

2022

Analiza zmienności temperatur z użyciem metod interpolacji przestrzennej – walidacja podzbiorem

PROJEKT WYKONANY W OPROGRAMOWANIU ARCGIS PRO
MONIKA MYSZKOWSKA

Spis treści

1. Wprowadzenie	3
2. Podstawowe pojęcia i wskaźniki geostatystyczne	5
a) Histogram	5
b) Wykres rozkładu normalnego	6
c) Semiwariogram	6
d) Deterministyczne i stochastyczne metody interpolacji przestrzennej	8
3. Część praktyczna – interpolacja przestrzenna	11
3.1. Walidacja z wykorzystaniem podzbioru 10%	12
a) Empirical Bayesian Kriging	16
b) Kriging Ordinary	17
c) Kriging Simple	18
d) Kriging Universal	19
e) Kriging Disjunctive	20
f) Inverse Distance Weighting.....	21
g) Radial Basis Functions	22
3.2. Walidacja z wykorzystaniem podzbioru 15%	23
a) Empirical Bayesian Kriging	27
b) Kriging Ordinary	28
c) Kriging Simple	29
d) Kriging Universal	30
e) Kriging Disjunctive	31
f) Inverse Distance Weighting.....	32
g) Radial Basis Functions	33
3.3. Walidacja z wykorzystaniem podzbioru 20%	34
a) Empirical Bayesian Kriging	38
b) Kriging Ordinary	39
c) Kriging Simple	40
d) Kriging Universal	41
e) Kriging Disjunctive	42
f) Inverse Distance Weighting.....	43
g) Radial Basis Functions	44
4. Zestawienie metod wykorzystujących podzbiory	45
4.1. Empirical Bayesian Kriging	45

4.2.	Kriging Ordinary	46
4.3.	Kriging Simple	47
4.4.	Kriging Universal	48
4.5.	Kriging Disjunctive	49
4.6.	Inverse Distance Weighting.....	50
4.7.	Radial Basis Functions	51
5.	Wybór optymalnej metody interpolacji	52
6.	Podsumowanie i wnioski	54
	Literatura.....	56
	Źródła internetowe	57
	Spis tabel	57
	Spis rysunków.....	58

1. Wprowadzenie

Celem badania było wskazanie optymalnej metody interpolacji przestrzennej na podstawie wartości średniej kwadratowej (RMSE, ang. Root Mean Square Error) oraz analiza zmienności temperatury powietrza z użyciem metod interpolacji przestrzennej. Wybór optymalnej metody opierał się na analizie wykorzystującej walidację podzbiorom. W ramach analizy wydzielono podzbiory 10%, 15% oraz 20% z danych punktowych. W analizie wykorzystano dane pomiarowe temperatury powietrza wykonane w stacjach meteorologicznych na całym świecie (źródło: National Oceanic and Atmospheric Administration, NOAA). Analizy dokonano dla danych punktowych położonych w Australii i Oceanii. Wykorzystane dane są reprezentatywne, niezależne i pozyskane z tego samego źródła. Mapy i parametry oceny opracowano z wykorzystaniem oprogramowania ArcGIS Pro. Procedura badawcza opierała się na następujących etapach:

1. Wstępna analiza danych,
2. Wydzielenie podzbiorów 10%, 15% i 20%,
3. Obliczenie wskaźników: histogram, wykres rozkładu normalnego, semiwariogram (dla pozostałych 90%, 85% i 80% całości),
4. Analiza zmienności temperatur z użyciem metod interpolacji przestrzennej dla zbiorów 90%, 85%, 80%,
5. Przeprowadzenie walidacji podzbiorom wykorzystując wydzielone podzbiory
6. Wyliczenie wartości wskaźnika RMSE dla poszczególnych metod interpolacji przestrzennej
7. Porównanie wyników i wybór optymalnej metody geostatystycznej dla każdego ze zbiorów.

Walidacja podzbiorom polega na podziale zbioru danych na dwa podzbiory, z których jeden jest testowym. Zwykle zbiór testowy stanowi mniej niż 1/3 próby. Zbiór główny służy do stworzenia semiwariogramu empirycznego, zbudowania modelu oraz estymacji wartości. Następnie wynik estymacji porównywany jest z rzeczywistymi wartościami ze zbioru testowego. Zaletą tego podejścia jest stosowanie danych niezależnych od estymacji do oceny jakości modelu. Wadą natomiast jest konieczność posiadania (relatywnie) dużego zbioru danych.¹

¹ <https://www.wbc.poznan.pl/Content/382516/ocena-jakosci-estymacji.html>

Wartość wskaźnika RMSE dla poszczególnych metod obliczono wykorzystując wzór:

$$RMSE = \sqrt{MSE}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - \tilde{x})^2$$

Większość estymacji i symulacji prowadzonych metodami geostatystycznymi ma charakter lokalny, oznacza to, że przypisanie nowej wartości w miejscu niepróbkowanym odbywa się na podstawie podzbioru punktów w pewnym „lokalnym otoczeniu” (URBAŃSKI, 2012). W badaniu wykorzystano metody takie jak:

- Empirical Bayesian Kriging (EBK),
- Kriging Ordinary,
- Kriging Simple,
- Kriging Universal,
- Kriging Disjunctive,
- Inverse Distance Weighting (IDW),
- Radial Basis Functions (RBF).

W wizualizacji interpolacji przestrzennej punktów we wszystkich badanych metodach wykorzystano 14 przedziałów klasowych.

Analiza ograniczała się do obszaru Australii i Oceanii. Australia to kraj leżący w trzech strefach klimatycznych – strefa podrównikowa, zwrotnikowa i podzwrotnikowa. Na wschodzie kontynentu znajdują się masywy górskie, na zachodzie rozciągają się rozległe wyżyny. Pomiędzy wyżynami i górami rozciągają się niziny a pod nimi znajduje się Wielki Basen Artezyjski. Na obszarze Australii występuje 6 dużych obszarów pustynnych lub półpustynnych. Na kształtowanie się warunków klimatycznych Australii wywiera wpływ położenie geograficzne, zawartość lądu i nieznaczne urozmaicenie ukształtowania powierzchni, prądy morskie i układ stałych ośrodków barycznych. Temperatura powietrza Australii wykazuje znaczny związek z szerokością geograficzną.

2. Podstawowe pojęcia i wskaźniki geostatystyczne

„Interpolacja przestrzenna to procedura szacowania wartości cechy w niepróbowanych punktach na obszarze objętym istniejącymi pomiarami” (Waters, 1989). Interpolacja przestrzenna pozwala na oszacowanie wielkości analizowanej zmiennej w dowolnym punkcie obszaru badań (SZYMANOWSKI, KRYZA, 2011). Termin estymacja stosowany jest jako synonim interpolacji przestrzennej, szczególnie w przypadku stosowania metod geostatystycznych. Pojęcie interpolacji przestrzennej, jako metody obejmującej szacowanie (estymację) czy też przybliżanie (aproksymację) wartości zmiennej, zarówno wewnętrz (interpolacja), jak i na zewnątrz (ekstrapolacja) obszaru badań jest powszechnie akceptowane w analizie przestrzennej i GIS. Głównym celem interpolacji jest otrzymanie ciągłej przestrzennie informacji o wartości zmiennej oraz zapewnienie możliwości jej graficznej prezentacji w formie mapy (SZYMANOWSKI I IN. 2020 ZA: LAM, 1983; MITAS, MITASOVA, 1999; DUBOIS, GALMARINI, 2004).

a) Histogram

Histogram stanowi graficzny sposób przedstawienia rozkładu liczebności wybranych zmiennych. Służy on do analizy zbiorowości, co oznacza, że jesteśmy w stanie odpowiedzieć między innymi na pytania: który przedział jest najliczniejszy (określenie dominanty), czy rozkład jest symetryczny czy może asymetryczny/skośny, czy występują przedziały mocno odstające od reszty. Skośność i kurtoza wskazują na rozkład normalny na histogramie. Przystępując do interpolacji, należy upewnić się czy rozkład danych jest zbliżony do normalnego.²

Skośność jest miarą asymetrii obserwowanych wyników. Informuje o tym jak wyniki dla danej zmiennej kształtuje się wokół średniej. Czy większość zaobserwowanych wyników jest z lewej strony średniej, blisko wartości średniej czy z prawej strony średniej? Innymi słowy, czy w analizowanym zbiorze obserwacji więcej jest wyników, które są niższe niż średnia dla całej grupy, wyższe czy równe średniej?³ Jeżeli rozkład wokół średniej jest symetryczny to skośność przyjmuje wartość zero. Rozkład zwiększający się asymetrycznie po prawej stronie średniej oznacza, że skośność przyjmuje wartość dodatnią.

² <https://www.statystyka-zadania.pl/histogram>

³ https://www.naukowiec.org/wiedza/statystyka/skosnosc_714.html

Kurtoza jest miarą koncentracji wyników. Kurtoza informuje o tym, na ile obserwacje, wyniki są skoncentrowane wokół średniej. Miara ta informuje jak dużo wyników/obserwacji jest zbliżonych do wartości średniej, czy większość z zaobserwowanych wyników ma wartość podobną do średniej.⁴ Im wyższa kurtoza tym większe skupienie zbiorowości wokół wartości średniej, co daje wyraz w większej smukłości krzywej rozkładu. Mała jej wartość daje efekt odwrotny czyli większy rozrzut wartości, słabą koncentrację i, co za tym idzie, spłaszczenie krzywej liczebności. Dla rozkładu normalnego przyjmuje się wartość kurtozy równą 3, dla wartości większych od 3 rozkład jest bardziej wysmukły, a dla wartości mniejszych bardziej spłaszczony.⁵

b) Wykres rozkładu normalnego

Rozkład normalny jest jednym z najważniejszych rozkładów teoretycznych prawdopodobieństwa. Pomiary wartości temperatury powietrza zostały poddane badaniu czy rozkład jest normalny (Gaussowski), czy nie. W ramach badania wykorzystano technikę tworzenia wykresu QQ (qqplot). Interpretacja jest następująca: jeżeli punkty wykresu leżą blisko prostej i są równomiernie rozłożone po jej jednej i drugiej stronie (np. naprzemiennie), to dane pochodzą z rozkładu normalnego.⁶ Im bardziej obserwacje zbliżone są do linii rozkładu normalnego, tym bardziej rozkład analizowanego zbioru danych zbliżony jest do rozkładu normalnego.

c) Semiwariogram

Semiwariogram to podstawowe narzędzie służące do estymacji i badania struktury zmienności badanych zjawisk w geostatystyce. Określa zależność między średnim zróżnicowaniem wartości wysokości (semivariancją) i odległością między punktami pomiaru (MARMOL 2002). Semiwariogram często w literaturze nazywany jest zamiennie wariogramem, jednak wariogram jest jej dwukrotnością. Semiwariogram jest miarą definiowaną jako połowa średniej kwadratowej różnicy między dwiema wartościami cechy mierzalnej, pomiędzy którymi odległość równa jest w przybliżeniu wektorowi h . Semiwariogram jest preferowaną

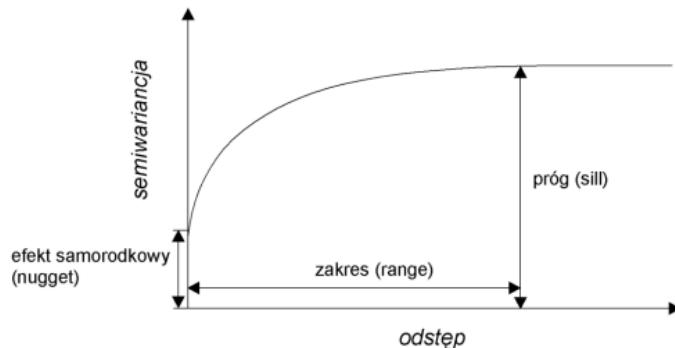
⁴ https://www.naukowiec.org/wiedza/statystyka/kurtoza_698.html

⁵ <https://mfiles.pl/pl/index.php/Kurtoza>

⁶ <https://idane.pl/minipodrecznik/testy-rozkladu.html>

miarą relacji przestrzennej, ponieważ wykazuje tendencję do lepszego wygładzania danych niż funkcja kowariancji. Semiwariogram składa się z trzech podstawowych elementów, są to⁷:

- nugget - efekt nuggetowy - pozwala na określenie błędu w danych wejściowych oraz zmienności na dystansie krótszym niż pierwszy odstęp,
- sill - semiwariancja progowa - oznacza wariancję badanej zmiennej,
- range - zasięg - to odległość do której istnieje przestrzenna korelacja.



Rysunek 1 Podstawowe pojęcia w analizie wariogramu

Źródło: Urbański, J. (2008). GIS w badaniach przyrodniczych. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego

Mapa semiwariogramu (zwana inaczej powierzchnią semiwariogramu) służy do określenia czy struktura przestrzenna zjawiska posiada anizotropię, a jeżeli tak to w jakim kierunku. Na podstawie mapy semiwariogramu identyfikuje się także parametry potrzebne do zbudowania semiwariogramów kierunkowych.⁸ Mapa pozwala na określenie liczby sektorów.

