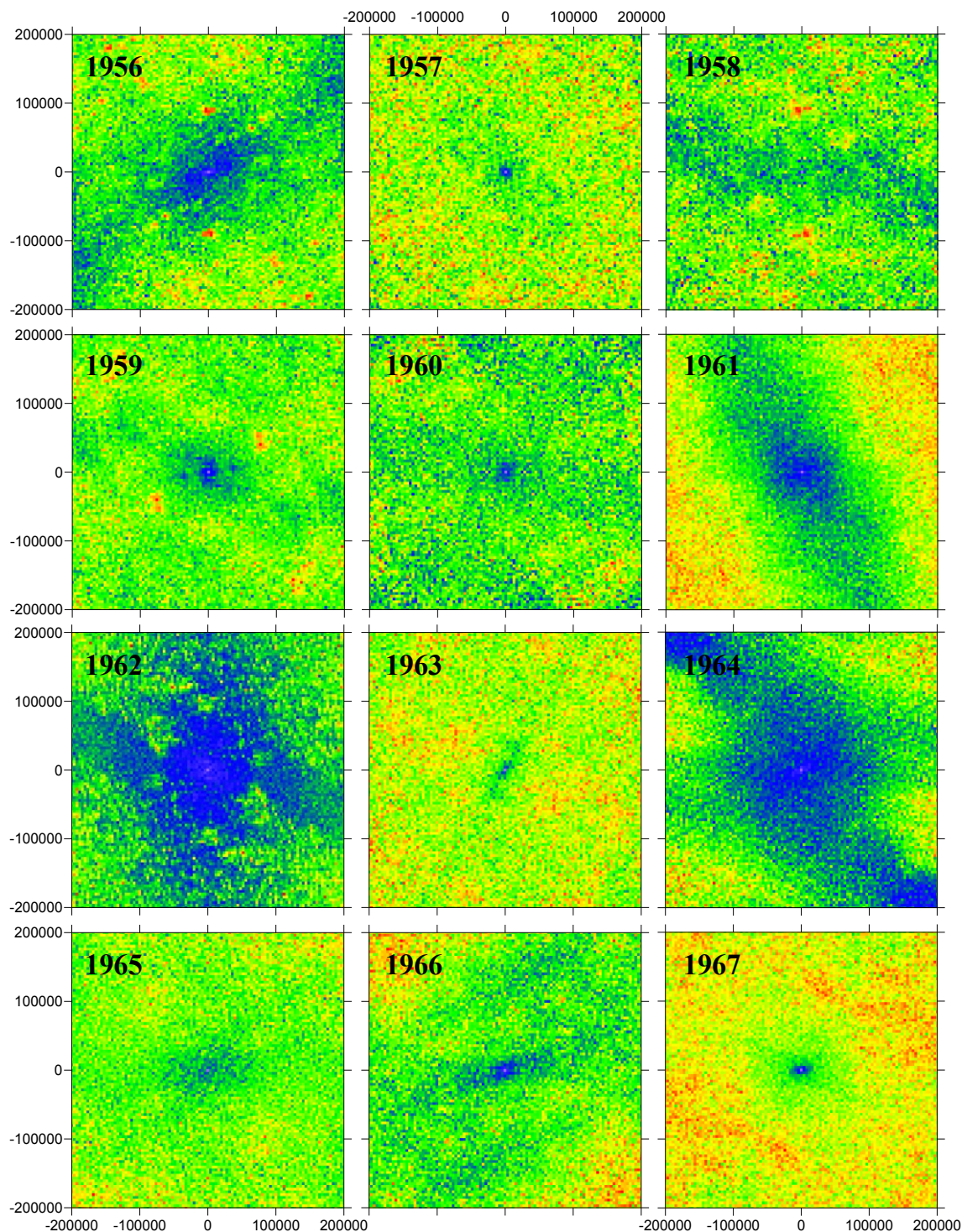
Jedną z najistotniejszych decyzji metodologicznych, jaką musiano podjąć, dotyczyła uwzględniania lub pominięcia anizotropii przy analizie struktury przestrzennej MSDO. Indywidualne opady wykazują najczęściej bardzo silną anizotropię (stosunek zasięgów zbliżony lub mniejszy niż 0,5), a jej azymut związany jest z kierunkiem adwekcji mas powietrza (Bacchi, Kottogoda 1995, Berne i in. 2006, Moszkowicz 2000). Wraz ze wzrostem długości czasu kumulacji opadów (doba, dekada, miesiąc, rok), i powierzchni analizowanego obszaru, maleje rola pojedynczych opadów, a pojawia się i zaczyna dominować wpływ kierunków i natężenia ogólnej cyrkulacji strefowej. Takiej prawidłowości można się spodziewać prowadząc analizę „pełnych” serii pomiarowych. Dane MSDO również w tym kontekście są specyficzne. Miesięczne i roczne MSDO to połączone, bardzo zmienne pod względem proporcji w stosunku do całości, podzbiory wielu sum dobowych zarejestrowanych na obszarze Polski w różnych terminach. Każdy z tych opadów charakteryzował się odmienną anizotropią wynikającą z jego genezy i przebiegu. Można się było spodziewać, że będzie to skutkowało, szczególnie w przypadku MSDO rocznych, występowaniem mało zróżnicowanej „mieszanki”. Z drugiej strony losowy charakter wysokich opadów mógł spowodować, że w przypadku danych miesięcznych pojedynczy, silnie anizotropowy, opad zanotowany w wielu stanowiskach zdominuje obraz struktury przestrzennej uzyskany dla całej Polski. Wszystko to przemawiało za pominięciem w analizie aspektu zróżnicowania kierunkowego MSDO.

Potwierdzeniem powyższych spekulacji są mapy rastrowe powierzchni semiwariogramów wykonane dla wszystkich rocznych MSDO oraz miesięcznych z jednego przykładowego roku z analizowanego wielolecia (ryc. 132-134). Wartości semiwariancji obliczono dla 100 odstępów po 4 km wzdłuż osi W – E (X) i N – S (Y). Każda wartość semiwariancji w komórce mapy jest średnią z od około 50-510 par danych⁶⁰. Przeciętna ilość par z których obliczona była wartość semiwariancji przekraczała 100, co zapewnia dostateczny poziom reprezentatywności statystycznej (Webster, Oliver 1992). W zastosowanej standardowej, tęczyowej paletce barw kolorem niebieskim zaznaczono niskie wartości semiwariancji, a więc wysokie podobieństwo danych, kolorem czerwonym – maksymalne jej wartości, czyli najmniejsze podobieństwo.

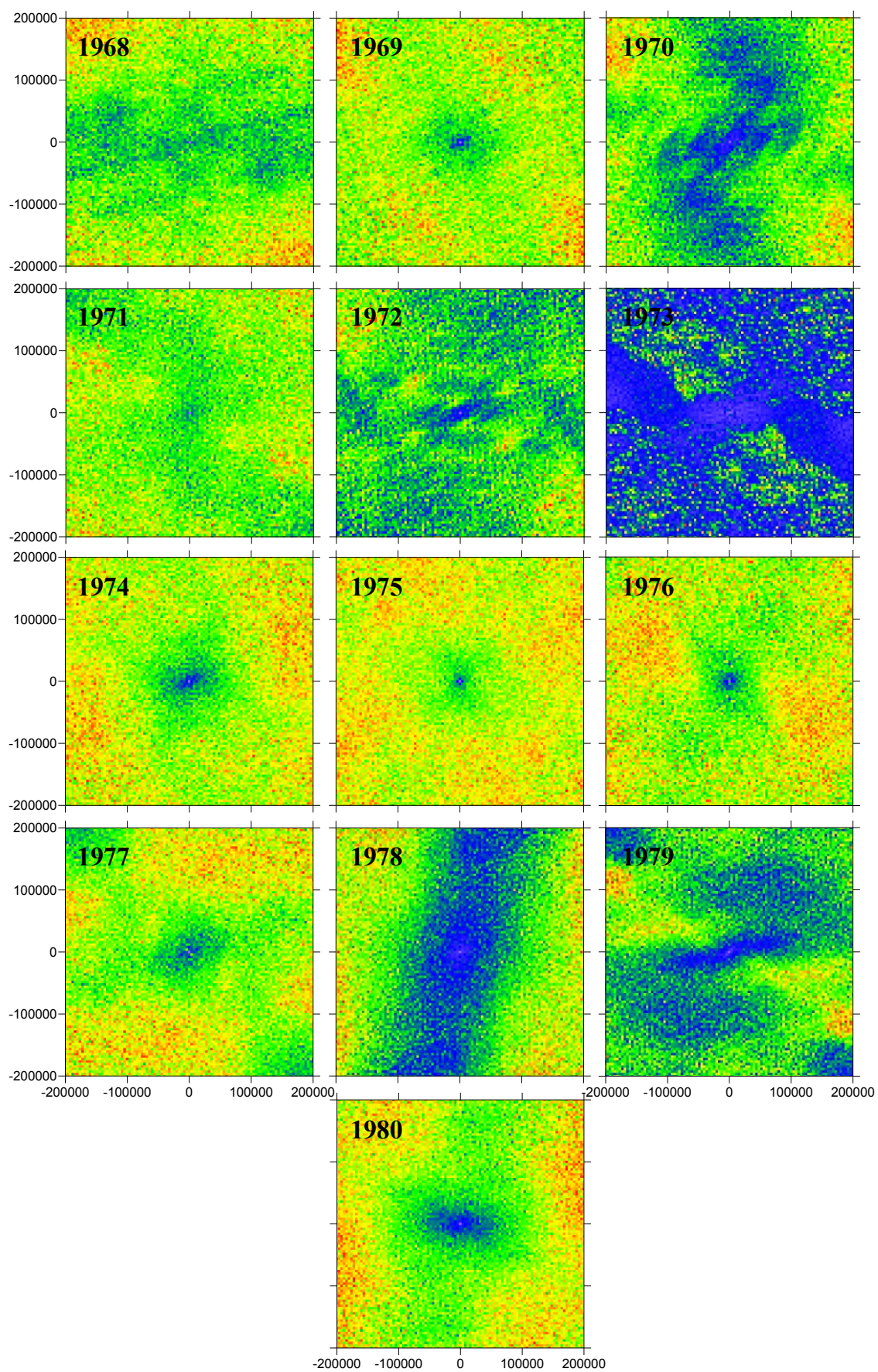
Anizotropia struktury przestrzennej rocznych MSDO (ryc. 132 i 133) w 9 latach z analizowanego wielolecia w ogóle się nie zaznacza (lata 1957, 1959, 1960, 1963, 1965, 1967, 1969, 1974, 1975), a w kolejnych 9 jest bardzo słabo zarysowana (lata 1958, 1962, 1966, 1968, 1971, 1972, 1973, 1976, 1977). W pozostałych 7 (lata 1956, 1961, 1964, 1970, 1978,

⁶⁰ Z wyjątkiem komórki centralnej, w której zawsze było ponad 2000 par.

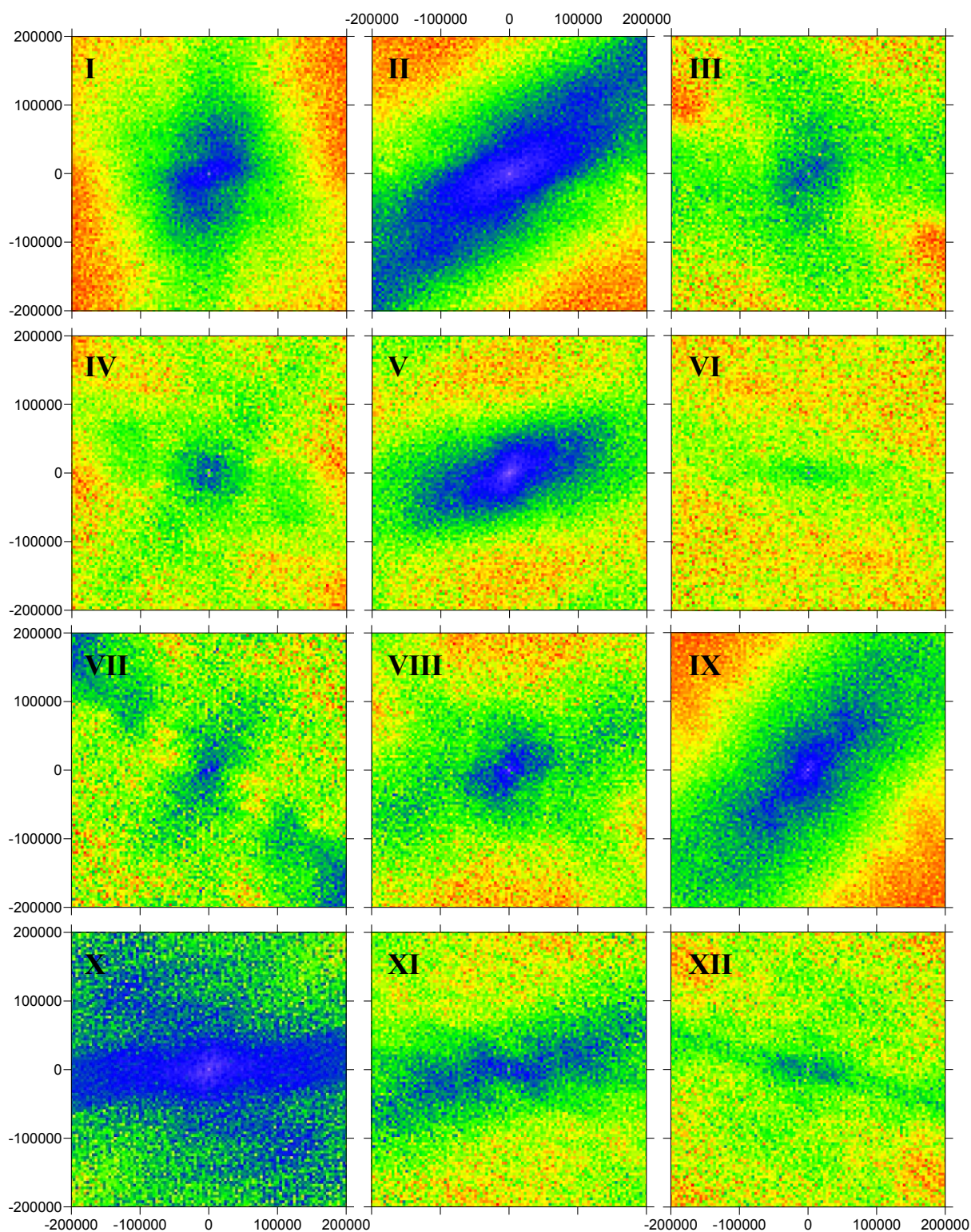
1979, 1980) jest lepiej widoczna, ale ma raczej charakter mocno „rozmażany”, świadcząc o złożonym, heterogenicznym charakterze analizowanego pola. Wyraźnie jednak dominują układy południkowe NW – SE (lata 1961, 1964, 1979, 1978, 1979, 1980). W drugiej grupie lat anizotropia ma głównie charakter geometryczny, w trzeciej zaś – geometryczno-strefowy.



Ryc. 132. Powierzchnie (mapy) semiwariogramów dla lat 1956-1967. Zakres wartości przedstawionych na każdej mapie jest różny. Każdy piksel ma 4 na 4 km. Szczegółowe objaśnienia w tekście.



Ryc. 133. Powierzchnie (mapy) semiwariogramów dla lat 1968-1980. Zakres wartości przedstawionych na każdej mapie jest różny. Każdy piksel ma 4 na 4 km. Szczegółowe objaśnienia w tekście.



Ryc. 134. Powierzchnie (mapy) semiwariancji dla poszczególnych miesięcy roku 1977. Zakres wartości przedstawionych na każdej mapie jest różny. Każdy piksel ma 4 na 4 km. Szczegółowe objaśnienia w tekście.

Na mapach semiwariogramów miesięcznych MSDO z roku 1977 obraz jest nieco inny (ryc. 134). Na większości z nich anizotropia zaznacza się bardzo wyraźnie, ale ma zazwyczaj charakter „przecinających” się elips (anizotropia geometryczna) bądź pasów (anizotropia

strefowa). Świadczy to o występowaniu w obrębie analizowanego obszaru podzbiorów danych pomiarowych o różnym ukierunkowaniu rozkładu zmierzonych MSDO i różnym zasięgu. Jedynie w lutym 1977 rozkład przestrzenny maksymalnych opadów dobowych charakteryzował się silną i w miarę jednorodną anizotropią strefową. Przykład losowo wybranego z całego wielolecia roku potwierdza przypuszczenie o skomplikowanym charakterze kierunkowych relacji podobieństwa miesięcznych MSDO, wynikającym z ich niesynchronizacji i ograniczonego, do niedużej części terytorium Polski, zasięgu.

Nie można również pominąć bardzo ważnego względu praktycznego. O ile obliczanie empirycznych kierunkowych miar struktury przestrzennej nie jest specjalnie kłopotliwe i nie wydłużyłoby w sposób znaczący czasu pracy, to etap modelowania byłby radykalnie odmienny. Tworzenie, przy pomocy stosowanej w niniejszej rozprawie metody półautomatycznej (por. podrozdz. III.2.5.3-III.2.5.6), złożonego modelu anizotropowego trwa kilka razy dłużej niż modelu izotropowego. Ponieważ w sumie wykonano 4550 modeli izotropowych i zajęło to ponad pół roku bardzo intensywnej pracy, uwzględnienie anizotropii uniemożliwiłoby zakończenie opracowania w okresie wynikającym z realizowanego projektu badawczego. Problem ten wymaga jednak dalszej uwagi i będzie tematem osobnego opracowania przy zastosowaniu odmiennej, nieparametrycznej, metodyki (Yao, Journal 1998, Ma, Yao 2001).

X.4. Normalizacja danych

X.4.1. Ogólna charakterystyka procedur normalizacji danych wykorzystywanych w geostatystyce

Istotnym etapem procedury symulacji metodą pola prawdopodobieństwa (ang. *p-field*)⁶¹ była normalizacja danych. Używane one były do opracowania modelu struktury przestrzennej symulowanego zjawiska (patrz rozdz. III.2.5.7. i VI). Model ten był następnie wykorzystany do wygenerowania inicjalnego, bezwarunkowego pola prawdopodobieństwa⁶². Przestrzennie skorelowane liczby losowe były potrzebne do próbkowania uzyskanych metodą krigingu wartości kodowanych warunkowych krzywych rozkładu prawdopodobieństwa (patrz dodatek X.2). Taka procedura zapewniała uzyskanie dwóch warunków nakładanych na wszystkie

⁶¹ Element ten, jak wspomniano we wprowadzeniu (rozdz. I) i omówieniu metodyki (rozdz. III.2, dodatek X.2), zostanie zaprezentowany w odrębnej publikacji.

⁶² Bezwarunkowa symulacja przestrzenna oparta jest jedynie na modelu struktury przestrzennej analizowanej cechy. Nie są brane pod uwagę wyniki ewentualnych pomiarów.

metody symulacji przestrzennej: a) zapewnienie zgodności rozkładu statystycznego danych symulowanych i empirycznych; b) zapewnienie zgodności struktury przestrzennej danych symulowanych i empirycznych. Normalizacja danych ma jednak w geostatystyce znacznie szersze zastosowania niż tylko w kontekście wykorzystywanej w niniejszej pracy metodyki (Chilés, Delfiner 1999, Goovaerts 1997). Nie należy jej mylić ze standaryzacją i/lub normalizacją w sensie statystycznym (StatSoft Inc., 2004). W naukach o Ziemi zmienne pomiarowe wykazują często asymetryczny rozkład z niewielką ilością bardzo dużych wartości (skośność dodatnia, ryc. 138A). W wielu procedurach geostatystycznych wymagane jest jednakże, albo bardzo korzystne, aby przetwarzane dane posiadały rozkład normalny. Normalizacja danych to nieliniowa odwracalna transformacja jakiegokolwiek ciągłej skumulowanej funkcji rozkładu (ang. *cumulative distribution function = cdf*) w standaryzowany rozkład gaussowski o średniej równej zero i odchyleniu standardowym równym jeden⁶³. Odwracalność transformacji zakłada możliwość uzyskania z powrotem z danych znormalizowanych oryginalnego, nieobciążonego *cdf*⁶⁴.

W geostatystyce stosowane są w praktyce dwa sposoby normalizacji danych: nieparametryczny i parametryczny. Dają one w zasadzie równoważne wyniki. Metoda nieparametryczna, zwana najczęściej *normal score transform* (Deutsch, Journel 1998, Goovaerts 1997, Olea 1999) polega na prostym, „graficznym” przekodowaniu kwantyli empirycznego rozkładu posiadanych danych na odpowiednie kwantyle standaryzowanego rozkładu normalnego. Zachowana tablica przekodowania umożliwia dokonanie w razie potrzeby transformacji odwrotnej. Taki sposób normalizacji danych jest dostępny w pakiecie GSLIB (Deutsch, Journel 1998). Słabością metody nieparametrycznej jest konieczność osobnego przeprowadzenia wygładzania rozkładu empirycznego, kiedy próbka jest zbyt mała i/lub ze względu na duże błędy pomiarowe jest on bardzo chaotyczny (szczególnie przy ogonach rozkładu). Parametryczna alternatywa jest tej wady pozbawiona. Wygładzanie rozkładu empirycznego jest jej integralną składową. Z tego względu właśnie ona została użyta w niniejszym opracowaniu.

⁶³ W ujęciu bardziej ogólnym, transformację jakiegokolwiek ciągłego *cdf* w dowolny inny *cdf*.

⁶⁴ „Normalizację” prawoskośnych danych bardzo często uzyskuje się przez logarytmowanie. Rzadko jednak jest to procedura w pełni skuteczna, to jest całkowicie „likwidująca” skośność. Poza tym, ma dwie poważne wady: a) może być zastosowana jedynie do danych dodatnich (wymaga przeskalowania wartości, jeśli są w nich równe lub mniejsze od zera), b) transformacja odwrotna (antylogarytmowanie) jest obciążona (Goovaerts 1997).

X.4.2. Normalizacja parametryczna z użyciem wielomianów Hermite

Normalizacja parametryczna zwana jest często w literaturze „anamorfozą gaussowską” (ang. *Gaussian anamorphosis*, Chilés, Delfiner 1999, Webster, Oliver 2001, Wackernagel 2003). Dokonuje się ona z użyciem serii wielomianów Hermite (Weisstein 2006). Punktem wyjścia jest chęć znalezienia takiej funkcji Φ , która umożliwiłaby transformację dowolnego rozkładu empirycznego $Z(\mathbf{x})$ w standardowy rozkład normalny $Y(\mathbf{x})$:

$$Y(\mathbf{x}) = \Phi[Z(\mathbf{x})] \quad [33]$$

Jej odwrotność z kolei Φ^{-1} umożliwia dokonanie konwersji w drugą stronę:

$$Z(\mathbf{x}) = \Phi^{-1}[Y(\mathbf{x})] \quad [34]$$

Wzór na funkcję gęstości prawdopodobieństwa standardowego rozkładu normalnego jest następujący:

$$g(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) \quad [35]$$

Wielomiany Hermite są powiązane z powyższą funkcją i mogą być przedstawione w postaci:

$$H_k(y) = \frac{1}{\sqrt{k!g(y)}} \cdot \frac{d^k g(y)}{dy^k} \quad [36]$$

gdzie k jest stopniem wielomianu mogącym przybrać wartości $0, 1, 2, \dots$, a $1/\sqrt{k!}$ – współczynnikiem standaryzacji. Dwa pierwsze wielomiany Hermite, to jest dla $k = 0$ i $k = 1$ są następujące:

$$H_0(y) = 1, \quad [37]$$

$$H_1(y) = -y, \quad [38]$$

i dlatego wielomiany wyższego stopnia stosują się do zasady rekurencyjnej:

$$H_k(y) = -\frac{1}{\sqrt{k}} y H_{k-1}(y) - \sqrt{\frac{k-1}{k}} H_{k-2}(y). \quad [39]$$

Tak więc, wielomiany te mogą być dla standardowego rozkładu normalnego obliczane do dowolnego stopnia.