Analizę semiwariancji należy powiązać z tzw. „Pierwszym Prawem Geografii Toblera”, który w 1970 roku stwierdził, że w „przestrzeni wszystko jest związane ze wszystkim innym, przy czym bliższe rzeczy są bardziej związane niż rzeczy odległe”. W związku z tym zależność przestrzenna (autokorelacja) występuje wtedy, gdy badane zjawiska w jednej jednostce przestrzennej powodują zwiększenie się lub zmniejszanie prawdopodobieństwa występowania tych zjawisk w sąsiednich jednostkach. Innymi słowy celem analizy przestrzennej (spatial analysis) jest wydobycie informacji o zależności między danymi oraz interakcjach pomiędzy wartościami badanych zmiennych (KRASOWSKA, BANASZUK 2013 ZA MILLER 2004, BIVAND 1980, ZAWADZKI 2011).

⁷ <https://www.wbc.poznan.pl/Content/382516/geostatystyka-prolog.html>

⁸ <https://www.wbc.poznan.pl/Content/382516/geostatystyka-prolog.html>

d) Deterministyczne i stochastyczne metody interpolacji przestrzennej

W ramach badania wykorzystano metody interpolacji przestrzennej takie jak:

1. Deterministyczne (modelują powierzchnię w sposób jednoznacznie określonymi funkcjami matematycznymi):

- Inverse Distance Weighting (IDW) - metoda odwrotnych odległości

Metoda średniej ważonej odlegością (IDW) wylicza wartość dla każdej komórki na podstawie wartości punktów obokległych ważonych odwrotnością ich odległości. W efekcie, im punkt jest bardziej oddalony, tym mniejszy jest jego wpływ na interpolowaną wartość. Wagę punktów ustala się z użyciem argumentu wykładnika potęgowego (ang. power).⁹ W metodzie IDW rola otaczających punkt estymowany danych w liczonej średniej jest zróżnicowana w zależności od odległości (Ogryzek, Kurowska 2016 Za ZAWADZKI, 2011):

$$\hat{Z}_j = \frac{\sum_{i=1}^n \frac{Z_i}{h_{ij}^\beta}}{\sum_{i=1}^n \frac{1}{h_i^\beta}}$$

gdzie:

\hat{Z}_j – wartość cechy Z estymowanej w punkcie j ,

Z_i – wartość cechy Z zmierzona w punkcie i (jednym z n punktów danych w otoczeniu),

h_{ij} – efektywna odległość między punktami i i j ,

β – wykładnik potęgowy – waga odległości.

- Radial Basis Functions (RBF) - metoda radialnych funkcji bazowych

Metoda radialnych funkcji bazowych RBF (Radial Basis Function), w której wartości rzeczywiste zależą od odległości od punktu początkowego. Metoda pozwala na wygenerowanie gładkiej powierzchni z dużych zbiorów danych (GIL I IN. 2018). Metody wykorzystujące radialne funkcje bazowe są szeregiem dokładnych technik interpolacyjnych, co oznacza, że powierzchnia musi przechodzić przez każdy punkt o znanej wartości. Metody RBF są rodzajem sztucznych sieci neuronowych. Metoda ta może wyznaczać wartości powyżej maksymalnej i poniżej minimalnej wartości danej (CICHOCIŃSKI 2011). Empiryczny kriging bayesowski to metoda interpolacji, która uwzględnia błąd w szacowaniu podstawowego semiwariogramu za pomocą powtarzanych symulacji.¹⁰

⁹ <https://bookdown.org/nowosad/geostatystyka/metody-interpolacji.html>

¹⁰ <https://desktop.arcgis.com/en/arcmap/10.5/tools/geostatistical-analyst-toolbox/empirical-bayesian-kriging.htm>

2. Stochastyczne (geostatystyczne, uwzględniają w tworzeniu modelu możliwą losowość w wartości zmiennej):

Kriging (metoda interpolacji oparta na geostatystyce) to grupa metod estymacji zaproponowana w latach 50. przez Daniela Krige. Główna zasada mówi, że prognoza w danej lokalizacji jest kombinacją obokległych obserwacji. Waga nadawana każdej z obserwacji jest zależna od stopnia (przestrzennej) korelacji - stąd też bierze się istotna rola semiwariogramów.¹¹ W metoda krigingu wyznacza się błąd interpolacji zwany wariancją krigingu. Algorytm krigingu jest skuteczny, ponieważ może zrekompensować dane w zbiorze, nadając tym obszarom mniejszą wagę w ogólnej prognozie. Pozwala również na ekstrapolację poza obszarem danych (GIL I IN. 2018). Estymacja za pomocą krigingu może się różnić ze względu na przyjęty model funkcji losowej.

- Kriging Ordinary (KO) – metoda krigingu zwyczajnego

Kriging zwyczajny to metoda estymowania powierzchni ciągłych za pomocą punktowych pomiarów danych ilościowych. Metoda ta umożliwia tworzenie map wartości oczekiwanej (średniej) oraz map błędu standardowego estymacji (odchylenia standardowego), który przyjmuje się, że jest równy pierwiastkowi zminimalizowanej wariancji. Wartości błędu estymacji są zależne od zmienności przestrzennej i rozmieszczenia punktów pomiarowych. Zwykły kriging zakłada model: $Z(S) = \mu + \varepsilon(s)$, gdzie μ jest nieznaną stałą.¹²

- Kriging Simple (KS) – metoda krigingu prostego

Kriging prosty zakłada, że średnia jest znana i stała na całym obszarze. Prosty kriging zakłada model: $Z(S) = \mu + \varepsilon(s)$, gdzie μ jest znaną stałą.

- Kriging Universal (KU) – metoda krigingu uniwersalnego

Kriging uniwersalny, określany również jako kriging z trendem zakłada, że nieznana średnia lokalna zmienia się stopniowo na badanym obszarze. W krigingu uniwersalnym możemy stosować zarówno zmienne jakościowe, jak i ilościowe.¹³ Używany jest w warunkach niestacjonarnych tzn. przy obecności dryftu (powolna zmiana wartości opisywanej zmiennej zregionalizowanej). Kriging uniwersalny stanowi rozszerzenie krigingu zwykłego, w którym usunięto konieczność wprowadzenia założenia o istnieniu stałej wartości oczekiwanej. Metoda

¹¹ <http://www.wbc.poznan.pl/Content/382516/estymacje-jednozmienne.html>

¹² <https://desktop.arcgis.com/fr/arcmap/latest/extensions/geostatistical-analyst/understanding-ordinary-kriging.htm>

¹³ <http://www.wbc.poznan.pl/Content/382516/estymacje-jednozmienne.html>

polega na zastosowaniu modelu przestrzennej regresji wielokrotnej, która rozdziela funkcję losową (stochastyczną) na liniową kombinację funkcji deterministycznych, określoną w każdym punkcie regionu oraz składnik losowości będący resztą funkcji losowej (ZASINA 2019 ZA: OLEA 1999; WACKERNAGEL 2003). Uniwersalny kriging zakłada model: $Z(s) = \mu(s) + \varepsilon(s)$, gdzie $\mu(S)$ jest jakąś funkcją deterministyczną.

- **Kriging Disjunctive (KD) – metoda krigingu dysjunktywnego**

Disjunctive kriging zakłada model: $f(Z(s)) = \mu_1 + \varepsilon(s)$, gdzie μ_1 jest nieznaną stałą, $f(Z(s))$ jest dowolną funkcją $Z(s)$. Kriging wskaźnikowy jest szczególnym przypadkiem krigingu disjunktywnego. W analizie geostatystycznej przewidzieć można samą wartość, albo wskaźnik z dysjunctive kriging. Kriging dysjunktywny może wykorzystywać semiwariogramy lub kowariance i przekształcenia, ale nie może uwzględniać błędu pomiaru.

- **Empirical Bayesian Kriging (EBK) - metoda empirycznego krigingu Bayesa**

Empiryczny Kriging Bayesowski (EBK) to geostatystyczna metoda interpolacji, która automatyzuje najtrudniejsze aspekty budowy prawidłowego modelu krigingu. EBK automatycznie oblicza niezbędne parametry poprzez proces podzbiorów i symulacji. EBK różni się od klasycznych metod krigingu uwzględnieniem błędu wprowadzonego przez estymację modelu semiwariogramu.

3. Część praktyczna – interpolacja przestrzenna

W części praktycznej wykonano interpolację przestrzenną dla danych pomiarów zaobserwowanych w Australii i Oceanii z wykorzystaniem walidacji podzbiorem 10%, 15% oraz 20%. Temperatura powietrza stanowi daną bezwzględną (dane bezwzględne są wynikiem pomiarów lub obserwacji bezpośrednich). Interpolowana temperatura powietrza przedstawiona została na mapie za pomocą skali skokowej w kolorze.

Informacją wejściową w procesie interpolacji są dane pomiarowe ze stacji meteorologicznych, które rozmieszczone są stosunkowo rzadko i nierównomiernie, a które z założenia mają być reprezentatywne dla znacznych obszarów. Dodatkowo stacje lokalizowane są tak, aby nie oddawały w nadmiernym stopniu efektów oddziaływania czynników meteorologicznych w skali lokalnej, co niejednokrotnie może stać na przeszkodzie w opracowaniu wiarygodnego rozkładu przestrzennego temperatury powietrza o wysokiej rozdzielczości przestrzennej (SZYMANOWSKI, KRYZA 2012).

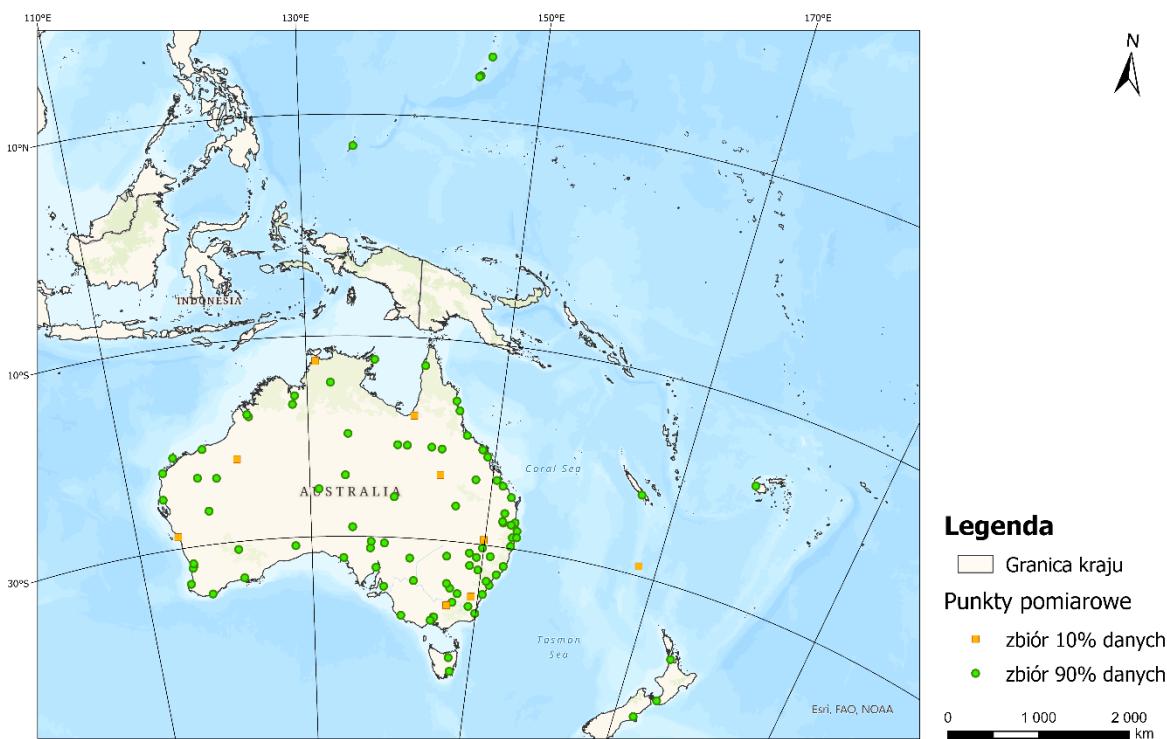
Temperatura przestrzenna może być traktowana jako zmienna zregionalizowana. Zmienne zregionalizowane można uznać za klasę zmiennych losowych, składających się z komponentów: strukturalnego, lokalnego i przypadkowego (SZYMANOWSKI ZA: STEIN 1999).

3.1. Walidacja z wykorzystaniem podzbioru 10%

Liczba punktów pomiarowych wynosi 101. Dane podzielono na dwa zbiory – 90% i 10%. Zbiór 90% zawiera 91 punktów pomiarowych, natomiast zbiór 10% zawiera 9 punktów pomiarowych, z czego 8 położone jest na obszarze Australii. Z uwagi na wyjście poza obszar analizy jednego punktu pomiarowego znajdującego się w podzbiorze 10% usunięto go.