Wielomiany Hermite są ortogonalne w odniesieniu do funkcji wazącej $\exp(-y^2/2)$ w przedziale od $-\infty$ do $+\infty$. Są one zatem niezależnymi składnikami rozkładu normalnego o ciągle rosnącej, wraz ze stopniem wielomianu, złożoności. Praktycznie każda funkcja $Y(\mathbf{x})$ może być przedstawiona jako ich suma:

$$f\{Y(\mathbf{x})\} = f_0 H_0\{Y(\mathbf{x})\} + f_1 H_1\{Y(\mathbf{x})\} + f_2 H_2\{Y(\mathbf{x})\} + \dots, \quad [40]$$

a ponieważ wielomiany Hermite są ortogonalne:

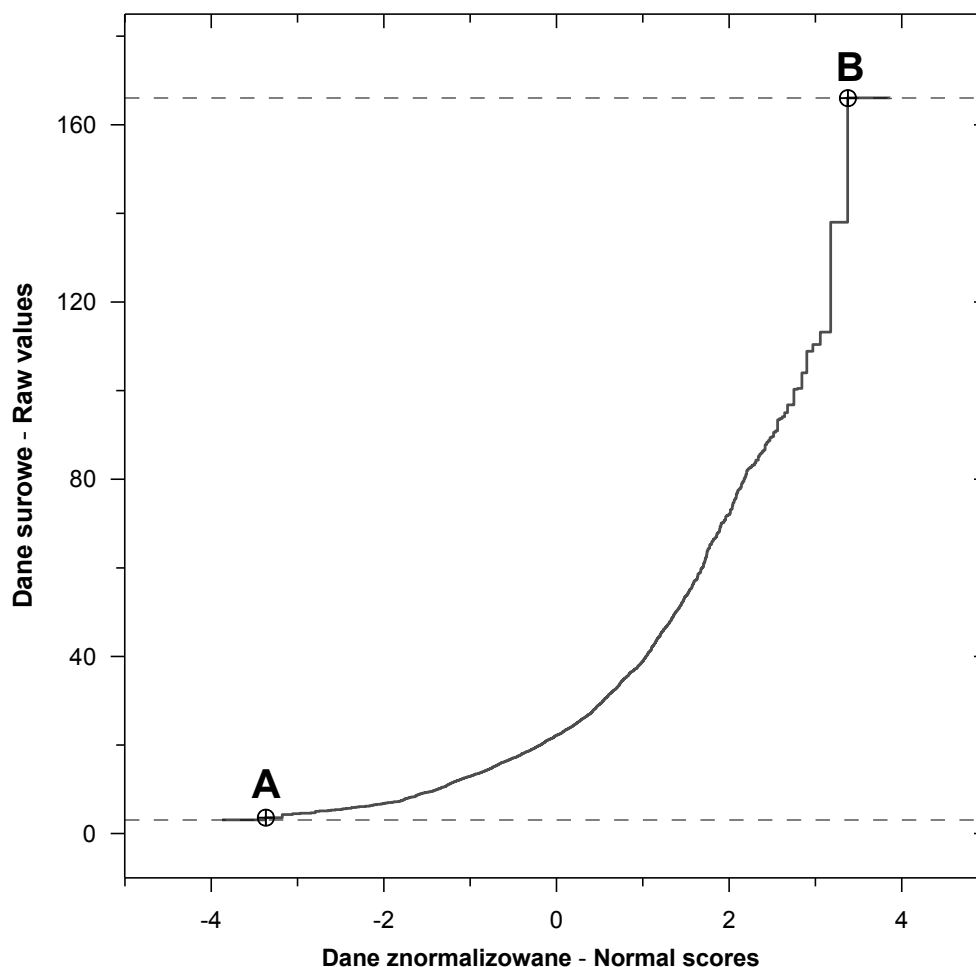
$$\begin{aligned} E\left[f\{Y(\mathbf{x})\} H_k\{Y(\mathbf{x})\}\right] &= E\left[H_k\{Y(\mathbf{x})\} \sum_{l=0}^{\infty} f_l H_l\{Y(\mathbf{x})\}\right] \\ &= \sum_{l=0}^{\infty} f_l E\left[H_l\{Y(\mathbf{x})\} H_k\{Y(\mathbf{x})\}\right] \\ &= f_k \end{aligned} \quad [41]$$

Pozwala to obliczyć współczynniki ϕ_k funkcji $\Phi[Y(\mathbf{x})]$ z równania [33] jako:

$$\begin{aligned} Z(x) &= \Phi[Y(\mathbf{x})] \\ &= \phi_0 H_0\{Y(\mathbf{x})\} + \phi_1 H_1\{Y(\mathbf{x})\} + \phi_2 H_2\{Y(\mathbf{x})\} + \dots \\ &= \sum_{k=0}^{\infty} \phi_k H_k\{Y(\mathbf{x})\} \end{aligned} \quad [42]$$

Transformacja dokonana z użyciem wielomianów Hermite jest oczywiście w pełni odwracalna [wzór 34], co oznacza, że wyniki obliczeń na danych znormalizowanych mogą być w każdym momencie przekonwertowane do oryginalnej skali pomiarowej.

Szczegóły algorytmu wyznaczania współczynników wielomianów Hermite opisane są przez Webstera i Oliver (2001) oraz Wackernagla (2003). W niniejszym opracowaniu anamorfozę gaussowską wykonano w programie ISATIS (Bleinès i in. 2007). Modelowanie anamorfozy rozpoczyna się od konstrukcji dyskretnej krzywej rozkładu danych surowych (pomiarowych, ryc. 135). Użytkownik ma możliwość modyfikowania kształtu modelu na jego krawędziach poprzez podanie wartości 2 punktów kontrolnych (A i B na ryc. 135). Ta opcja ma znaczenie tylko wtedy, kiedy liczba pomiarów jest niewielka. Poza tym można podać interwał danych pomiarowych (ang. *Authorized Interval on the Raw Variable*) zdefiniowany jako różnica pomiędzy wartością minimum $Z_{a_{\min}}$ i maksimum $Z_{a_{\max}}$, oraz rząd rozwinięcia wielomianów Hermite (ilość wielomianów). Domyślnymi wartościami interwału są minimum i maksimum danych empirycznych. Przy takiej konfiguracji dwa wymienione wyżej punkty kontrolne nie powodują zmiany krzywej empirycznej.

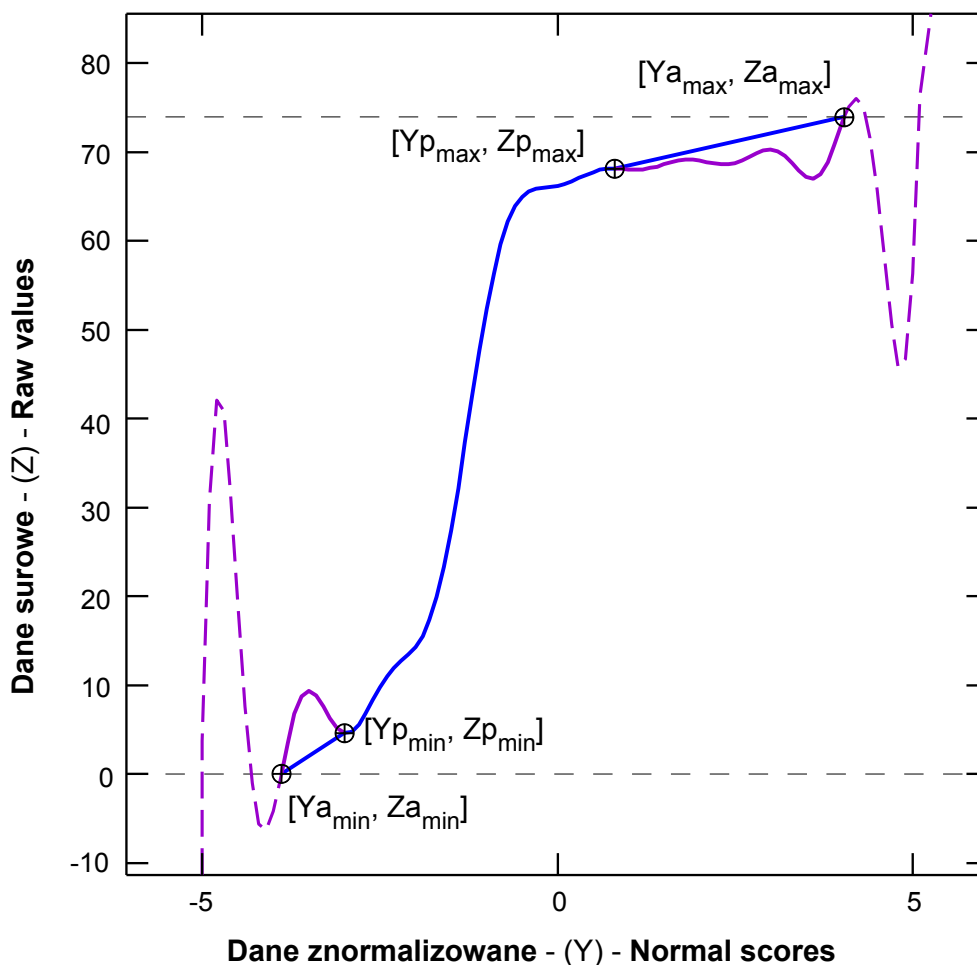


Ryc. 135. Usytuowanie dwóch punktów kontrolnych anamorfozy gaussowskiej z użyciem wielomianów Hermite (Bleinès i in. 2007). Szersze objaśnienia w tekście.

Po zdefiniowaniu warunków anamorfozy dyskretnej program oblicza współczynniki ψ_i rozwinięcia wielomianów Hermite. Wykreślana jest krzywa i obliczany jest „praktyczny” oraz „absolutny interwał definicji” (ang. *Practical Interval of Definition*, *Absolute Interval of Definition*, ryc. 136). Stąd, wiadomo, że granicami:

- „praktycznego interwału definicji” są punkty $[Y_{p_{\min}}, Z_{p_{\min}}]$ i $[Y_{p_{\max}}, Z_{p_{\max}}]$ (ryc. 136). Są to miejsca gdzie modelowana krzywa przecina górny i dolny kres interwału danych pomiarowych ($Z_{a_{\min}}, Z_{a_{\max}}$), lub też punkty gdzie wartość na krzywej empirycznej nie zmienia się przy zmianach Y .
- „absolutnego interwału definicji” są punkty $[Y_{a_{\min}}, Z_{a_{\min}}]$ i $[Y_{a_{\max}}, Z_{a_{\max}}]$ (ryc. 136). Zlokalizowane są one na przecięciu modelowanej krzywej z horyzontalnymi liniami określającymi interwał danych pomiarowych. Wartości generowane przez funkcję anamorfozy poza „absolutnym interwałem definicji” nie powinny być w

żadnym przypadku uwzględniane. Krzywa wykazuje bowiem tam najczęściej chaotyczne fluktuacje (ryc. 136 i 137). Może to stanowić istotne ograniczenie w sytuacji, kiedy symulowane w przestrzeni danych znormalizowanych wartości wykraczają poza „absolutny interwał definicji”. Próba konwersji do oryginalnej skali pomiarowej daje wtedy nierealistyczne wyniki. Istnieje jednak podawane przez Wackernagla (2003) proste rozwiązanie tego problemu.



Ryc. 136. Schemat określający „praktyczny” i „absolutny interwał definicji” modelu anamorfozy gaussowskiej (Bleinès i in. 2007). Dokładny opis w tekście.

W praktycznej realizacji anamorfoza jest zdefiniowana jako funkcja $Y: Z(\mathbf{x}) = \Phi[Y(\mathbf{x})]$ (czyli odwrotnie niż we wzorach [33] i [34]), więc aby dokonać transformacji danych surowych w gaussowskie należy dokonać jej inwersji do postaci $Y(\mathbf{x}) = \Phi^{-1}[Z(\mathbf{x})]$. W programie ISATIS można dokonać tego trzema sposobami (Bleinès i in. 2007), z których zalecana jest inwersja na bazie frekwencji (ang. *frequency inversion*). Na początek program sortuje wartości pomiarowe. Dla każdej próbki FC_i , poczynając od najmniejszej obliczana jest skumulowana frekwencja [43]:

$$FC_i = FC_{i-1} + W_i \quad [43]$$

Zmienna ważąca W_i jest podawana przez operatora lub też obliczana jako $W_i = 1/N$. Dzięki temu dwie próbki o identycznych wartościach empirycznych będą miały inne skumulowane frekwencje. Finalne obliczenie danej znormalizowanej polega na wyznaczeniu średniej arytmetycznej dwóch kolejnych wyrazów szeregu:

$$Y_i = \frac{\Phi^{-1}(FC_i) + \Phi^{-1}(FC_{i-1})}{2} \quad [44]$$

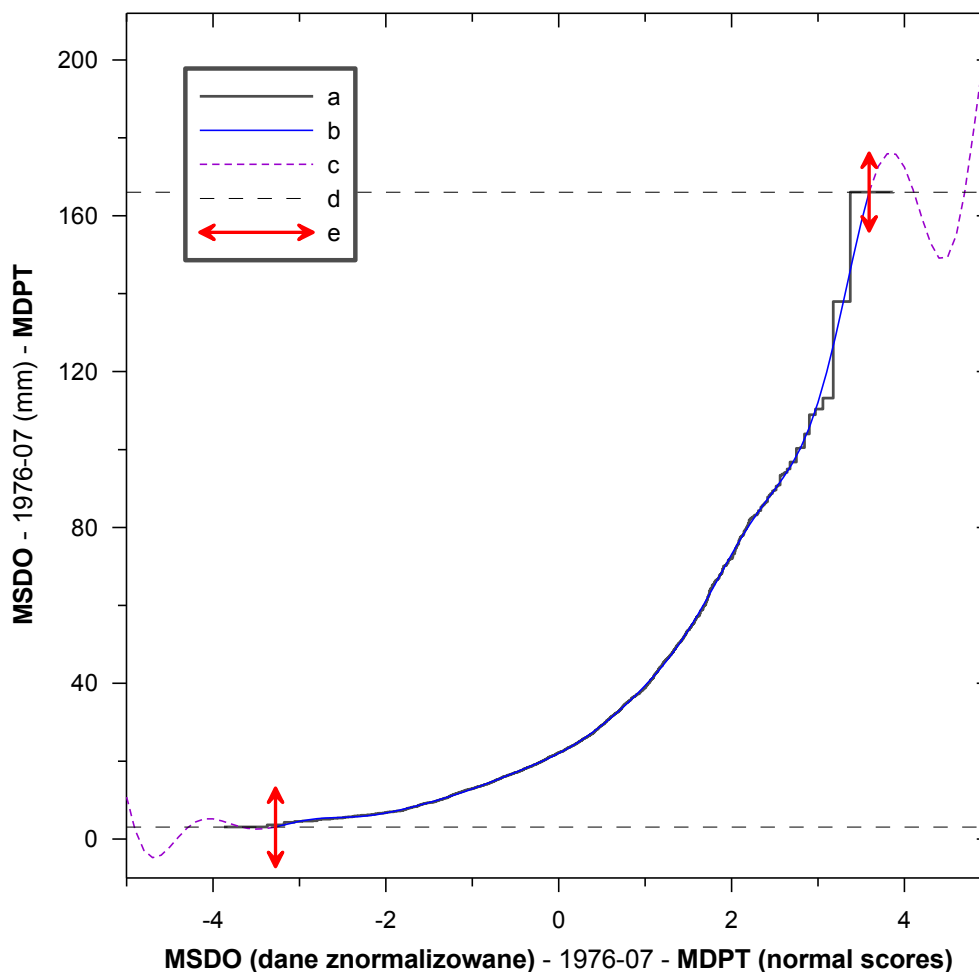
Konsekwentnie zatem dwie identyczne dane pomiarowe będą miały odmienne wartości znormalizowane.

X.4.3. Przykład normalizacji danych MSDO

Do zilustrowania zagadnienia normalizacji danych użyto zbioru miesięcznych MSDO z lipca 1976. Charakteryzowały się one typowym skośnym rozkładem, z wysoką koncentracją wartości niskich i średnich (62,3% w przedziale 10-30 mm) i maksimum 166,1 mm (ryc. 138A, tab. 30). Anamorfozę przeprowadzono posługując się standardowym ustawieniem opcji, to jest z interwałem danych pomiarowych (ang. *Authorized Interval on the Raw Variable*) zgodnym z rzeczywistymi (empirycznymi) wartościami minimum i maksimum analizowanego zbioru danych. Przyjęto również 30 rząd rozwinięcia wielomianów Hermite, zalecany, w programie ISATIS, jako najlepszy w większości przypadków. W praktyce optymalne rozwiązanie mieści się zazwyczaj w zakresie 15-50 rzędu (Wackernagel 2003). Możliwe jest iteracyjne poszukiwanie najlepszego wariantu, lecz nie ma to dużego wpływu na jakość transformacji, a w znaczący sposób wydłuża obliczenia. Uzyskane znormalizowane *cdf* (ryc. 137) rozciąga się w przedziale od -3,43 do +3,43. Dopasowanie modelu do danych empirycznych, do granicy około 90 mm, jest prawie idealne. Powyżej nieciągłe, „schodkowe” *cdf* empiryczne zostało przez funkcję anamorfozy wygładzone.

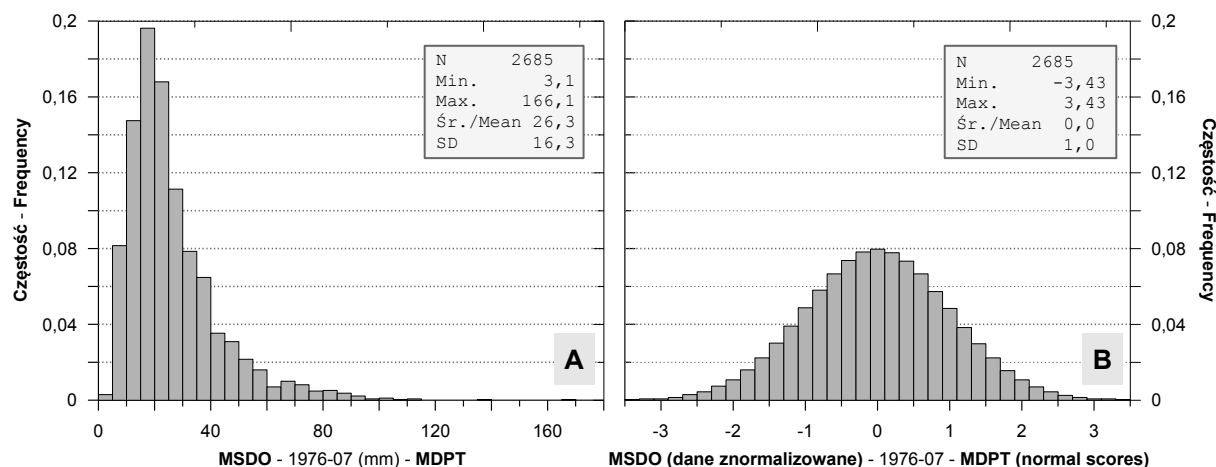
O skuteczności transformacji świadczy nie tylko uzyskanie zbioru danych o średniej równej zero i odchyleniu standardowym równym jeden oraz symetrycznego histogramu (ryc. 138B, tab. 30). Bardziej wymowne są inne charakterystyki ilościowe – skośność, która w wyjściowej próbce MSDO z lipca 1976 wynosiła 1,905 po transformacji równa się dokładnie 0. Kurtoza zmniejszyła się z 5,874 w oryginalnym zbiorze danych do -0,021.

Prawdopodobieństwo normalności rozkładu oszacowane za pomocą testu Shapiro-Wilka wynosiło dla wyjściowych danych mniej niż $2,2^{-16}$, a po ich transformacji – dokładnie 1.



Ryc. 137. Przykład normalizacji danych (ang. *Gaussian anamorphosis*) MSDO z lipca 1976: a – skumulowana empiryczna funkcja rozkładu (*cdf*), b – krzywa rozkładu danych znormalizowanych uzyskana przy użyciu 30 wielomianów Hermita, c – nierealistyczna ekstrapolacja rozkładu poza zasięgiem danych empirycznych, d – granice rozkładu danych empirycznych (3,1-166,1 mm), e – granice przedziału danych znormalizowanych (od -3,43 do 3,43).

Sprawdzono również jakość konwersji odwrotnej, to jest z powrotem z danych znormalizowanych do oryginalnej skali pomiarowej. Statystyki opisowe wszystkich trzech zbiorów danych (oryginalne, znormalizowane, po konwersji zwrotnej) zestawiono w tabeli 30. Wygładzenie górnej części rozkładu empirycznego poskutkowało niewielkim obniżeniem odchylenia standardowego, skośności i kurtozy przy zachowaniu praktycznie idealnym średniej. Maksimum zmalało o 14,2 mm. Przekształcenie do oryginalnej skali jest rzeczywiście nieobciążone – średnie odchylenie tylko minimalnie różni się od zera (-0,003 mm). Średnie absolutne odchylenie (RMSE) wynosi 0,383 mm, a współczynnik korelacji równa się 1.



Ryc. 138. Histogramy i statystyki opisowe MSDO z lipca 1976 roku: A – surowe dane pomiarowe w mm, B – dane znormalizowane według metodyki opisanej w tekście.

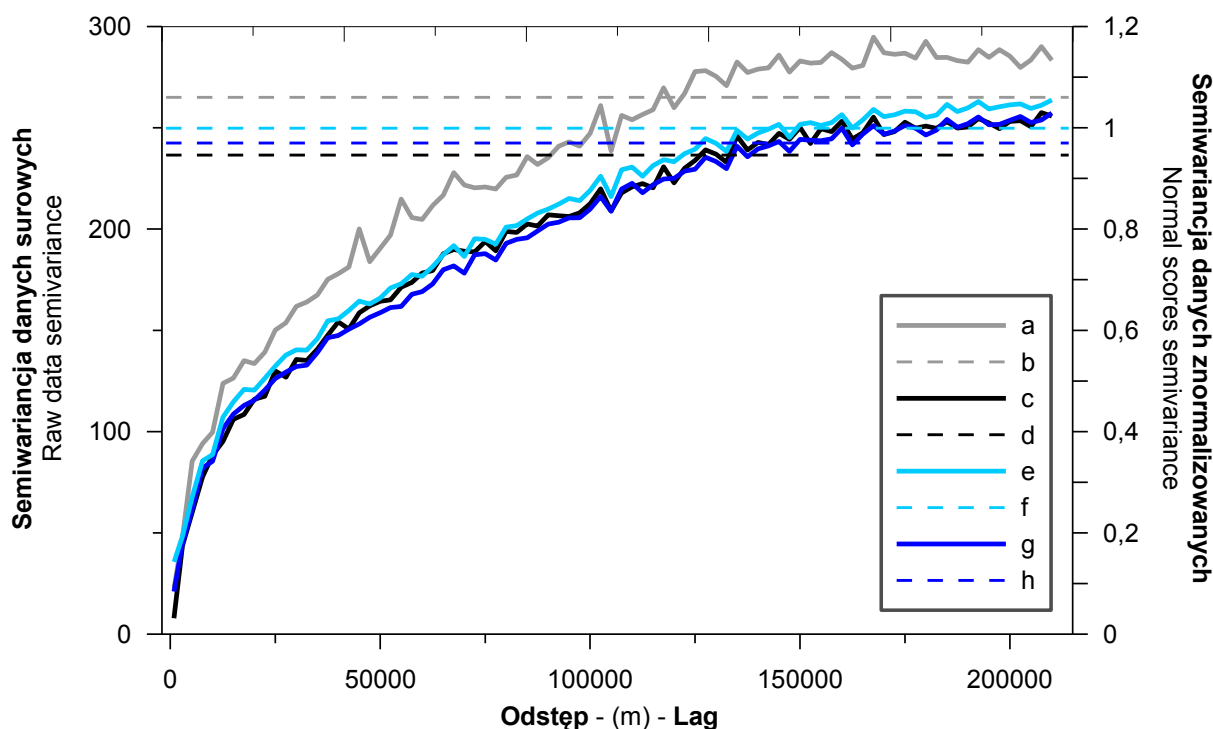
Tabela 30. Statystyki opisowe danych MSDO z lipca 1976: VII-76 - oryginalne dane pomiarowe, VII-76n – dane znormalizowane, VII-76bt – dane po konwersji zwrotnej.

| Parametr – Parametr | Zmienna - Variable | | |
|---|--------------------|---------|----------|
| | VII-76 | VII-76n | VII-76bt |
| Średnia - Mean | 26,269 | 0,000 | 26,266 |
| Błąd standardowy - Standard Error | 0,314 | 0,019 | 0,314 |
| Mediana - Median | 22,2 | 0,0 | 22,0 |
| Odchylenie standardowe - Standard Deviation | 16,283 | 1,000 | 16,251 |
| Kurtoza - Kurtosis | 5,874 | -0,021 | 5,309 |
| Skośność - Skewness | 1,905 | 0,000 | 1,863 |
| Zakres - Range | 163,0 | 6,86 | 148,8 |
| Minimum - Minimum | 3,1 | -3,43 | 3,1 |
| Maksimum - Maximum | 166,1 | 3,43 | 151,9 |

X.4.4. Wpływ normalizacji danych na ocenę ich struktury przestrzennej

Normalizacja danych nie zmienia w najmniejszym stopniu ich struktury przestrzennej reprezentowanej przez semiwariogram. Wręcz przeciwnie, taka transformacja korzystnie zmniejsza czułość semiwariogramu na dane ekstremalne, potencjalnie błędne, lub należące do osobnej populacji. Właściwość ta została zilustrowana na rycinie 139. Przedstawia ona cztery semiwariogramy empiryczne policzone dla MSDO z lipca roku 1976. Krzywe „a” i „e” zostały wyliczone z całego zbioru danych (2685 przypadków), odpowiednio dla danych surowych i znormalizowanych. Krzywe „c” i „g” to z kolei semiwariogramy jak wyżej, ale uzyskane tym razem z danych wyselekcjonowanych (2672 przypadków) po procedurze

interaktywnego czyszczenia (patrz dodatek X.5). Układ obu skal rzędnych (Y) został tak dobrany, aby uzyskać jak najlepszą zgodność przebiegu krzywych „c” i „g”. Takie zestawienie wykresów pokazuje dobitnie, że sama normalizacja całego zbioru danych (krzywa „e”) daje już obraz struktury przestrzennej bardzo zbliżony do uzyskanego po wyeliminowaniu 13 odstających przypadków (krzywe „c” i „g”). Różni się on znacznie od wyjściowego semiwariogramu pełnego zbioru danych surowych (krzywa „a”).



Ryc. 139. Struktura przestrzenna danych MSDO z lipca 1976: a, b – semiwariogram i wariacja całego zbioru danych surowych (2685 przypadków), c, d – semiwariogram i wariacja danych surowych wyselekcjonowanych (2672 przypadków), e, f – semiwariogram i wariacja całego zbioru danych po normalizacji, g, h – semiwariogram i wariacja wyselekcjonowanych danych znormalizowanych.

Zaletą normalizacji danych empirycznych jest także możliwość bezpośredniego porównywania struktury przestrzennej wielu zbiorów danych różniących się skalą pomiarów i/lub zakresem ich zmienności. Ominięcie tego etapu przygotowania danych spowodowałoby konieczność wprowadzenia innej procedury – standaryzacji semiwariogramów empirycznych i ich modeli⁶⁵.