Położenie wszystkich bazowych punktów pomiarowych temperatury powietrza w Australii i Oceanii ukazane jest na rysunku poniżej. Ich rozkład jest nierównomierny, największa liczba punktów pomiarowych znajduje się w Australii, pomiędzy równoleżnikami 10°S - 40°S. W obrębie Australii większość punktów położona jest wzdłuż wschodniego wybrzeża. Na obszarze Nowej Zelandii zlokalizowane są 3 punkty pomiarowe, pozostałe punkty położone są na wyspach Oceanu Spokojnego na wschód od Australii (Mikronezja, Malezja, Polinezja).

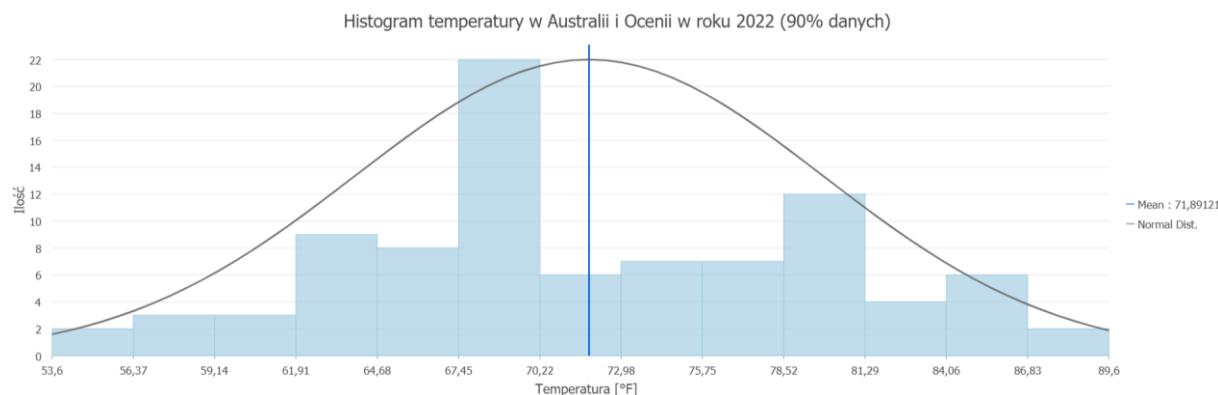
Położenie punktów pomiarowych temperatury powietrza w Australii i Oceanii (90% danych)



Rysunek 2 Położenie punktów pomiarowych temperatury powietrza w Australii i Oceanii dla zbiorów 90% i 10%

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

W przeprowadzonym badaniu oś pozioma oznacza przedział temperatury wyrażony w stopniach Farenheita, natomiast oś pionowa oznacza liczebność zbioru. Rozkład częstotliwości oraz statystyki podsumowujące przedstawiono na rysunkach poniżej. Rozkład częściowy będący wykresem słupkowym ukazuje liczebność przyporządkowanych zmiennych do przedziału wartości. Szerokości przedziałów są równe i wynoszą 2,77 stopnia. Analizując histogram przestawiony na rysunku poniżej można zauważyc, że największą liczebnością zbioru – dominantą – odznacza się przedział 67,45–70,22°F stopni, natomiast najmniejszą liczebność posiada pierwszy przedział 53,60–56,37°F oraz ostatni przedział 86,83–89,6°F. Zauważalne są również dwa wierzchołki o wyraźnie zwiększonej liczebności klas, tzw. rozkład dwumodalny. Na podstawie histogramu zauważalna jest niewielka asymetria (skośność) rozkładu, co potwierdzają dane statystyczne wyliczone w oprogramowaniu ArcGIS Pro. Więcej zaobserwowanych wyników położonych jest z lewej strony od średniej. Skośność wynosi 0,079, a kurtoza 2,34. W rozkładzie normalnym parametry te wynoszą odpowiednio 0 i 3.



Rysunek 3 Histogram temperatury w Australii i Oceanii w 2022 roku – 90 % danych

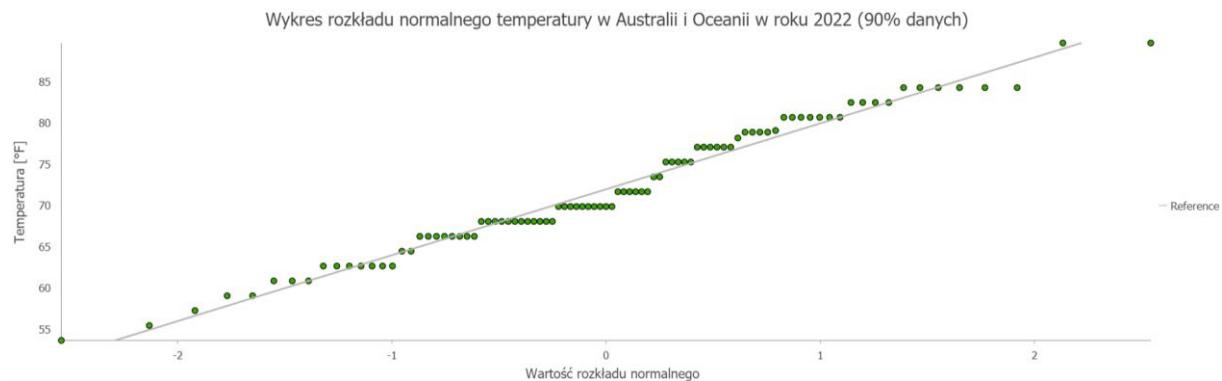
Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

Statistics	
Mean	71,9
Median	69,8
Std. Dev.	8,0
Rows	91
Count	91
Nulls	0
Min	53,6
Max	89,6
Sum	6 542,1
Skewness	0,079
Kurtosis	2,34

Rysunek 4 Wartości statystyczne histogramu temperatury powietrza dla zbioru 90%

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

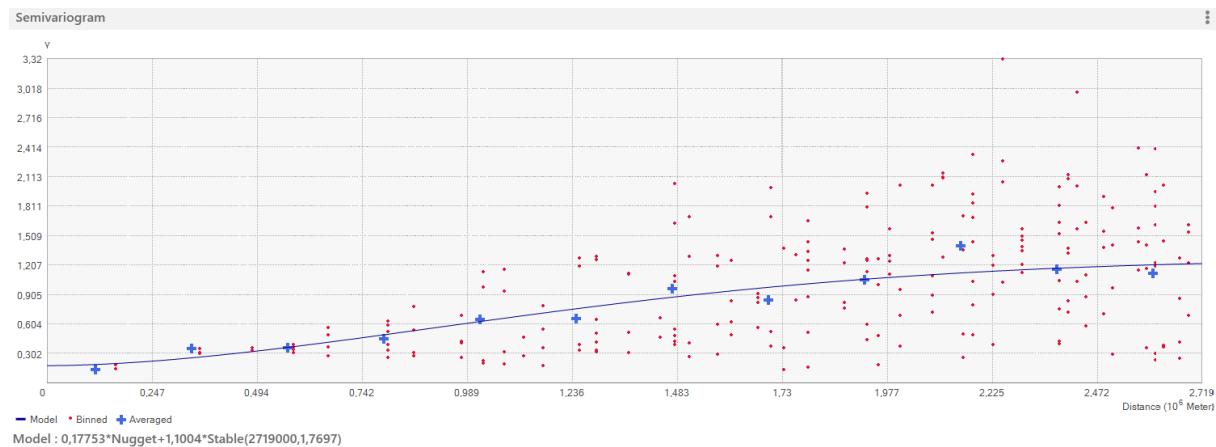
Analizując wykres poniżej można zauważyc, że widać drobne odstępstwa w przypadku punktów na początku i końcu wykresu – punkty te leżą dalej od prostej niż pozostałe punkty, jednak odstępstwo to jest na tyle małe, że uznać można, iż dane pochodzą z rozkładu normalnego.



Rysunek 5 Wykres rozkładu normalnego temperatury powietrza w Australii i Oceanii w 2022 roku dla 90% danych

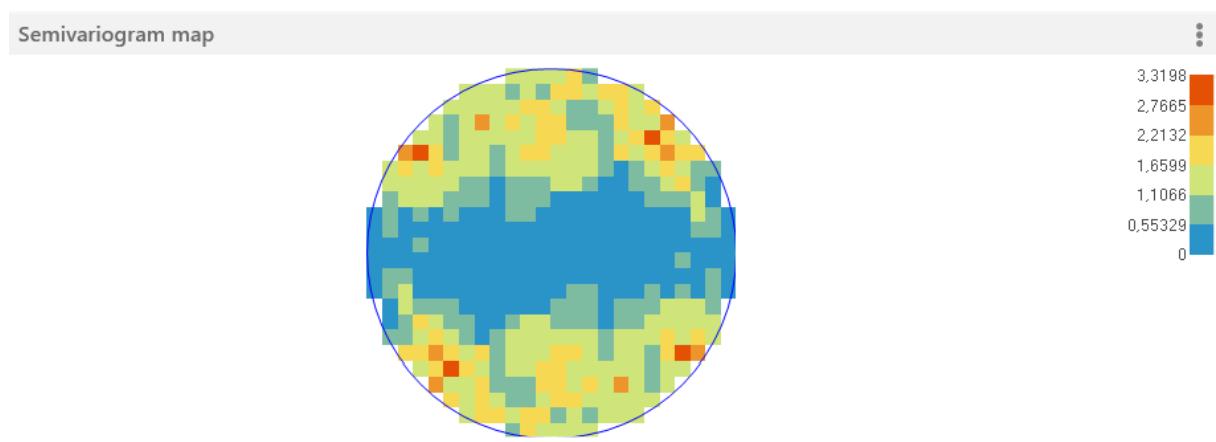
Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

Wykres poniżej przedstawia semiwariancje dla danych pomiarowych temperatury powietrza w Australii i Oceanii. Semiwariancja powstała poprzez utworzenie punktów na wykresie, gdzie oś rzędnych oznacza różnice w wysokości, a oś odciętych różnice w odległości. Otrzymany wykres ma charakter wykresu rozrzutu punktów. Na wykresie można zauważyc wzrost zmienności wraz ze zwiększeniem się odległości pomiędzy pomiarami, aż do osiągnięcia pewnego stałego poziomu. Analizując kształt wykresu można zauważyc, że otrzymana funkcja zbliżona jest wykresem do typowego modelu gaussowskiego. Nie zauważalna jest losowość danych, efekt samorodków (nugget effect) jest niewielki. Wyniki analizy semiwariancji pokazują, że pozyskane dane są wystarczające do wykonania analiz przestrzennych. Analizując mapę semiwariogramu można zauważyc, że badana powierzchnia wykazuje najmniejsze zróżnicowanie z zachodu na wschód.



Rysunek 6 Semivariogram dla analizowanych 90% danych

Źródło: opracowanie własne

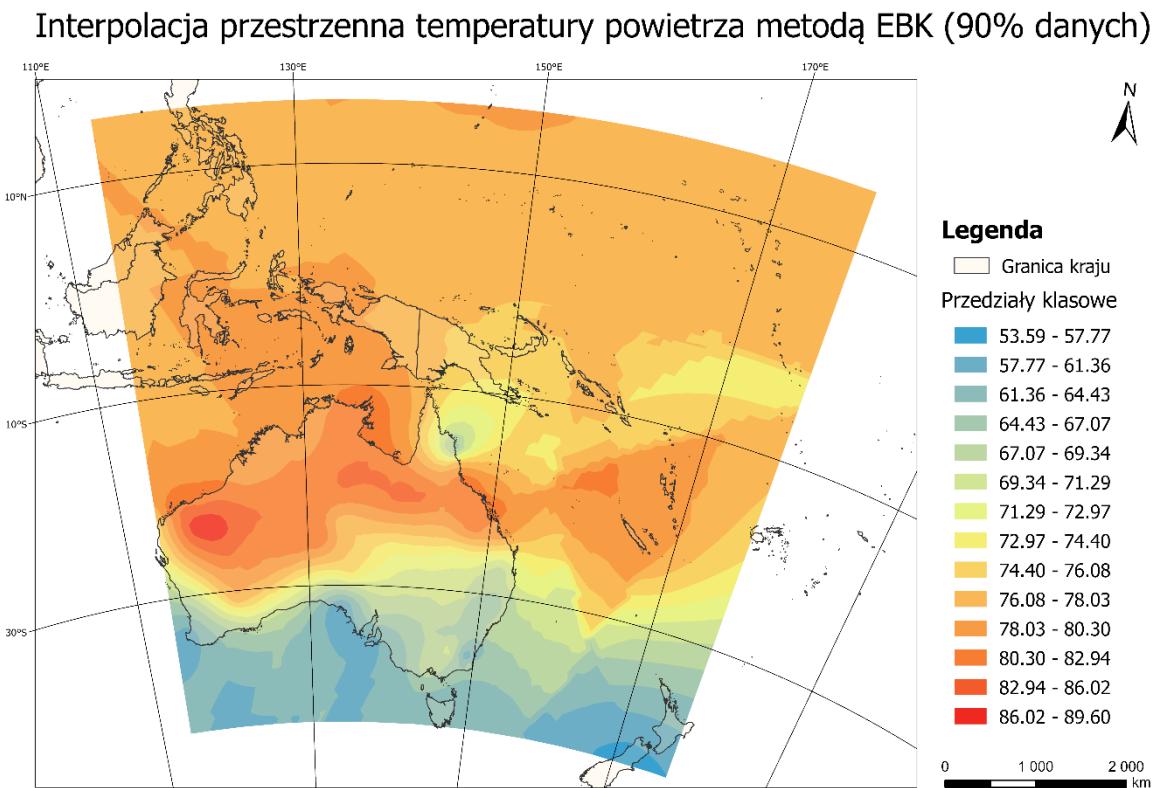


Rysunek 7 Mapa semiwariancji dla analizowanych 90% danych

Źródło: opracowanie własne

a) Empirical Bayesian Kriging

Rysunek 8 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą EBK (90% danych)