Cenną właściwością anamorfozy gaussowskiej przy użyciu wielomianów Hermite jest możliwość obliczenia, na podstawie współczynników wielomianów (ψ_i), wariacji danych

⁶⁵ Wykonuje się to poprzez podzielenie wartości semiwariogramów empirycznych i ich modeli przez wariację próby.

surowych (Bleinès i in. 2007). Dokonuje się tego na podstawie prostej relacji pomiędzy kowariancją znormalizowanych danych a kowariancją danych surowych:

$$C(\mathbf{h}) = \sum_{i=1}^{\infty} \psi_i^2 \rho^i(\mathbf{h}) \quad [45]$$

gdzie:

- $\rho(\mathbf{h})$ – kowariancja danych znormalizowanych,
- $C(\mathbf{h})$ – kowariancja danych surowych.

Relacja ta jest spełniona wówczas, gdy para zmiennych ($Z(x), Z(x + \mathbf{h})$) może być uważana za dwuzmienną normalną (ang. *bivariate normal*, Goovaerts 1997). Na podstawie zależności istniejącej pomiędzy kowariancją $C(\mathbf{h})$ a semiwariancją $\gamma(\mathbf{h})$:

$$\gamma(\mathbf{h}) = C(0) - C(\mathbf{h}) \quad [46]$$

gdzie $C(0)$ oznacza wariancję danych, a zatem można uzyskać formułę na konwersję semiwariogramów pomiędzy przestrzenią danych znormalizowanych a wyjściową skalą pomiarową. Jest to cenna właściwość w kontekście zademonstrowanej powyżej mniejszej czułości semiwariogramu zmiennej gaussowskiej na wartości odstające i ekstremalne, a co za tym idzie prostszym jego modelowaniu. W niniejszej pracy nie było jednakże potrzeby korzystania z tej opcji. Zachowane parametry anamorfozy gaussowskiej (współczynniki wielomianów Hermite) umożliwiają jednak jej zastosowanie w razie potrzeby w przyszłości.

X.5. Maskowanie danych w trakcie analizy struktury przestrzennej

X.5.1. Wpływ danych ekstremalnych i odstających na określenie struktury przestrzennej zjawiska

Podobnie jak inne statystyki będące pochodną wariancji, semiwariancja jest czuła na wartości ekstremalne (ze względu na operację potęgowania, Kerry, Oliver 2007). W wypadku semiwariancji negatywny wpływ na ustalenie prawidłowości struktury przestrzennej mają także wartości lokalnie odstające⁶⁶. Pojedyncze wyniki pomiarów mogą zniekształcić wynik analizy dla setek, czy tysięcy danych. Jest to szczególnie znaczące dla 1-3 pierwszych odstępów, które zazwyczaj zawierają znacznie mniejszą ilość par punktów niż kolejne. Aby

⁶⁶ Są to takie wartości, które nie stanowiąc ekstremów analizowanej próby, wyraźnie odbiegają od sąsiadujących w przestrzeni i/lub w czasie wyników pomiarów.

uniknąć takich sytuacji, po upewnieniu się, że wartości owe nie są po prostu efektem błędów pomiarowych, stosuje się różne zabiegi (Goovaerts 1997).

Jeśli wartości ekstremalne wykazują sąsiedztwo w przestrzeni, można je uznać za osobną populację i wyłączyć z analizy. Stosuje się również transformację danych (np. logarytmiczną) zmniejszającą skalę zmienności, a przez to redukującą wpływ wartości ekstremalnych. Kolejnym wariantem jest wykorzystanie takich miar struktury przestrzennej mniej czułych na obecność wartości odstających i ekstremalnych, jak madogram czy rodogram (Deutsch, Journel 1998, Goovaerts 1997). Ostatnią z możliwości stanowi tak zwane interaktywne czyszczenie semiwariogramu (Goovaerts 1997, Pannatier 1996). Wykorzystuje się do tego celu wykresy tzw. chmur semiwariogramu (ang. *variogram cloud*) lub rozrzutu z przesunięciem (ang. *h-scatterplots*). Pierwszy z nich przedstawia wszystkie wyliczone wartości kwadratów różnic między parami pomiarów jako funkcję dzielącą je odległości, natomiast drugi – relację między wartościami poszczególnych par pomiarów dla jednego określonego przedziału dzielących je odległości. Procedura polega na wizualnej identyfikacji na wymienionych powyżej wykresach odstających punktów – czyli par pomiarów. Są one następnie przez operatora zaznaczane przy wykorzystywaniu graficznego interfejsu programu komputerowego i maskowane. Maskowanie polega na tymczasowym ich usuwaniu z obliczeń semiwariogramu (przeliczaniu ponownie wartości semiwariogramu z wyłączeniem zamaskowanych wartości). Usunięcie pary danych z obliczeń konkretnej wartości semiwariogramu nie oznacza ich całkowitej eliminacji ze analizowanego zbioru. Te dwa wyniki pomiarów wciąż mogą być użyte do obliczeń w innych parach tego samego przedziału odległości lub innych przedziałów. Tego typu procedura ma liczne zalety, szczególnie przy analizowaniu małych zbiorów danych. W niniejszym opracowaniu dysponowano jednak bardzo licznymi, jak na typową analizę geostatystyczną, zbiorami pomiarów. Dlatego, wykorzystując program ISATIS (Bleinès i in. 2007), czyszczenie semiwariogramu wykonano drugim sposobem (na podstawie chmury semiwariogramu), eliminując całkowicie z obliczeń i modelowania struktury przestrzennej MSDO sprawiające „problemy” dane.

X.5.2. Metodyka procedury interaktywnego czyszczenia semiwariogramu

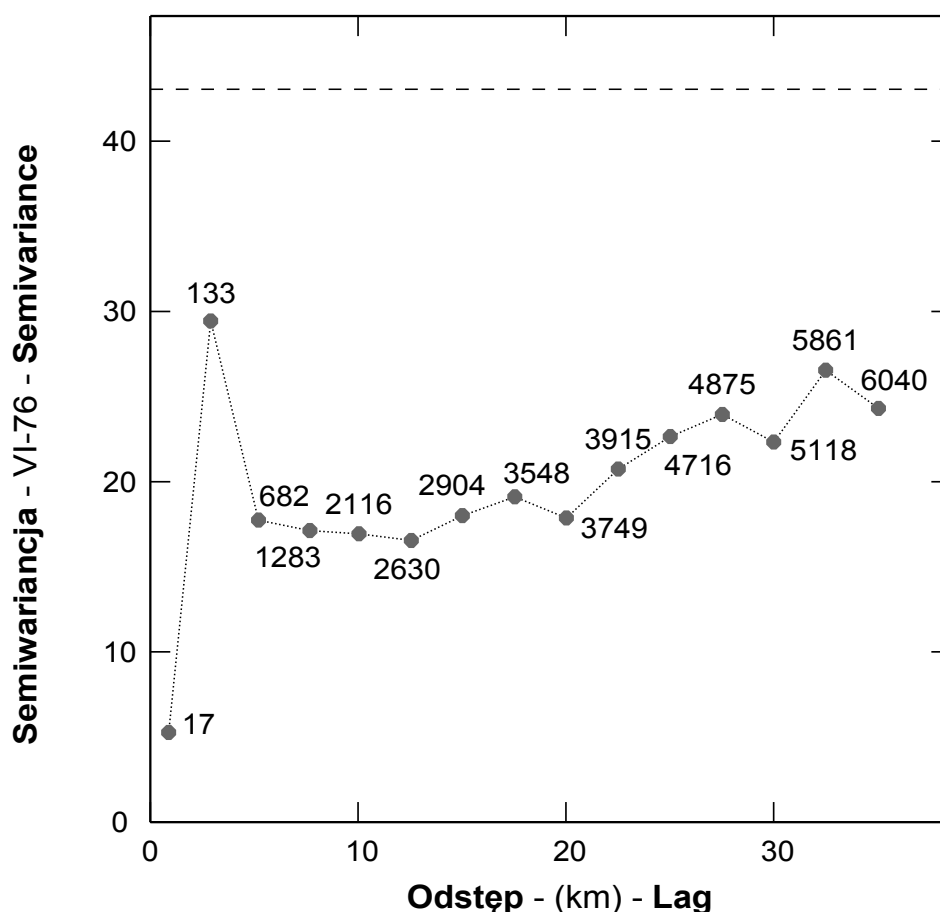
Kolejne etapy zastosowanej procedury interaktywnego czyszczenia semiwariogramu przedstawiono na ryc. 140-142. W przykładzie wykorzystano surowe dane MSDO (w mm)

zanotowane w czerwcu 1976. Dla uproszczenia obliczenia wykonywano jedynie dla 15 odstępów o szerokości 2,5 km.

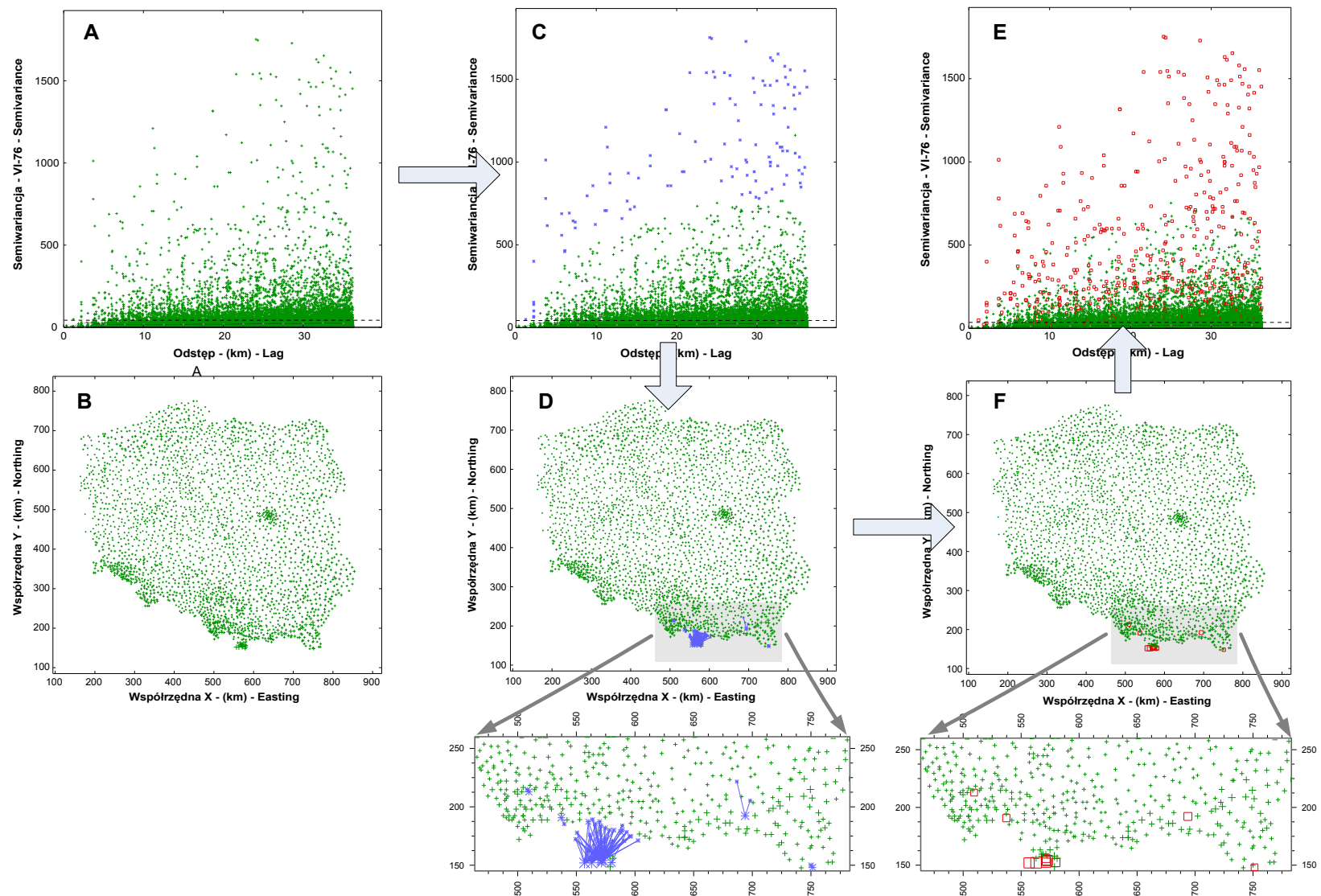
Wykres semiwariogramu MSDO z czerwca 1976 roku wykonany dla wszystkich dostępnych danych, przedstawiono na rycinie 140. Pokazuje ona klarowny przykład struktury przestrzennej zjawiska zaburzonej przez występowanie odstających (być może ekstremalnych) danych. Wartości semiwariancji jako miary niepodobieństwa powinny rosnać wraz ze wzrostem odległości pomiędzy porównywanymi wynikami pomiarów. Tymczasem, na wykresie jest widoczne, że po gwałtownym wzroście w pierwszym odstępnie (od ok. 5^2 do blisko 30^2 mm) wartości semiwariancji dla następnych czterech odstępów (do ok. 12,5 km) zamiast rosnać – maleją. Aby zidentyfikować przyczynę tej anomalii, wykonuje się wykres chmury semiwariogramu (A na ryc. 141). Wyraźnie rysuje się na nim szereg, rozproszonych, odstających od głównej chmury, bardzo wysokich wartości, podniesionych do kwadratu, różnic wyników pomiarów. Identyfikacja miejsc, gdzie znajdują się punkty pomiarowe, na których te opady zostały zmierzone, i jakie konkretnie są to wartości, dokonywana jest za pomocą narzędzia połączonych okien (ang. *linking windows*). Dzięki temu zaznaczone na jednym wykresie lub mapie dane są również widoczne w innym oknie programu zawierającym odmienną ich reprezentację graficzną. W odrębnym oknie wyświetlana jest mapa lokalizacyjna z symbolami proporcjonalnymi do zmierzonych w czerwcu 1976 MSDO (B na ryc.141). Na wykresie chmury semiwariogramu zaznacza się, za pomocą kursora myszki, wszystkie budzące wątpliwość wysokie wartości semiwariancji (niebieskie gwiazdki na części C ryc.141). Wykonanie tej czynności skutkuje natychmiast zaznaczeniem na powiązanej z wykresem mapie wszystkich par lokalizacji, które brały udział w obliczeniach zaznaczonych wartości semiwariancji (połączone liniami niebieskie gwiazdki na części D ryc. 141; widoczne dokładniej na powiększonym fragmencie). Widać dzięki temu wyraźnie, że występują one wyłącznie w Karpatach, a zwłaszcza w Tatrach i na ich przedpolu. Jest to zatem efekt bardzo wysokich opadów orograficznych, które zdecydowanie pod względem relacji przestrzennych „nie pasują” do reszty danych z całej Polski. Dokładna inspekcja mapy w dużym powiększeniu umożliwia precyzyjne zlokalizowanie miejsc, w których wystąpiły anomalne opady, i podjęcie decyzji, które z nich zostaną „zamaskowane”, a zatem usunięte czasowo z dalszych obliczeń.