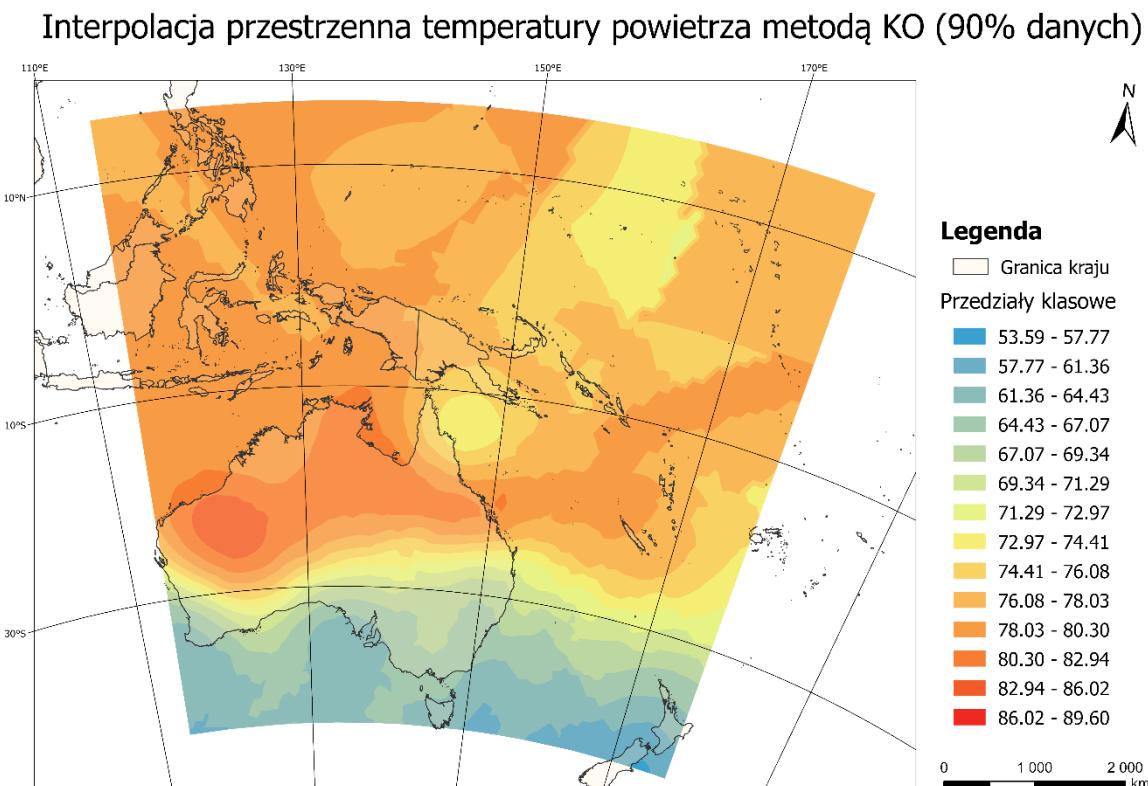


Źródło: opracowanie własne

Pierwszą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda EBK (Empiryczny Kriging Bayesowski). Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najniższe wartości znajdują się pod równoleżnikiem 30°S , najwyższe w okolicy punktu $20^{\circ}\text{S}, 120^{\circ}\text{E}$. Zauważalna jest anomalia wartości na północno-wschodnim wybrzeżu Australii, która może być spowodowana występowaniem występowaniem łańcucha górskiego Wielkich Górz Wododziałowych, biegnących wzdłuż wschodniej granicy Australii). Wysoka wartość temperatury w północno-zachodniej części Australii spowodowana może być występowaniem na tym obszarze rozległych obszarów pustynnych. Analizując rysunek powyżej można zauważać skoki pomiędzy przedziałami, w szczególności na obszarach znajdujących się pod równoleżnikiem 10°S . Wartość RMSE EBK wynosi 2,70.

b) Kriging Ordinary

Rysunek 9 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Ordinary (90% danych)

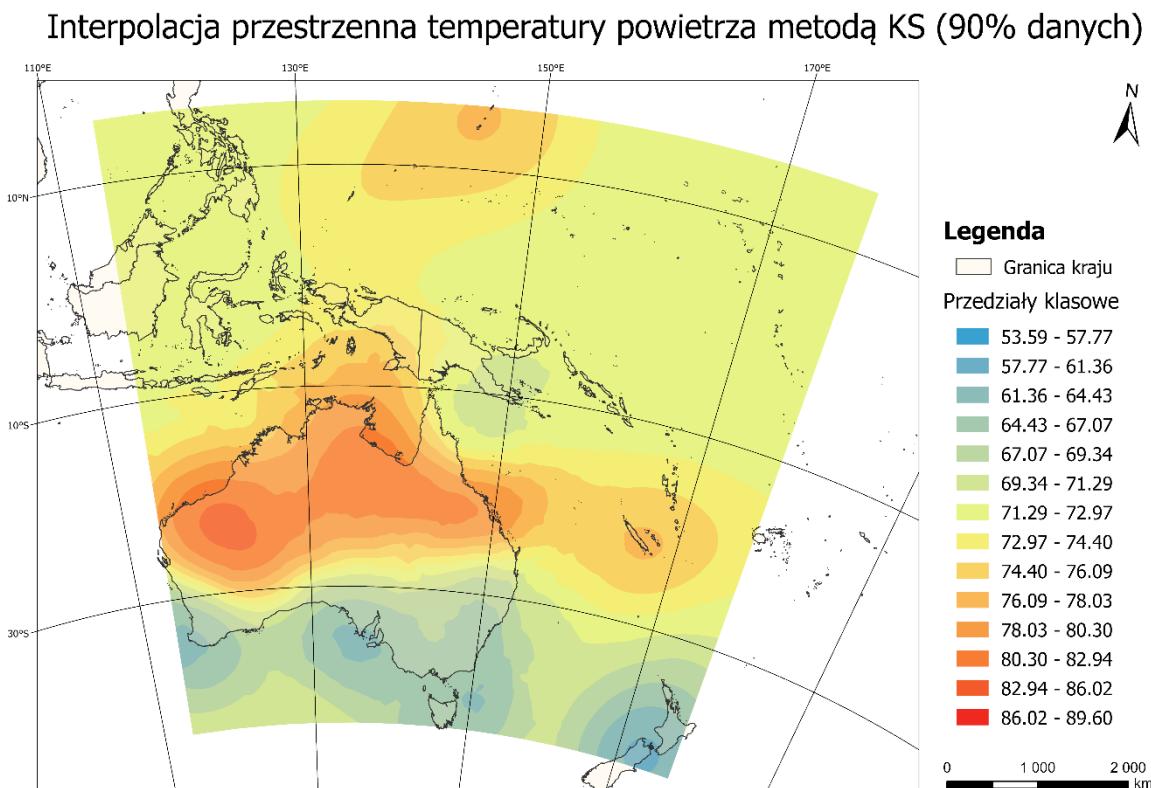


Źródło: opracowanie własne

Drugą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Kriging Ordinary (Kriging zwykły). Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najwyższe wartości temperatur znajdują się w północnej części Australii. Najniższe wartości znajdują się na obszarze Australii. Granice pomiędzy przedziałami są wyraźne – ostre, w szczególności w północno-wschodniej części obszaru. Wartość RMSE Kriging Ordinary wynosi 2,72. W przypadku tej metody anomalia nie jest tak silna w porównaniu do metody EBK, dodatkowo cały obszar badawczy jest silnie zróżnicowany, a w przypadku metody EBK północna część należy głównie do jednego przedziału klasowego.

c) Kriging Simple

Rysunek 10 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Simple (90% danych)



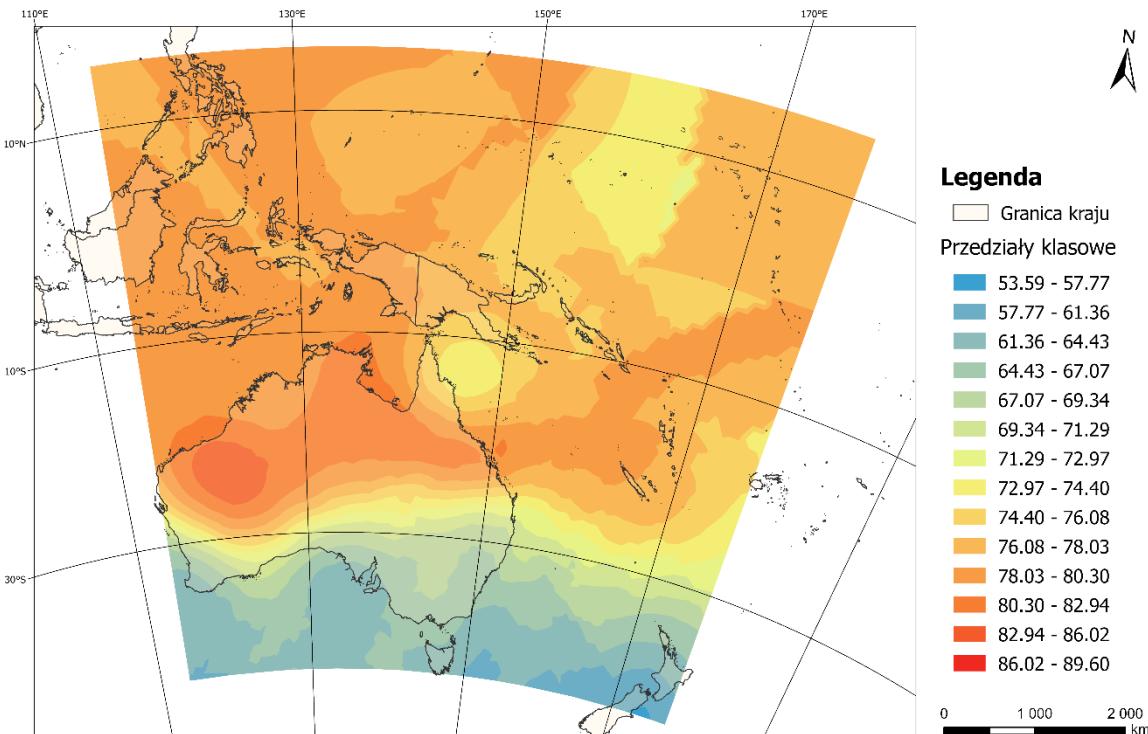
Źródło: opracowanie własne

Trzecią metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Kriging Simple. Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. W przypadku tej metody rozpiętość barw na mapie jest mniejsza. Większość obszaru należy do 3 – 4 przedziałów klasowych. Wartości w powyższych analizach zostały sklasyfikowane do innych przedziałów niż w powyższych analizowanych metodach. Analizując mapę można zauważyć duże obszary znajdujące się w obrębie jednego przedziału klasowego. Najwyższe wartości znajdują się na obszarze Australii północno-zachodniej. Wartość RMSE Kriging Simple wynosi 3,14 i jest najwyższa spośród wszystkich analizowanych metod.

d) Kriging Universal

Rysunek 11 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Universal (90% danych)

Interpolacja przestrzenna temperatury powietrza metodą KU (90% danych)

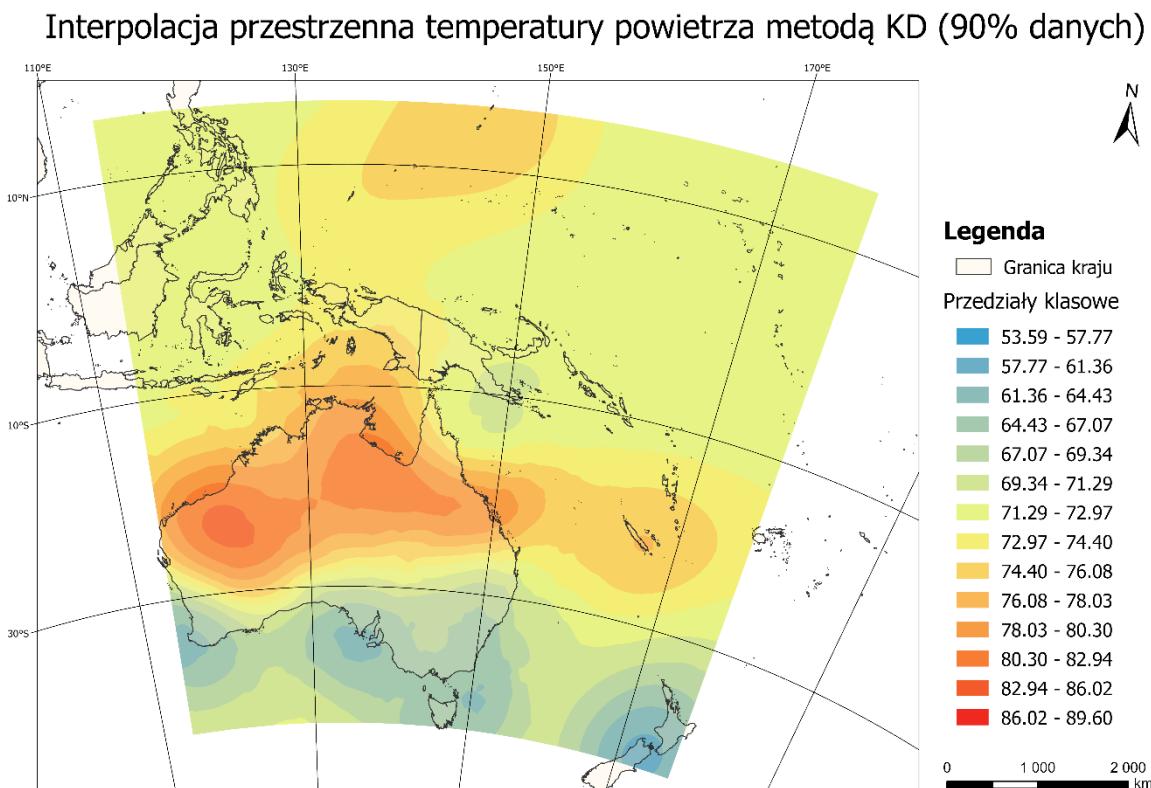


Źródło: opracowanie własne

Czwartą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Kriging Universal. Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najwyższe wartości temperatury znajdują się w Australii północno-zachodniej. Analizując interpolację przestrzenną przedstawioną na rysunku powyżej można zauważyć, że najniższe wartości temperatur znajdują się pod równoleżnikiem 30°S. Wartość RMSE Kriging Universal wynosi 2,72, czyli tyle samo co w przypadku metody Kriging Ordinary. Wizualizacje mapowe obu metod są bardzo zbliżone.

e) Kriging Disjunctive

Rysunek 12 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Disjunctive (90% danych)

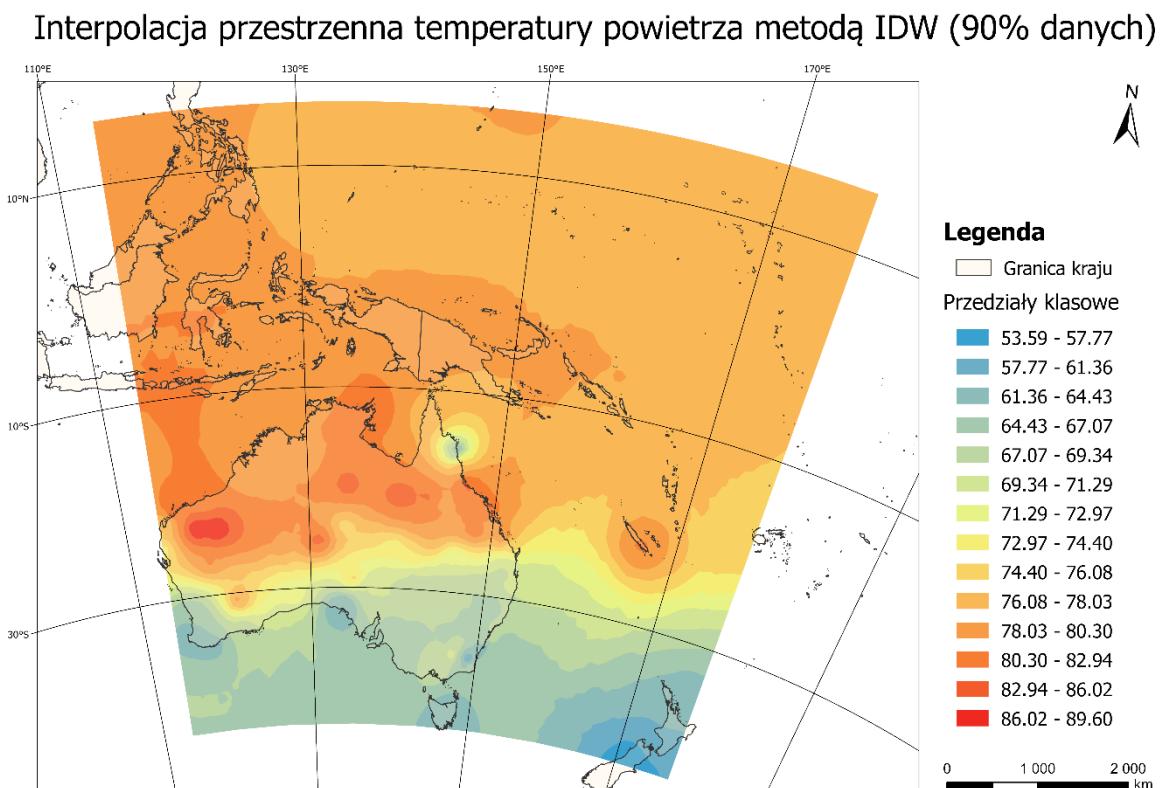


Źródło: opracowanie własne

Piątą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Kriging Disjunctive. Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Większość obszaru znajduje się w obrębie jednego przedziału, wyjątkiem jest obszar Australii północnej, gdzie zauważalna jest podwyższona wartość temperatury powietrza - zauważalne są wysypy wartości wysokich temperatur oraz obszar pod równoleżnikiem 30°S, gdzie zauważalne są niższe temperatury niż na pozostałych obszarach. Wartość RMSE Kriging Disjunctive wynosi 3,12, co jest szóstym najwyższym wynikiem wśród analizowanych 7 metod interpolacji przestrzennej.

f) Inverse Distance Weighting

Rysunek 13 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą IDW (90% danych)

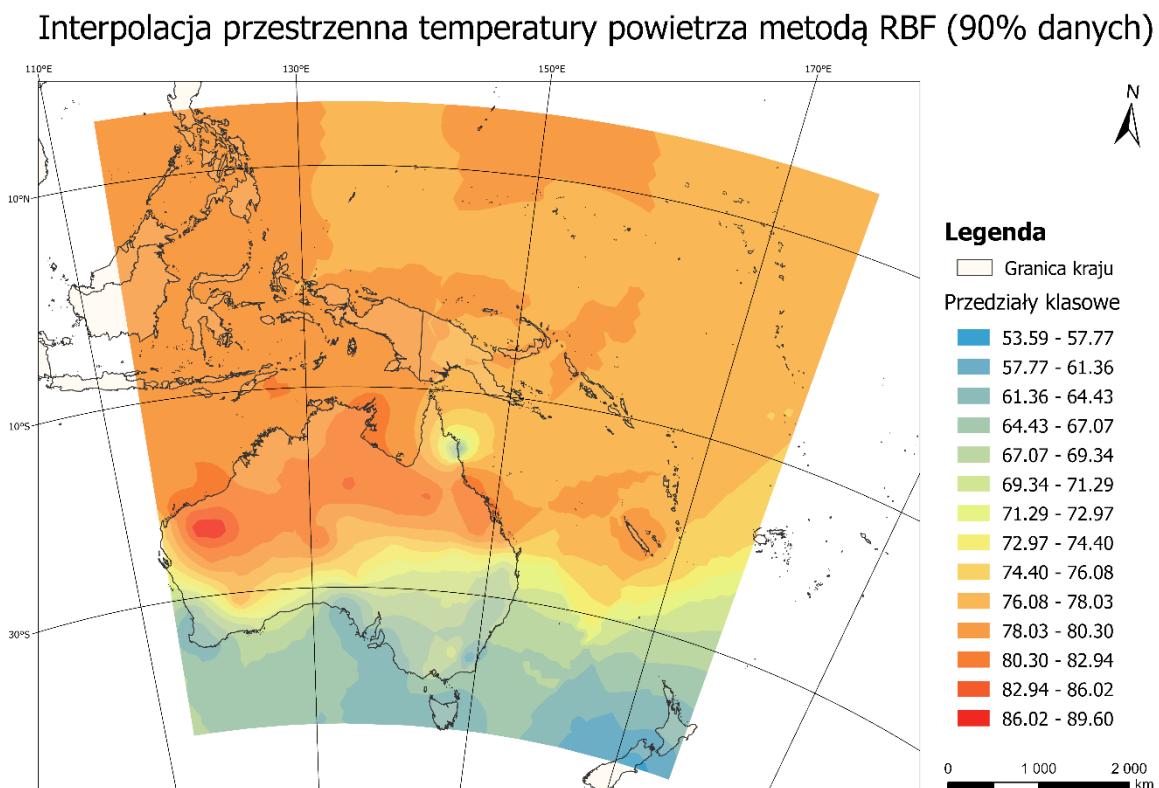


Źródło: opracowanie własne

Szóstą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Inverse Distance Weighting (IDW, metoda odwrotnej odległości). Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najniższe wartości znajdują się na obszarach znajdujących się pod równoleżnikiem 30°S, natomiast najwyższe w północno – zachodniej części analizowanego obszaru. W przypadku tej metody cała część obszaru powyżej równoleżnika 30°S zakwalifikowana została do przedziałów o wysokich temperaturach. Wyjątkiem jest anomalia na wschodnim wybrzeżu Australii, gdzie występuje punktowy obszar o niższej temperaturze w porównaniu do okolicznych obszarów. Negatywną cechą IDW jest tendencja do tworzenia tzw. efektów oczkowych (SZYMANOWSKI 2014). Analizując interpolację przestrzenną przedstawioną na rysunku powyżej można zauważyc, że występują duże obszary należące do jednego przedziału, przerywane kilkoma punktowymi obszarami o wyższej lub niższej klasie temperatury. Wartość RMSE IDW wynosi 2,84.

g) Radial Basis Functions

Rysunek 14 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą RBF



Źródło: opracowanie własne

Siódmą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Radial Basis Functions (RBF, Radialna Funkcja Bazowa). Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najniższe wartości znajdują się pod równoleżnikiem 30°S. Analizując interpolację przestrzenną przedstawioną na rysunku powyżej można zauważyc, że występują dosyć silne wahania. Na mapie widoczne jest wiele obszarów różnych klas oraz występują silne załamania na granicach tych obrębów. Wartość RMSE RBF wynosi 2,70 i jest jednym z najniższych wartości wśród analizowanych metod dla tego zbioru.

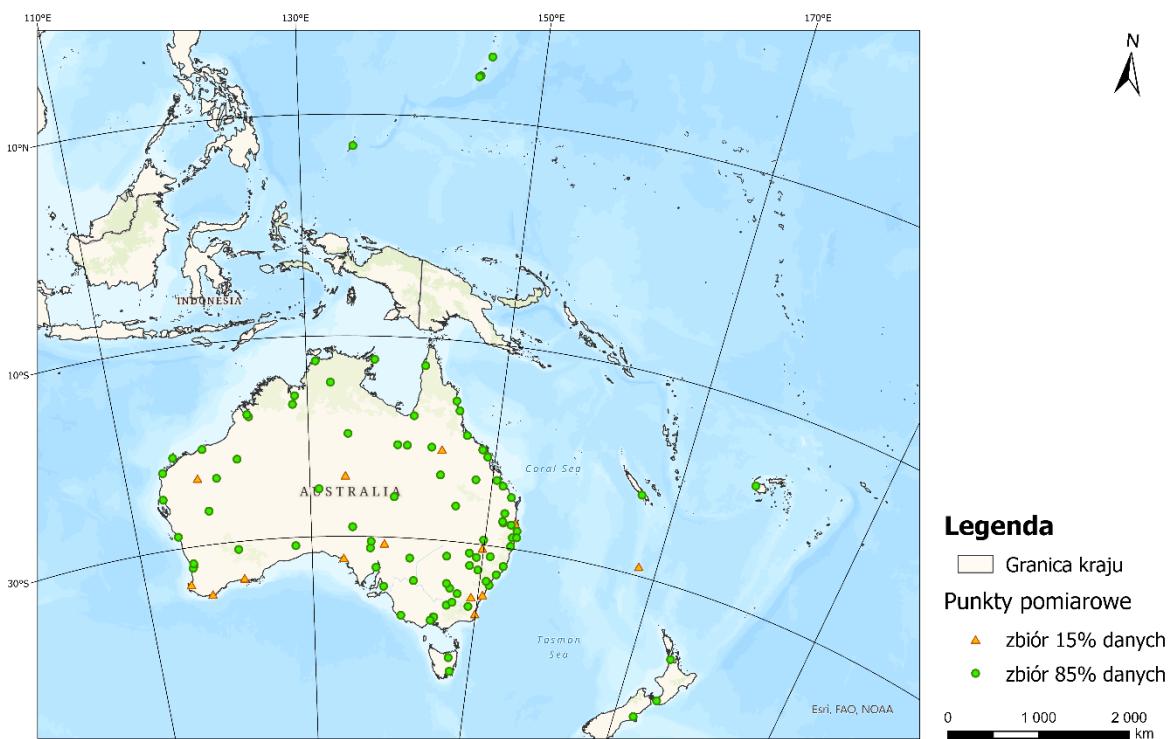
Analizując wszystkie interpolacje przestrzenne można zauważyć, że tam gdzie sieć punktów jest gęstsza (obszar Australii, w szczególności południowo-zachodniej) otrzymany obraz jest bardziej zróżnicowany, w obszarze tym występuje wiele klas, a przejście między nimi są wyraźnie widoczne, natomiast mniejsza ilość punktów bazowych (obszar morskie) charakteryzuje się zmniejszoną liczbą klas.