W wypadku czerwca 1976 roku zdecydowano o maskowaniu danych z 11 stanowisk (F na ryc. 141), co wobec całego, liczącego 2686 lokalizacji, zbioru danych stanowiło nieco ponad 0,4%. Na owe 11 punktów 7 zlokalizowane było w Tatrach. Należały do nich m.in.

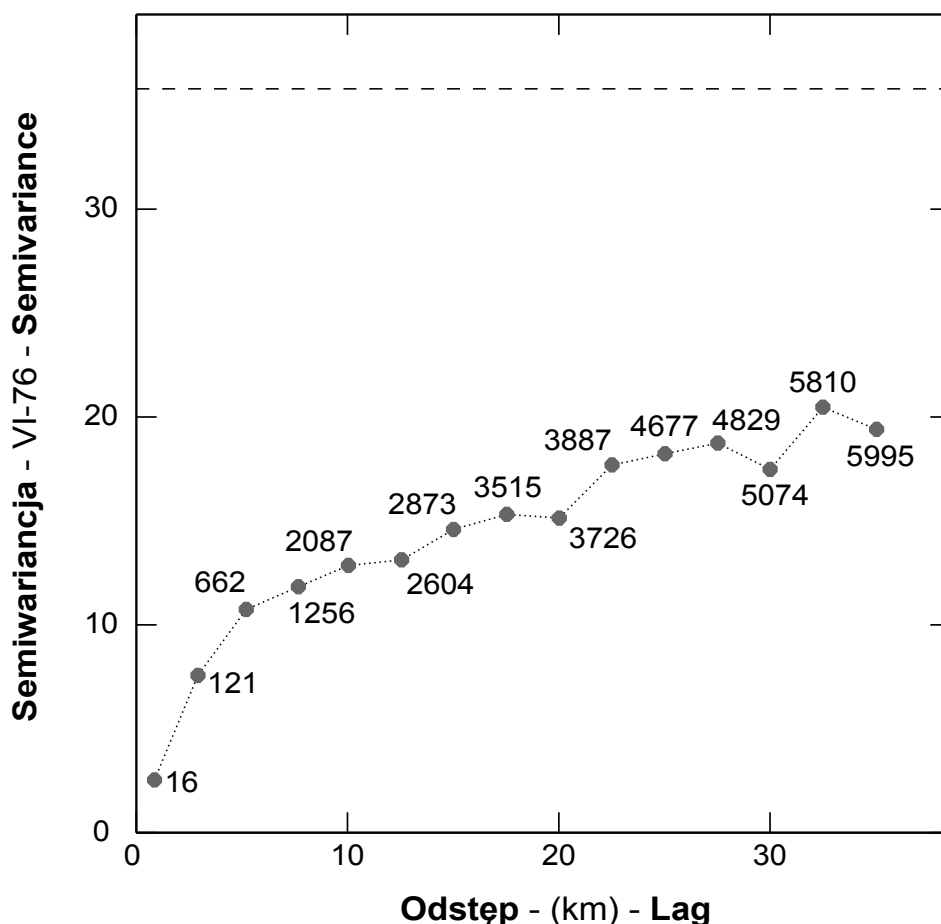
posterunki Hala Ornak, Polana Chochołowska, Hala Gąsienicowa, Kasprowy Wierch, Myślenickie Turnie, gdzie maksymalne opady dobowe przekraczały 60 mm i były najwyższe w tym miesiącu w całej Polsce. Do grupy tej włączono jednak także Babią Górę, Magurkę i Szyndzielnię, oraz Duklę i Wetlinę na których notowano opady o wysokościach wyraźnie niższych (w kolejności od maksymalnego opadu pozycje 16, 20, 50, 10 i 22). O „anomalności” wartości zmierzonego opadu nie decyduje bowiem jego bezwzględna wysokość, tylko relacja względem pomiarów na stacjach sąsiednich. Eliminacja z obliczeń semiwariancji 11 wyjściowych wartości (surowych pomiarów) spowodowało, że liczba porównywanych par danych na badanym dystansie zmalała o 455 z wyjściowych 47 587 (0,96%, E na ryc. 141). Semiwariancje dla kolejnych odstępów zostały obliczone z pominięciem od jednej do maksymalnie 51 par danych (ryc. 142). Redukcja ta zawierała się w przedziale 0,6-9,0%, ale uzyskany skorygowany semiwariogram (ryc. 142) daje klarowny, reprezentatywny dla całej Polski, obraz struktury przestrzennej MSDO z czerwca 1976.



Ryc. 140. Bezkielkowy (izotropowy) semiwariogram empiryczny surowych danych pomiarowych MSDO z czerwca 1976 roku. Obliczono go dla 15 odstępów o szerokości 2,5 km. Przy każdej wartości semiwariancji podano liczbę par danych, z których została wyliczona.



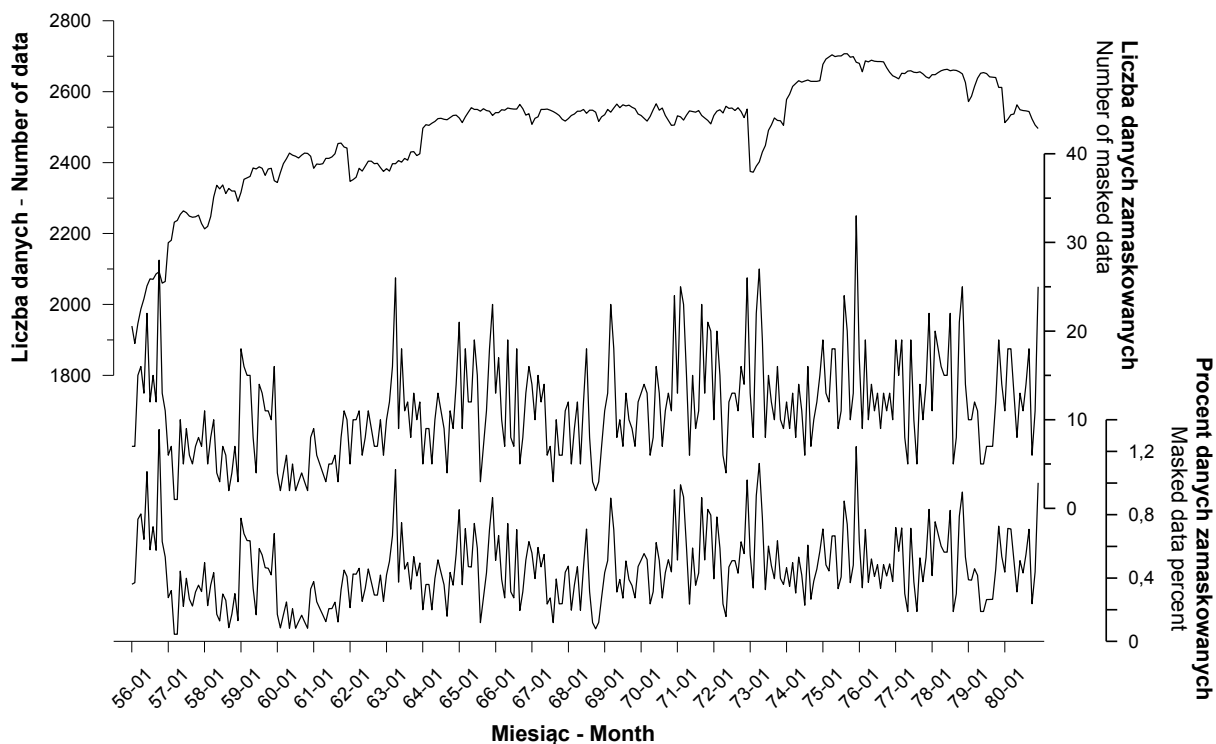
Ryc. 141. Schemat procedury interaktywnego czyszczenia semiwariogramu opartej na wykresie chmury semiwariogramu (A) i mapie lokalizacyjnej z symbolami proporcjonalnymi do wyników pomiarów (B). Szczegółowe omówienie w tekście.



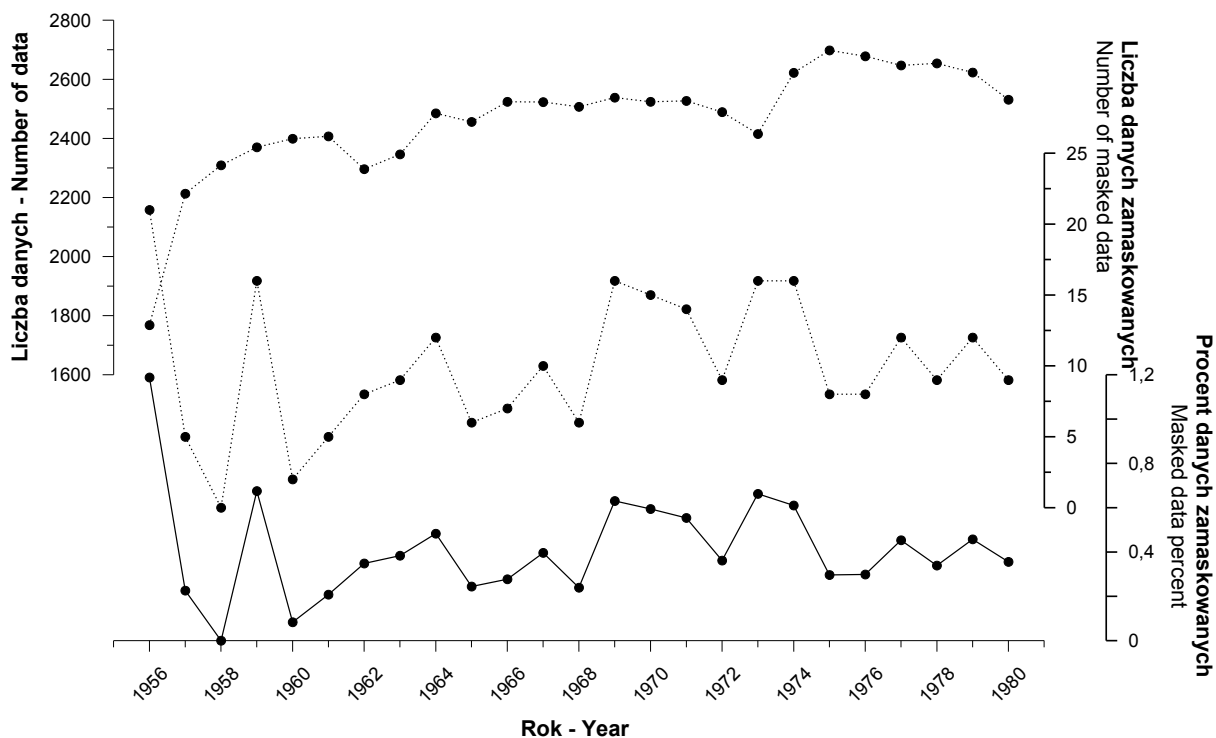
Ryc. 142. Skorygowany bezkierunkowy (izotropowy) semiwariogram empiryczny surowych danych pomiarowych MSDO z czerwca 1976 roku – porównaj z ryciną 140. Przy każdej wartości semiwariancji podano liczbę par danych z których, została wyliczona; porównanie z ryciną 140 daje informację o liczbie usuniętych (zamaskowanych) par.