3.2. Walidacja z wykorzystaniem podzbioru 15%

Liczba punktów pomiarowych wynosi 101. Dane podzielono na dwa zbiory – 90% i 15%. Zbiór 85% zawiera 86 punktów pomiarowych, natomiast zbiór 15% zawiera 14 punktów pomiarowych, z czego 13 położone jest na obszarze Australii. Z uwagi na wyjście poza obszar analizy jednego punktu pomiarowego znajdującego się w podzbiorze 15% usunięto go.

Położenie wszystkich bazowych punktów pomiarowych temperatury powietrza w Australii i Oceanii ukazane jest na rysunku poniżej. Ich rozkład jest nierównomierny, największa liczba punktów pomiarowych znajduje się w Australii, pomiędzy równoleżnikami 10°S - 40°S. W obrębie Australii większość punktów położona jest wzdłuż wschodniego wybrzeża. Na obszarze Nowej Zelandii zlokalizowane są 3 punkty pomiarowe, pozostałe punkty położone są na wyspach Oceanu Spokojnego na wschód od Australii (Mikronezja, Malezja, Polinezja).

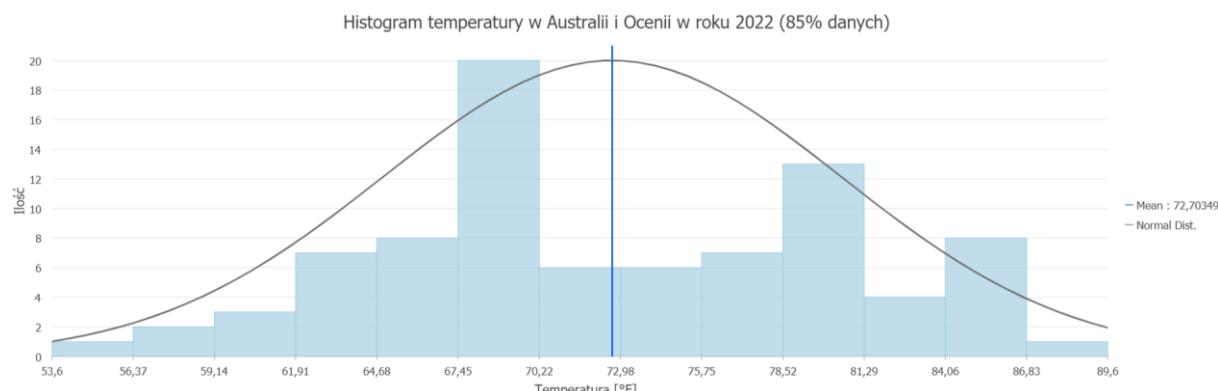
Położenie punktów pomiarowych temperatury powietrza w Australii i Oceanii (85% danych)



Rysunek 15 Położenie punktów pomiarowych temperatury powietrza w Australii i Oceanii dla zbiorów 85% i 15%

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

W przeprowadzonym badaniu oś pozioma oznacza przedział temperatury wyrażony w stopniach Farenheita, natomiast oś pionowa oznacza liczebność zbioru. Rozkład częstotliwości oraz statystyki podsumowujące przedstawiono na rysunkach poniżej. Rozkład częściowy będący wykresem słupkowym ukazuje liczebność przyporządkowanych zmiennych do przedziału wartości. Szerokości przedziałów są równe i wynoszą 2,77 stopnia. Analizując histogram przestawiony na rysunku poniżej można zauważyc, że największą liczebnością zbioru – dominantą – odznacza się przedział 67,45–70,22°F stopni, natomiast najmniejszą liczebność posiada pierwszy przedział 53,60–56,37°F oraz ostatni przedział 86,83–89,6°F. Zauważalne są również dwa wierzchołki o wyraźnie zwiększonej liczebności klas, tzw. rozkład dwumodalny. Na podstawie histogramu nie zaobserwowano asymetrii (skośności) rozkładu, co potwierdzają dane statystyczne wyliczone w oprogramowaniu ArcGIS Pro. Skośność wynosi 0,0. Kurtoza wynosząca 2,2 informuje, że analizowane dane są w mniejszym stopniu skoncentrowane wokół średniej niż w przypadku podzbioru 10%. W rozkładzie normalnym parametry te wynoszą odpowiednio 0 i 3.



Rysunek 16 Histogram temperatury w Australii i Oceanii w 2022 roku – 85 % danych

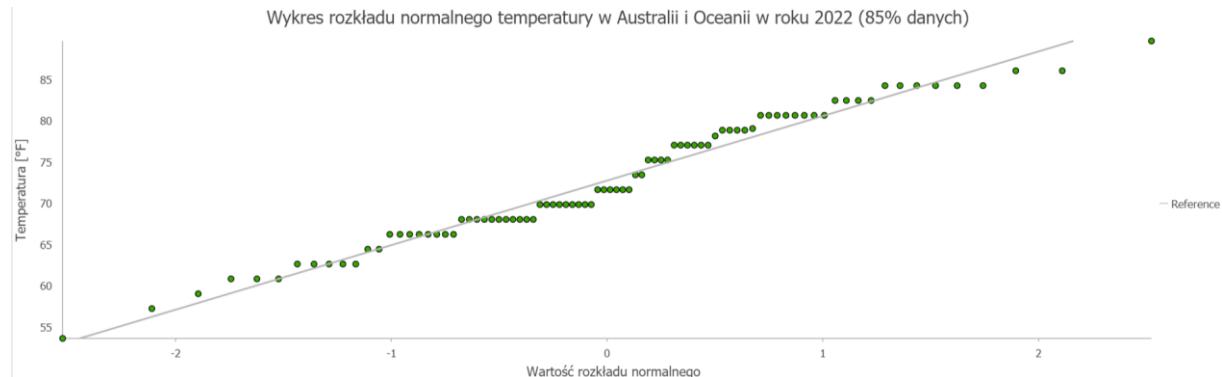
Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

Statistics	
Dataset	
<input checked="" type="checkbox"/> Mean	72,7
<input type="checkbox"/> Median	71,6
<input type="checkbox"/> Std. Dev.	7,8
Rows	86
Count	86
Nulls	0
Min	53,6
Max	89,6
Sum	6 252,5
Skewness	0,0
Kurtosis	2,2

Rysunek 17 Wartości statystyczne histogramu temperatury powietrza dla zbioru 85%

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

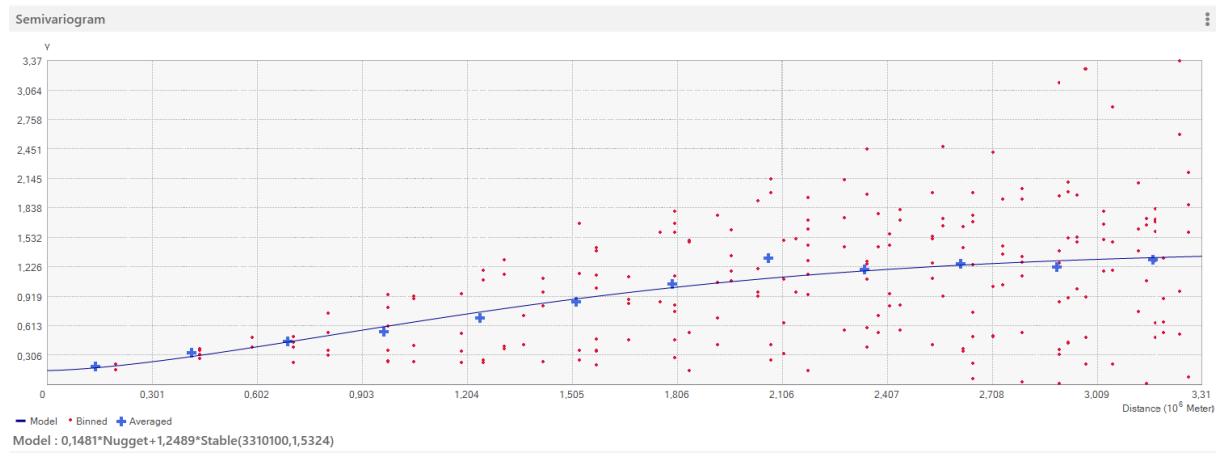
Analizując wykres poniżej można zauważyc, że widać drobne odstępstwa w przypadku punktów na początku i końcu wykresu – punkty te leżą dalej od prostej niż pozostałe punkty, jednak odstępstwo to jest na tyle małe, że uznać można, iż dane pochodzą z rozkładu normalnego.



Rysunek 18 Wykres rozkładu normalnego temperatury powietrza w Australii i Oceanii w 2022 roku dla 85% danych

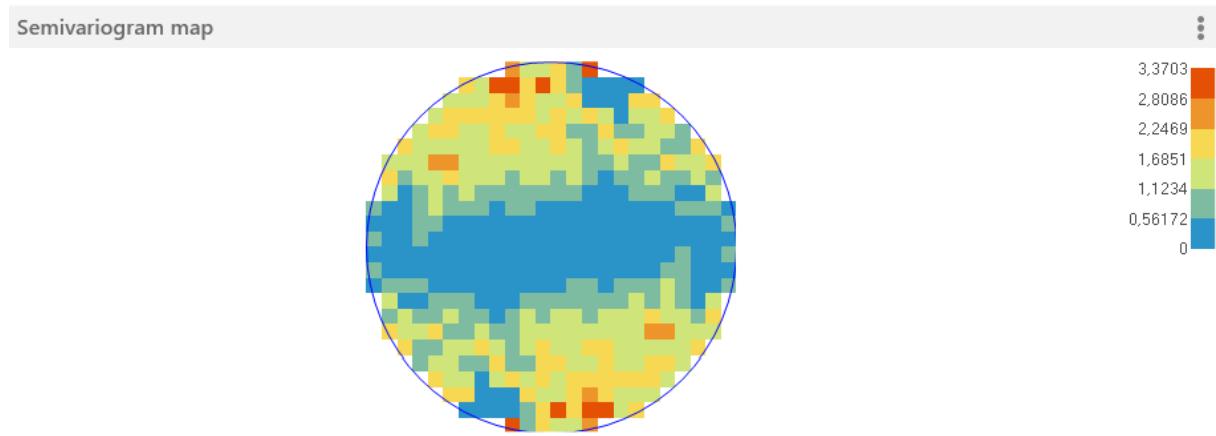
Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

Wykres poniżej przedstawia semiwariancje dla danych pomiarowych temperatury powietrza w Australii i Oceanii. Semiwariancja powstała poprzez utworzenie punktów na wykresie, gdzie oś rzędnych oznacza różnice w wysokości, a oś odciętych różnice w odległości. Otrzymany wykres ma charakter wykresu rozrzutu punktów. Na wykresie można zauważyc wzrost zmienności wraz ze zwiększeniem się odległości pomiędzy pomiarami, aż do osiągnięcia pewnego stałego poziomu. Analizując kształt wykresu można zauważyc, że otrzymana funkcja zbliżona jest wykresem do typowego modelu gaussowskiego. Nie zauważalna jest losowość danych, efekt samorodków (nugget effect) jest niewielki. Wyniki analizy semiwariancji pokazują, że pozyskane dane są wystarczające do wykonania analiz przestrzennych. Analizując mapę semivariogramu można zauważyc, że badana powierzchnia wykazuje najmniejsze zróżnicowanie z zachodu na wschód. Model ten wykazuje najlepsze dopasowanie spośród trzech analizowanych zbiorów.