X.5.3. Statystyki procedury czyszczenia semiwariogramu dla wszystkich analizowanych zbiorów danych MSDO

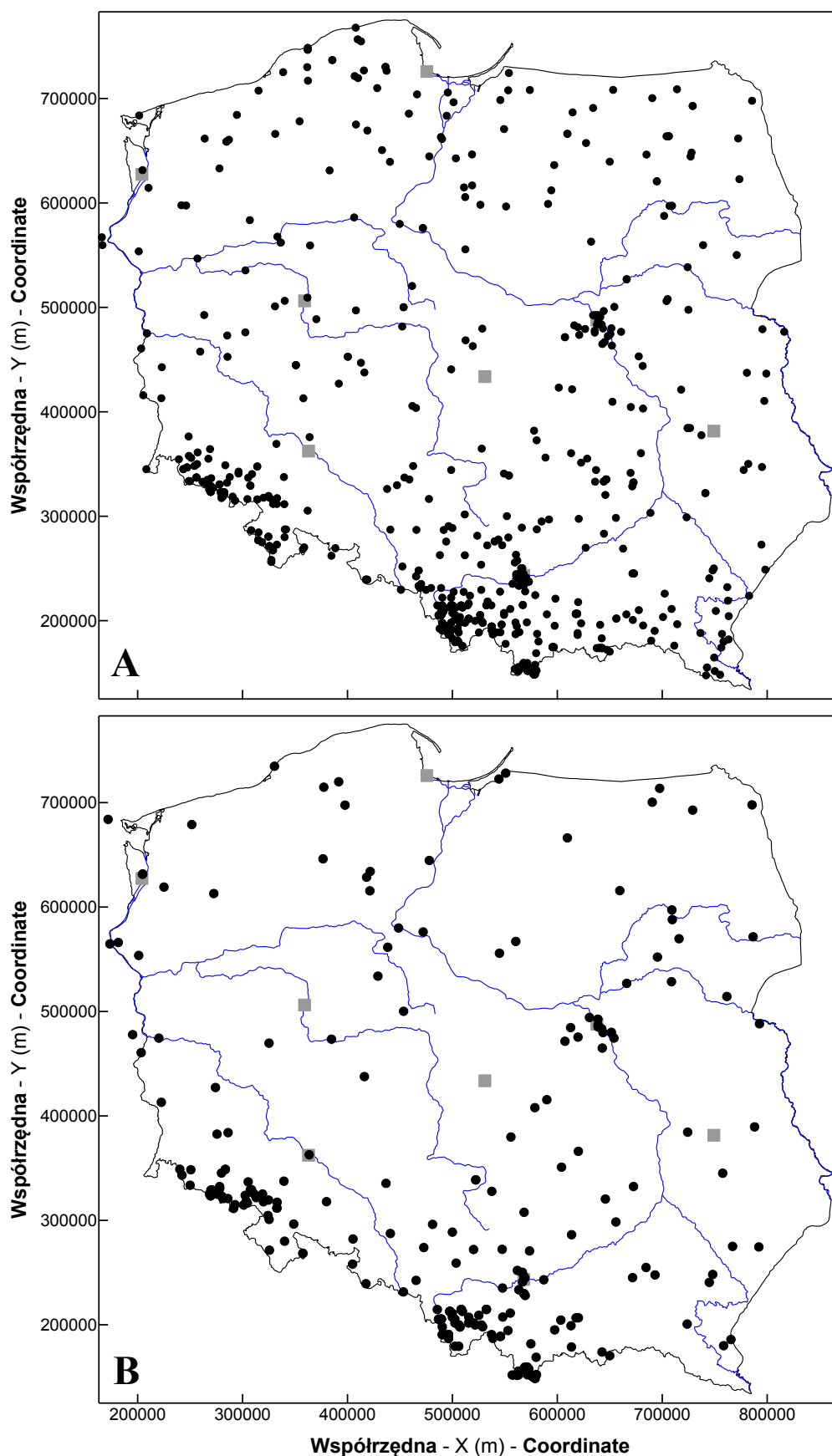
Czyszczenie semiwariogramu było konieczne w wypadku wszystkich analizowanych zbiorów danych oprócz rocznych MSDO z roku 1958 (ryc. 143 i 144). W trakcie obliczania semiwariancji miesięcznych MSDO maskowano 1-33 stanowisk (średnio około 11), co stanowiło 0,05 do 1,34% całości (średnio 0,45%). Te same charakterystyki w odniesieniu do rocznych zbiorów danych wynosiły odpowiednio: 0-21 (średnio ok. 10) i 0-1,19% (średnio 0,4%). W przebiegu czasowym zaznacza się okres lat 1957-1962, gdzie maskowanych danych było nieco mniej. W pozostałej części wielolecia liczba takich przypadków zmienia się dość chaotycznie z miesiąca na miesiąc, bez żadnej wyraźnej sezonowości ani trendu.



Ryc. 143. Liczba i odsetek danych zamaskowanych w trakcie analizy struktury przestrzennej w zestawieniu z całkowitą liczbą danych poszczególnych miesięcznych zbiorów MSDO.



Ryc. 144. Liczba i odsetek danych zamaskowanych w trakcie analizy struktury przestrzennej w zestawieniu z całkowitą liczbą danych poszczególnych rocznych zbiorów MSDO.



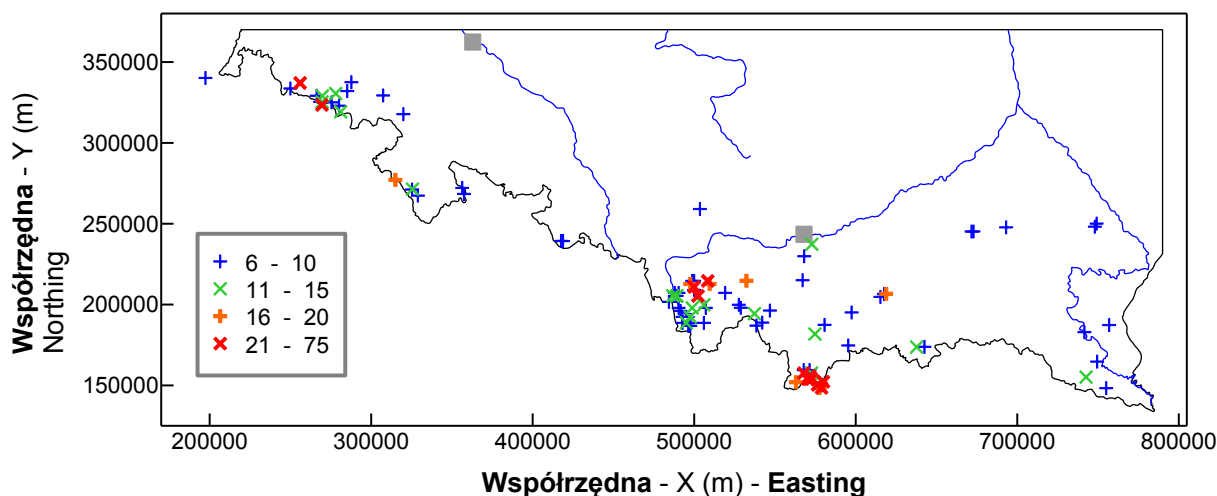
Ryc. 145. Rozmieszczenie posterunków opadowych, na których dokonano pomiarów MSDO zamaskowanych w trakcie analizy struktury przestrzennej: A – zbiory miesięczne, B – zbiory roczne. Dla uproszczenia na mapie A zaznaczono punkty pomiarowe jedynie z 5 wybranych lat: 1956, 1960, 1965, 1970 i 1975.

W przykładzie przedstawionym w poprzednim podrozdziale wartości zaburzające globalny obraz struktury przestrzennej dla całego kraju były ewidentnie związane z lokalnymi opadami orograficznymi. I jakkolwiek w innych zbiorach danych takie sytuacje były bardzo częste, to jednak, anomalne w sensie przestrzennej struktury podobieństwa opady notowane były na terenie całego kraju (ryc. 145). Jakaś ich część, raczej niewielka, to z pewnością wartości błędne, nie zidentyfikowane na etapie archiwizacji danych. Ponieważ jednak w zdecydowanej większości przypadków nie było jednoznacznych podstaw do podjęcia decyzji o ich całkowitej eliminacji z bazy danych, wszystkie wyniki pomiarów wykluczone na etapie opracowania modelu struktury przestrzennej były używane w trakcie obliczeń estymacji i symulacji (por. dodatek X.2).

Zamieszczone mapy (ryc. 145) pokazują, że wykluczone z obliczeń semiwariancji MSDO występowały głównie w Sudetach i Karpatach (zwłaszcza Zachodnich i Tatrach). Na pozostałym obszarze kraju przypadki takich opadów są raczej rozmieszczone losowo, jakkolwiek widać także skupienia związane z miejskimi sieciami pomiarowymi w Krakowie i Warszawie, czy relatywnie nieco większą ich gęstość na Wyżynie Krakowsko-Częstochowskiej oraz w północno-wschodniej części Pojezierza Pomorskiego.

Należy również zwrócić uwagę, że „niepasujące”, ze względu na relacje przestrzenne, MSDO zdarzały się w tych samych lokalizacjach w analizowanym wieloleciu często więcej niż jeden raz (ryc. 146). Na przykład, wśród 251 zakwestionowanych przypadków maksymalnych rocznych sum dobowych było takich sytuacji 26. Sześciokrotnie w danym miejscu zanotowano dwa takie przypadki, trzykrotnie – trzy, a raz – aż pięć (stanowisko Równica w Beskidzie Śląskim – bądź „schronisko”, bądź „wieś”). W zbiorach miesięcznych takie sytuacje powtarzały się zdecydowanie częściej. Na 3399 zamaskowanych wartości MSDO 1274 lokalizacji pojawiło się tylko raz. Można przypuszczać, że wśród nich znaczący odsetek to pojedyncze przypadki zanotowanych tylko w jednej lokalizacji głównie na Nizu, wysokich opadów konwekcyjnych oraz zapewne w jakiejś części pomiary błędne. Pozostałe 2125 usuniętych z analizy struktury przestrzennej powtarzało się w 564 lokalizacjach. W 467 przypadkach powtórzeń było od 2 do 5; 38 razy w jednym stanowisku takich przypadków było więcej niż 10. Maksymalna liczba powtórzeń, bo aż 75, związana jest z pomiarami wykonanymi na Kasprowym Wierchu – najwyższym stałym stanowisku pomiarów opadów atmosferycznych w Polsce (1991 m n.p.m.). Kolejne 4 punkty stanowią następną w kolejności malejącej wysokość: Dolina Pięciu Stawów (1670 m n.p.m. – 53 powtórzenia), Śnieżka (1603 m n.p.m. – 45 powtórzeń), Hala Gąsienicowa (1520 m n.p.m. – 45 powtórzeń), Myślenickie Turnie (1360 m n.p.m. – 39 powtórzeń). Tylko 4 z 97 przypadków, kiedy powtórzeń było więcej niż 5, zlokalizowano na Nizu (ryc. 146).

Pozostałe usytuowane były w Sudetach i Karpatach oraz na ich przedgórzu. Największym zagęszczeniem takich stanowisk charakteryzowały się Karkonosze, Beskid Śląski i Tatry.



Ryc. 146. Rozmieszczenie posterunków opadowych, na których dokonano więcej niż 5 pomiarów MSDO zamaskowanych w trakcie analizy struktury przestrzennej zbiorów miesięcznych. W pozostałej, nie przedstawionej na mapie części Polski, takich lokalizacji było tylko 4: Szczecin – Warszewo (6 razy), Warszawa – Muzeum Narodowe (8), Wyszaków nad Bugiem (7), Klusy na południowo-wschodnim krańcu Krainy Wielkich Jezior Mazurskich (7).

Omówione w niniejszym podrozdziale zagadnienie pozwala spojrzeć na kwestię „ekstremalnych” opadów w trochę innym świetle. Zazwyczaj termin ten rezerwuje się dla zdarzeń, w trakcie których zanotowano bardzo wysokie sumy lub intensywności, których częstość występowania jest bardzo niska. Decyduje tu zatem kryterium czasowe. Z przeprowadzonej analizy wynika jednakże, że notowane są w Polsce opady o bardzo ograniczonym zasięgu przestrzennym, niekoniecznie skrajnie wysokie, które są wyraźnymi anomaliami z punktu widzenia struktury globalnego ich pola. Zagadnienie to zarówno w aspekcie czasowym, jak i przestrzennym będzie tematem osobnego opracowania.

X.6. Rozdzielczość modeli przestrzennych

X.6.1. Ogólne kryteria wyboru rozdzielczości rastrowych modeli przestrzennych

Analizy GIS, których efektem są rastrowe modele zmienności przestrzennej wymagają na etapie planowania metodyki podjęcia decyzji o ich rozdzielczości, czyli inaczej mówiąc – o wymiarach oczka siatki (ang. *grid*). Bardzo często problem ten jest pomijany milczeniem, a decyzja – podejmowana bez uwzględniania jakichkolwiek obiektywnych kryteriów. Jasne jest, że najważniejsze z nich muszą dotyczyć zbioru danych, na podstawie którego budowany jest model. Decydujące znaczenie ma typ próbkowania (punktowe bądź obszarowe; losowe, regularne, preferencyjne, profilowe itp.) i

zagęszczenie danych (liczba na jednostkę powierzchni) oraz charakterystyki zmienności przestrzennej analizowanej cechy (statystyki lokalne, a w tym także parametry autokorelacji). Powinno się również uwzględnić rozkład statystyczny błędów zarówno samych pomiarów, jak i określenia lokalizacji stanowisk pomiarowych. Można stwierdzić intuicyjnie, że tworzenie modelu o rozdzielczości 100 m, kiedy lokalizacje danych są od siebie odległe o kilkadziesiąt kilometrów, ma niewielki sens.