Rysunek 19 Semivariogram dla analizowanych 85% danych

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

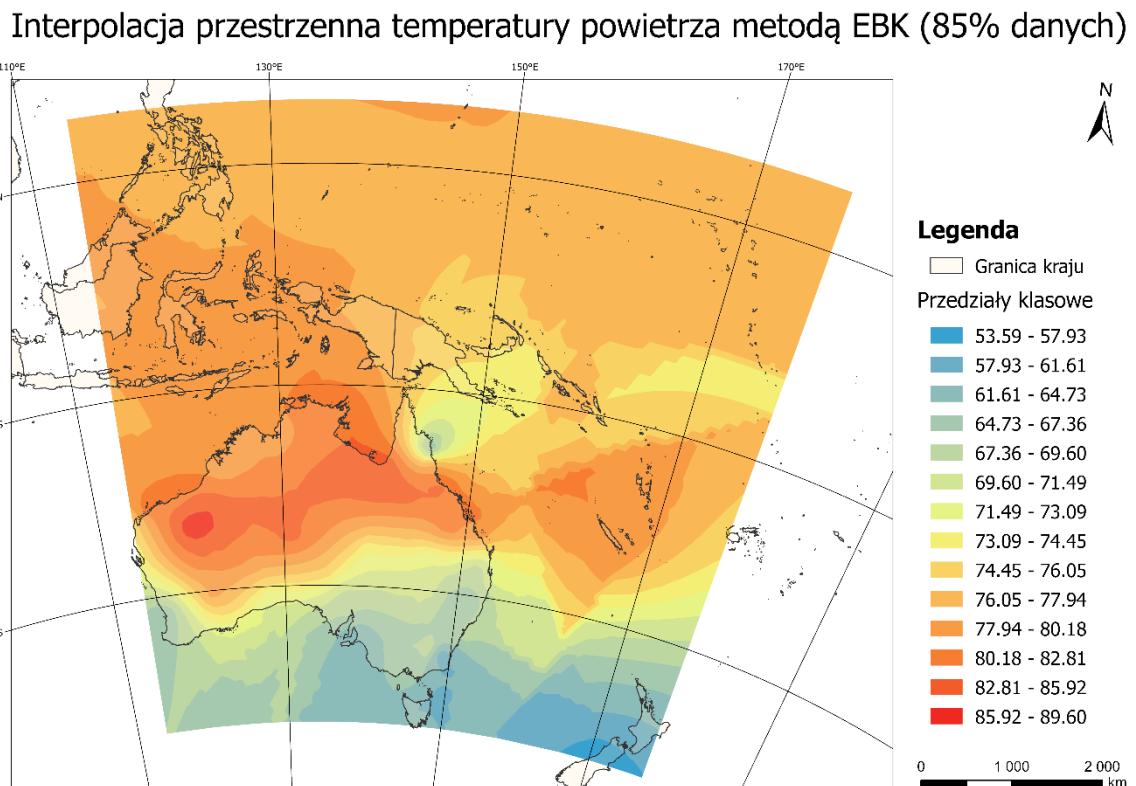


Rysunek 20 Mapa semiwariancji dla analizowanych 85% danych

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

a) Empirical Bayesian Kriging

Rysunek 21 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą EBK

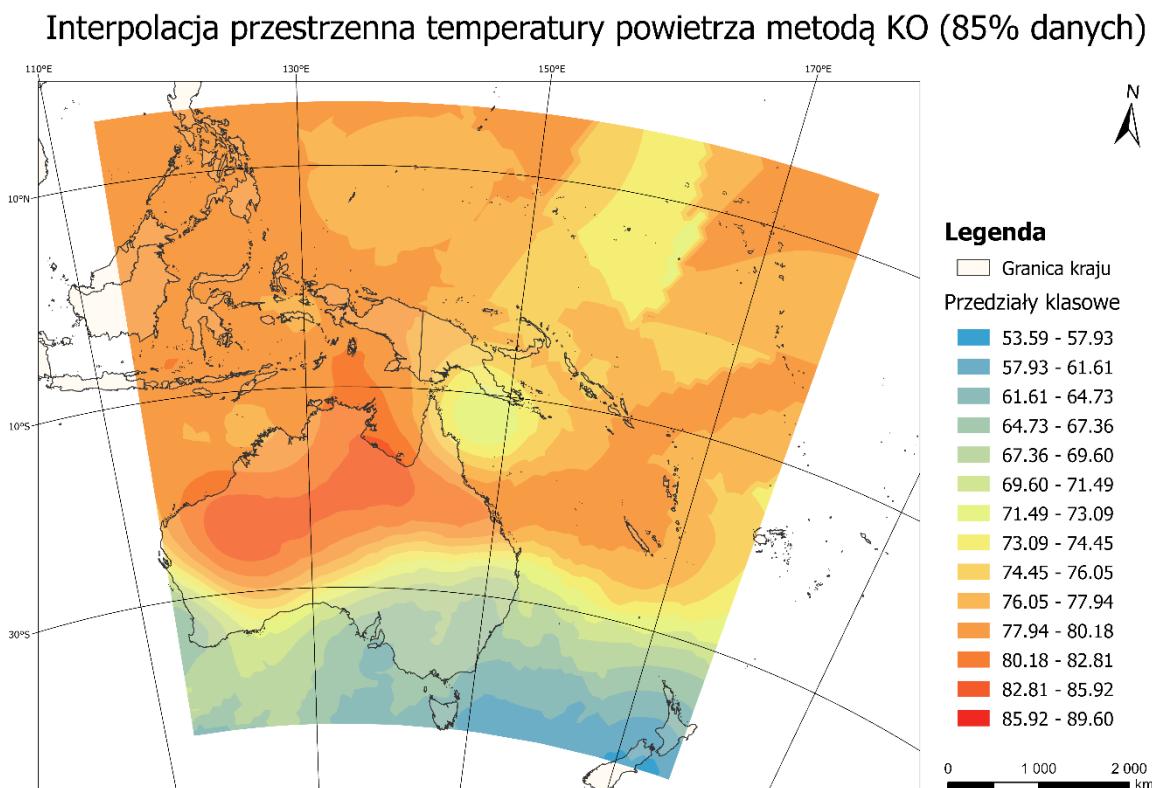


Źródło: opracowanie własne

Pierwszą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda EBK (Empiryczny Kriging Bayesowski). Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najniższe wartości znajdują się pod równoleżnikiem 30°S, najwyższe w okolicy punktu 20°S, 120°E. Zauważalna jest anomalia wartości na północno-wschodnim wybrzeżu Australii, która może być spowodowana występowaniem występowaniem łańcucha górskiego Wielkich Górz Wododziałowych, biegnących wzdłuż wschodniej granicy Australii). Wysoka wartość temperatury w północno-zachodniej części Australii spowodowana może być występowaniem na tym obszarze rozległych obszarów pustynnych. Analizując rysunek powyżej można zauważać skoki pomiędzy przedziałami. Wartość RMSE EBK wynosi 5,82.

b) Kriging Ordinary

Rysunek 22 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Ordinary

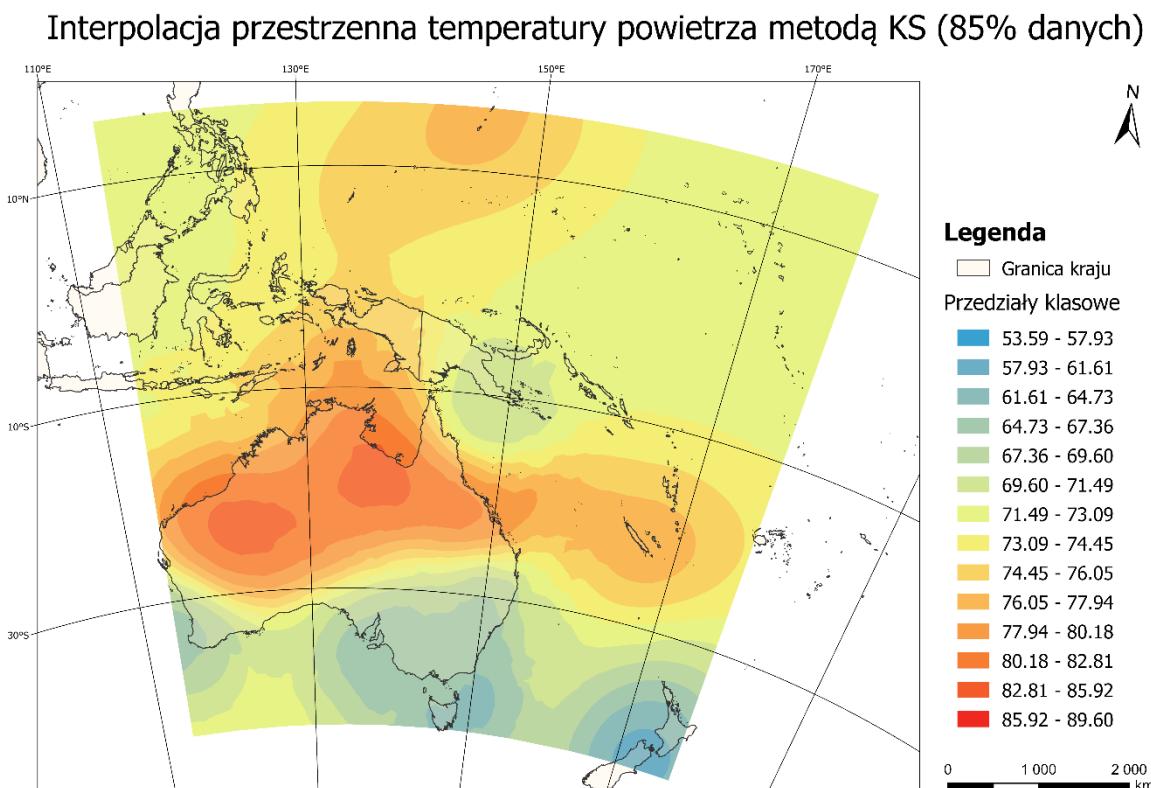


Źródło: opracowanie własne

Drugą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Kriging Ordinary (Kriging zwykły). Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najwyższe wartości temperatur znajdują się w północnej części Australii. Najniższe wartości znajdują się na obszarze Australii. Duże nagromadzenie i nachodzenie się obszarów z innych przedziałów klasowych zauważalne jest w północno-wschodniej części analizowanego obszaru. Granice pomiędzy przedziałami na tym obszarze są ostre. Wartość RMSE Kriging Ordinary wynosi 5,58. W przypadku tej metody anomalia nie jest tak silna w porównaniu do metody EBK, dodatkowo cały obszar badawczy jest silnie zróżnicowany, a w przypadku metody EBK północna część należy głównie do jednego przedziału klasowego.

c) Kriging Simple

Rysunek 23 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Simple

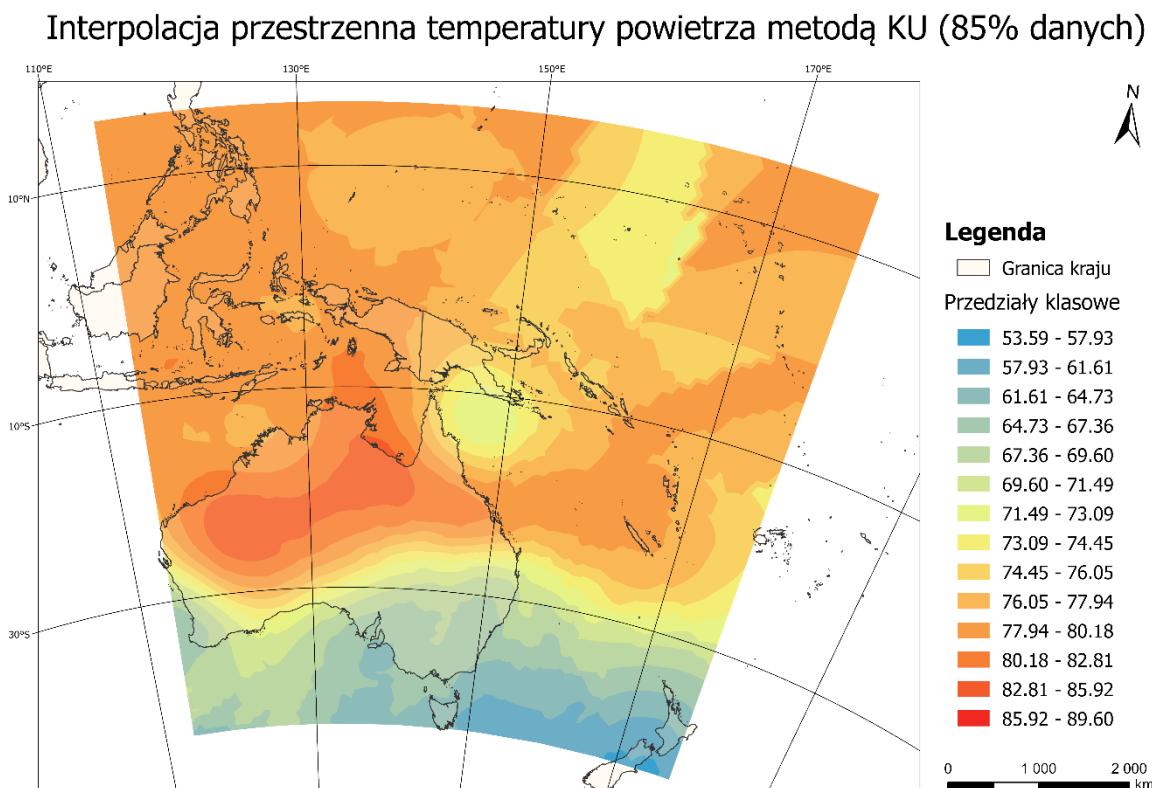


Źródło: opracowanie własne

Trzecią metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Kriging Simple. Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. W przypadku tej metody rozpiętość barw na mapie jest mniejsza. Większość obszaru należy do 3–4 przedziałów klasowych. Wartości w powyższych analizach zostały sklasyfikowane do innych przedziałów niż w powyższych analizowanych metodach. Analizując mapę można zauważać duże obszary znajdujące się w obrębie jednego przedziału klasowego. Największe zróżnicowanie temperaturowe zauważalne jest w centralnej części analizowanego obszaru. Najwyższe wartości znajdują się na obszarze Australii północno-zachodniej. Wartość RMSE Kriging Simple wynosi 5,77, co sprawia, że jest drugą najbardziej odpowiednią metodą interpolacji dla tego zbioru danych.