Nie można również pominąć aspektów technicznych przy podejmowaniu decyzji o rozdzielczości modelu. Pierwszy z nich związany jest ze skalą docelowej mapy. Te dwa elementy powinny być ze sobą zharmonizowane, a skala tak dobrana, aby można było przy danej rozdzielczości, zobaczyć wszystkie istotne szczegóły modelu. I odwrotnie, nie ma sensu tworzyć modelu o wysokiej rozdzielczości, jeśli przy założonej z góry skali mapy finalny efekt będzie musiał ulec generalizacji. Drugi aspekt techniczny to ograniczenia pojemności pamięci i szybkości przetwarzania wykorzystywanego komputera. Parametry te zmieniają się z miesiąca na miesiąc, ale wciąż nie można ich ignorować. W tym momencie (styczeń 2008) górną granicę⁶⁷ dla przeciętnego komputera osobistego o pamięci rzędu 1 GB stanowi dla obszarów o powierzchni Polski rozdzielczość rzędu 10-50 m ($x \cdot 10^8$ - $x \cdot 10^9$ węzłów).

Trzecim elementem który należy wziąć pod uwagę to przeznaczenie modelu. Jeśli ma posłużyć jedynie do celów naukowych – charakterystyki i wyjaśniania zmienności przestrzennej zjawiska – zazwyczaj ważne są raczej proporcje i relacje niż bezwzględne wartości. Często też ignoruje się wówczas, przy analizach obejmujących cały kraj, zmienność zachodzącą w małej skali przestrzennej. Zastosowania praktyczne stawiają w wielu aspektach wyższe wymagania dotyczące rozdzielczości, precyzji i dokładności modelu. Różnią się one oczywiście znacznie w zależności od celu. Inne muszą być przy obliczeniach bilansów wodnych zlewni, inne przy projektowaniu sieci kanalizacji burzowej w mieście itp.

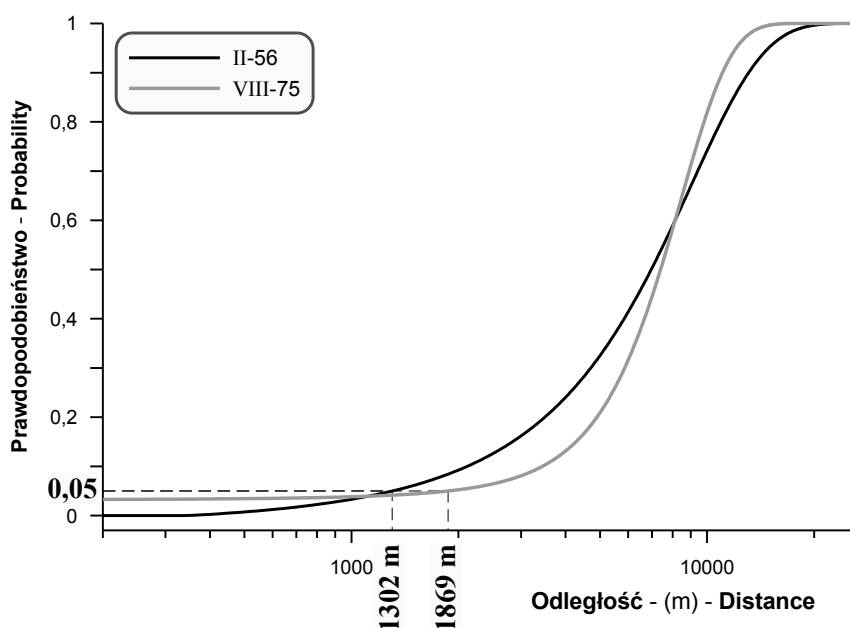
X.6.2. Ocena optymalnej rozdzielczości modeli z wykorzystaniem charakterystyk analizowanego zbioru danych

Przy konstruowaniu modeli przestrzennych maksymalnych sum dobowych opadów na obszarze Polski, tworzonych w niniejszej rozprawie, stosowano zalecenia zawarte w publikacji Hengla (2006), a dotyczące optymalizacji ich rozdzielczości. Nie są to jakieś sztywne reguły pozwalające w sposób jednoznaczny określić optymalną wartość. Autor cytowanej pracy zestawiał dla kryteriów związanych z właściwościami analizowanego zbioru danych, a także z docelową skalą mapy, formuły pozwalające

⁶⁷ Spore znaczenie ma optymalizacja algorytmów – czyli oprogramowanie.

oszacować optymalny zakres rozdzielczości, a także jej wartość zalecaną. Przy każdym z kryteriów są to jednakże nieco inne liczby.

Ocenę zaczęto od parametru najmniej ważnego – docelowej skali map. Niniejsza publikacja wydrukowana jest w formacie B5. Oznacza to, że maksymalny wymiar mapy⁶⁸, biorąc pod uwagę kształt Polski, może na kierunku W-E wynosić 12,5 cm. Żeby przedstawić całe terytorium należy zastosować skalę około 1 : 6 000 000. Według podanych przez Hengla (2006) i omówionych szczegółowo reguł, optymalna rozdzielczość dla takiej skali powinna się zawierać w przedziale od 600 do 15 000 m (od $SM \times 0,0001$ do $SM \times 0,0025$). Wartość zalecana wynosi zaś 3 000 m ($SM \times 0,0005$).



Ryc. 147. Krzywe prawdopodobieństwa znalezienia jednego punktu danych w określonym promieniu (odległości) od innego punktu dla miesięcy z najmniejszą (luty 1956) i największą (sierpień 1975, szczegóły w rozdz. III) liczbą stanowisk. Dla obu zbiorów danych zaznaczono odległość 95% prawdopodobieństwa, że ów punkt położony jest dalej.

Drugie kryterium związane jest gęstością danych (ang. *inspection density*), to jest ich liczbą (N) w stosunku do powierzchni analizowanego obszaru (A), niezależnie od ich rozkładu przestrzennego. Zgodnie z informacjami podanymi w rozdziale V, dysponowano dla analizowanego wielolecia średnio 2492 wynikami pomiarów miesięcznych MSDO i 2478 – rocznych. Ponieważ liczby te różnią się niewiele, zdecydowano że w obliczeniach użyta zostanie wartość 2480. Tworzone modele obejmowały powierzchnię⁶⁹ 319 114 km². Według Hengla (2006) zakres rozdzielczości (p)⁷⁰ tworzonego modelu

⁶⁸ Wykluczono, ze względu na ilość map i wykorzystanie koloru, opcję druku wkładek lub załączników o większym formacie.

⁶⁹ Terytorium państwa plus Zalew Szczeciński, Zatoka Pucka i Zalew Wiślany oraz strefa o szerokości 2 km wokół granic.

⁷⁰ W niniejszym podrozdziale zastosowano notację używaną przez Hengla (2006) co oznacza, że niektóre symbole literowe mają inne znaczenie niż w pozostałej części pracy. Są one jednakże w tekście dokładnie objaśnione.

powinien mieścić się w przedziale $\left(0,05 \cdot \sqrt{\frac{A}{N}} \leq p \leq 0,1 \cdot \sqrt{\frac{A}{N}}\right)$, a zalecana wartość powinna wynosić

$p = 0,0791 \cdot \sqrt{\frac{A}{N}}$. Użycie podanych powyżej danych we wzorach daje zakres 567-1134 m i wartość zalecaną 897 m.

Kryterium zagęszczenia danych jest bardzo proste, bo nie uwzględnia w ogóle ich rozkładu przestrzennego. Bardziej wyrafinowana metoda oceny optymalnej rozdzielczości modelu opiera się na charakterystykach statystycznych rozkładu punktów (ang. *point pattern analysis*). W zależności od generalnego typu rozmieszczenia danych – regularnego lub losowego – używa się nieco innych formuł. Obliczenia, których wyniki przedstawiono w podrozdziale V.4 dają podstawy do stwierdzenia, że przez cały okres analizowanego wielolecia, niezależnie od zmiany liczby i rozmieszczenia stanowisk pomiarowych, rozkład ich może być uważany za losowy (wartości wskaźnika Clarka-Evansa oscylujące wokół 1). W tej sytuacji potrzebne są dwie charakterystyki statystyczne analizowanego zbioru danych (Hengl 2006): średnia odległość do najbliższego sąsiada (\bar{h}_{ij}) i odległość od dowolnego punktu danych, w której z prawdopodobieństwem 5% znaleziony zostanie inny punkt ($h_{ij(p=5\%)}$). Odpowiednie wzory wyznaczające zakres optymalnych rozdzielczości modelu są wówczas następujące:

$$h_{ij(p=5\%)} \leq p \leq \frac{\bar{h}_{ij}}{2}, \quad [47]$$

a wartość zalecaną określa się z poniższej formuły:

$$p = \frac{\sqrt{\frac{A}{N}}}{4} \quad [48]$$

Średnia odległość do najbliższego sąsiada (\bar{h}_{ij}) oscylowała w analizowanych zbiorach danych MSDO w zakresie 6,8-7,8 km, wynosząc przeciętnie 7,3 km (por. podrozdz. V.4). Drugi potrzebny parametr ($h_{ij(p=5\%)}$) określono tylko dla miesięcy o skrajnej ilości i rozkładzie stanowisk pomiarowych (por. podroz. V.4): lutego 1956 roku i sierpnia 1975 (ryc. 147). Wynosił on odpowiednio 1302 i 1869 m. Tak więc, biorąc, z przytoczonych powyżej, skrajne wartości, optymalny zakres rozdzielczości siatek estymacji i symulacji powinien, ze względu na rozkład przestrzenny danych zawierać się w przedziale 1302 m - 3900 m, a zalecany rozmiar – 2836 m.

Ostatnie kryterium dotyczy zasięgu autokorelacji danych – odległości podobieństwa wyników pomiarów (ang. *spatial dependence structure*). Daje ono możliwość oceny ich nadmiarowości (redundancji w języku teorii informacji), a poprzez to ocenę skali dokładności interpolacji. Do obliczeń potrzebny jest zasięg autokorelacji (h_R), liczba (m) par punktów danych znajdujących się w jego zasięgu, a także zdefiniowany powyżej parametr rozkładu przestrzennego stanowisk pomiarowych $h_{ij(p=5\%)}$. Zakres optymalnych rozdzielczości modelu określa się przy użyciu następujących wzorów:

$$h_{ij(p=5\%)} \leq p \leq \frac{h_R}{2}, \quad [49]$$

a wartość zalecaną określa się z poniższej formuły:

$$p = h_R \cdot m^{-\frac{1}{3}} \quad [50]$$

Określenie zakresu p w tym przypadku nastęczyło nieco trudności. Struktura przestrzenna dominującej większości analizowanych zbiorów miesięcznych i rocznych MSDO jest złożona i bardzo zmienna (patrz rozdz. VI). Ponadto, w ponad 63% miesięcy i 44% lat definitywny zasięg autokorelacji nie został określony – przekraczał maksymalny uwzględniony w analizie odstęp danych, czyli 212,5 km. Zdecydowano zatem przyjąć do obliczeń zasięg autokorelacji najczęściej występującej struktury krótkodystansowej istniejącej w 285 analizowanych zborach miesięcznych i wszystkich (25) – zbiorach rocznych. Wynosiła ona średnio 15 500 m (8 000-40 000 m, SD = 6 370 m). Taką i mniejszą odległość między sobą miały 7770-9690 par punktów danych. Podstawienie do wyżej podanych wzorów średnich wartości obu potrzebnych parametrów ($h_R = 15\,500$ i $m = 8700$ oraz $h_{ij(p=5\%)} = 1560$ m) dało przedział 1560-7750 m i wartość zalecaną 753 m. Jest to oczywiście wynik sprzeczny – wartość zalecana znajduje się poza zakresem optymalnym. Wynika on z omówionych powyżej trudności w parametryzacji obliczeń według kryterium zasięgu autokorelacji.

Zestawienie zalecanych rozdzielczości modeli rastrowych estymacji i symulacji danych MSDO uzyskanych dla czterech kryteriów – skali mapy, zagęszczenia danych, ich rozkładu przestrzennego i zasięgu autokorelacji – daje odpowiednio wartości około 3000, 900, 2840 i 750 m. Wszystkie, przy objęciu analizą całego terytorium kraju, są w zakresie aktualnych możliwości przetwarzania i gromadzenia danych. Biorąc pod uwagę prostotę struktury modelu, a także zakres błędów lokalizacji stanowisk pomiarowych wynoszącą od około 600 do 900 m (por. podrozdz. V.2), zdecydowano się na przyjęcie konsekwentnie we wszystkich przypadkach kilometrowej rozdzielczości (1×1 km) tworzonych map rastrowych.