d) Kriging Universal

Rysunek 24 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Universal

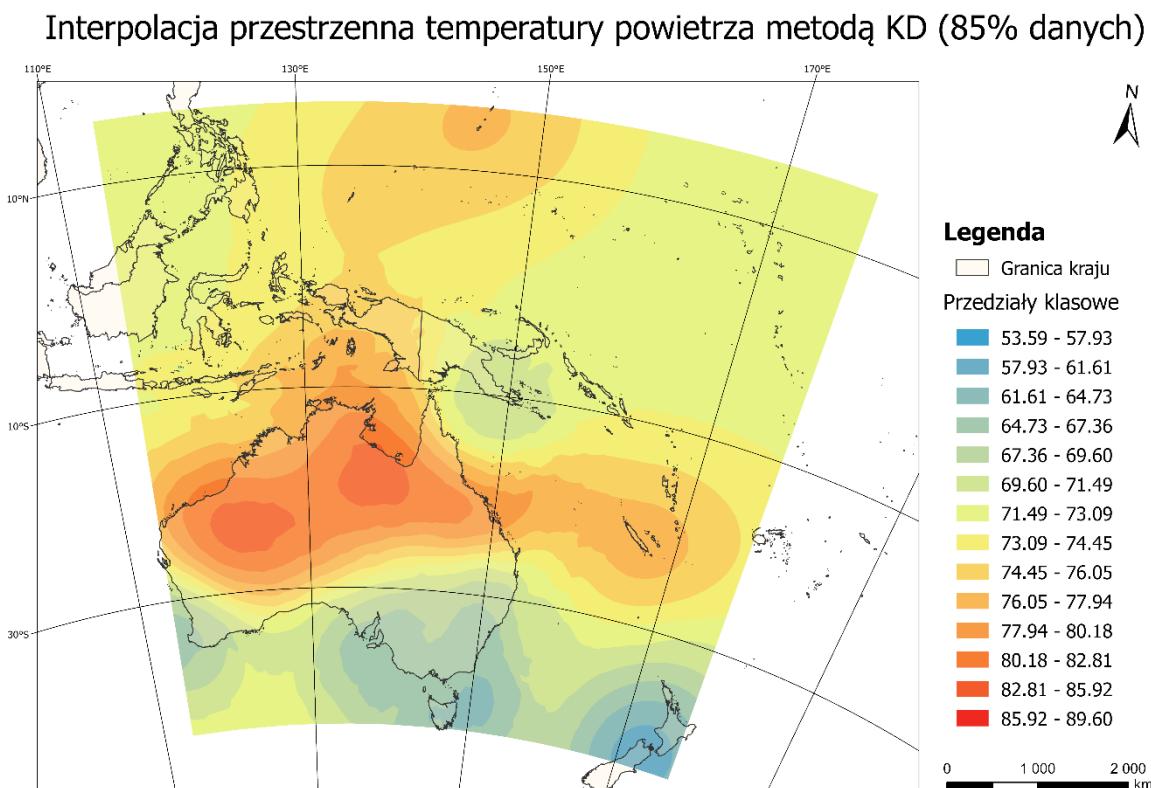


Źródło: opracowanie własne

Czwartą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Kriging Universal. Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najwyższe wartości temperatury znajdują się w Australii północno-zachodniej. Analizując interpolację przestrzenną przedstawioną na rysunku powyżej można zauważać dużą powierzchnię anomalii występującą na obszarze Papua-Nowa Gwinei. Wartość RMSE Kriging Universal wynosi 5,58, czyli tyle samo co w przypadku metody Kriging Ordinary. Obie metody w przypadku tego zbioru posiadają najniższy wskaźnik RMSE. Wizualizacje mapowe obu metod są bardzo zbliżone.

e) Kriging Disjunctive

Rysunek 25 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Disjunctive

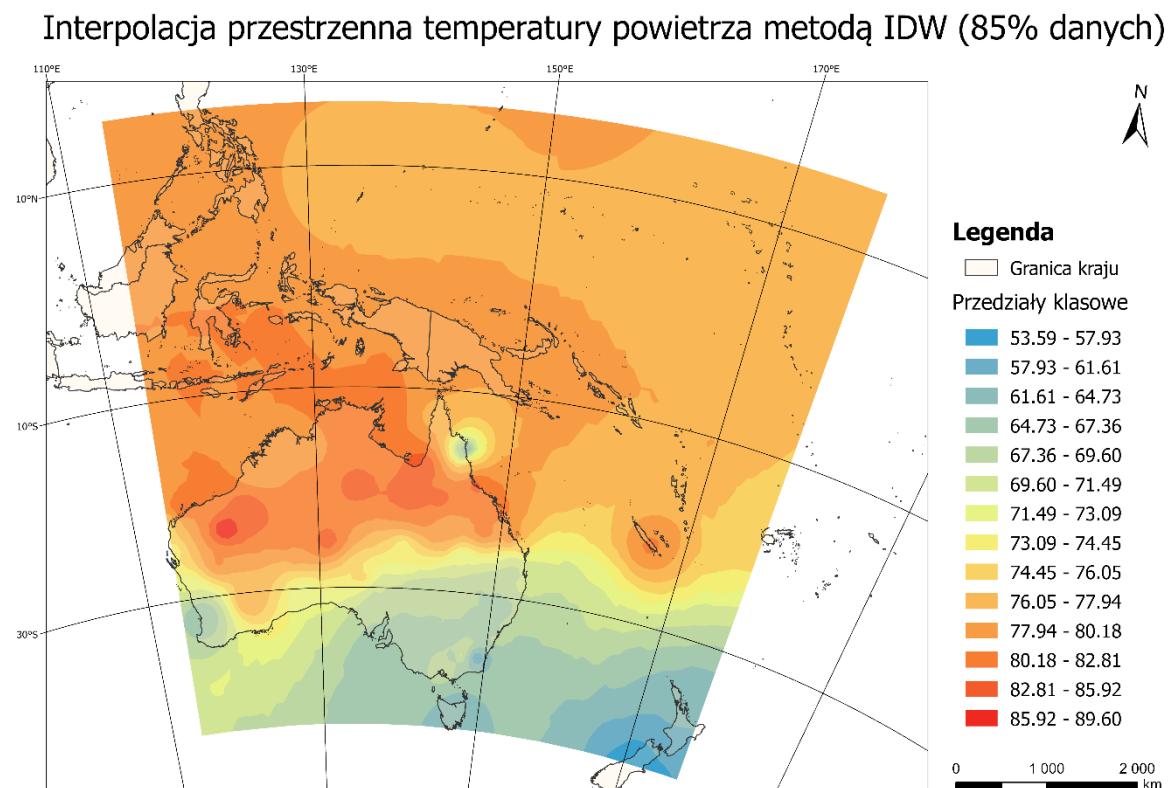


Źródło: opracowanie własne

Piątą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Kriging Disjunctive. Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najniższe wartości znajdują się w północno-wschodniej części analizowanego obszaru, natomiast najwyższe w południowej części analizowanego obszaru. Analizując interpolację przestrzenną przedstawioną na rysunku 12 można zauważyć, występowanie wyspy wartości wysokich temperatur na obszarze Australii północnej. Wartość RMSE Kriging Disjunctive wynosi 4,22, co jest szóstym najwyższym wynikiem wśród analizowanych 7 metod interpolacji przestrzennej.

f) Inverse Distance Weighting

Rysunek 26 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą IDW

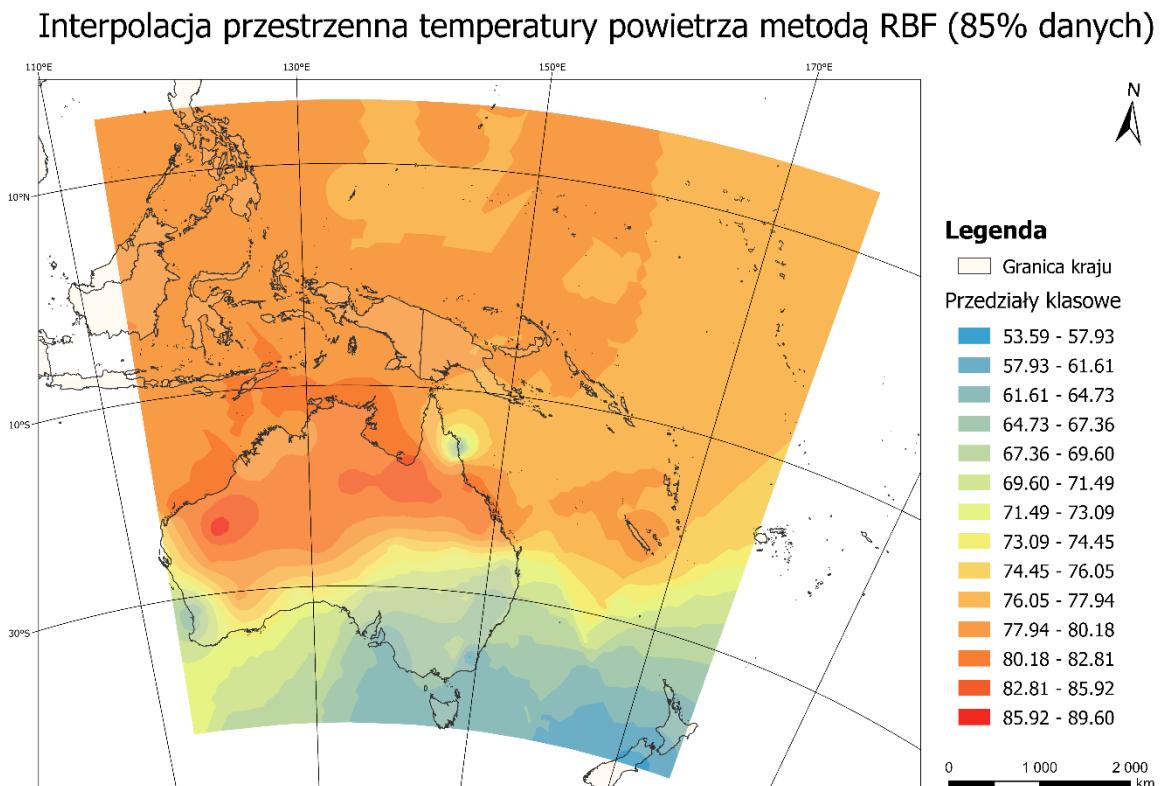


Źródło: opracowanie własne

Szóstą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Inverse Distance Weighting (IDW, metoda odwrotnej odległości). Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najniższe wartości znajdują się na obszarach znajdujących się pod równoleżnikiem 30°S , natomiast najwyższe w północno – zachodniej części analizowanego obszaru. W przypadku tej metody cała część obszaru powyżej równoleżnika 30°S zakwalifikowana została do przedziałów o wysokich temperaturach. Wyjątkiem jest anomalia na wschodnim wybrzeżu Australii, gdzie występuje punktowy obszar o niższej temperaturze w porównaniu do okolicznych obszarów. Negatywną cechą IDW jest tendencja do tworzenia tzw. efektów oczkowych (SZYMANOWSKI 2014). Analizując interpolację przestrzenną przedstawioną na rysunku powyżej można zauważyc, że występują duże obszary należące do jednego przedziału, przerywane kilkoma punktowymi obszarami o wyższej lub niższej klasie temperatury. Wartość RMSE IDW wynosi 6,01.

g) Radial Basis Functions

Rysunek 27 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą RBF



Źródło: opracowanie własne

Siódmą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Radial Basis Functions (RBF, Radialna Funkcja Bazowa). Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najniższe wartości znajdują się pod równoleżnikiem 30°S. Analizując interpolację przestrzenną przedstawioną na rysunku powyżej można zauważyc, że znaczna część obszaru sklasyfikowana została do przedziału 76,05-77,94 oraz 77,94-80,18. Na mapie widoczne jest wiele obszarów różnych klas oraz występują silne załamania na granicach tych obrębów. Wartość RMSE RBF wynosi 6,47, co jest najniższą wartością wśród analizowanych metod dla tego zbioru.

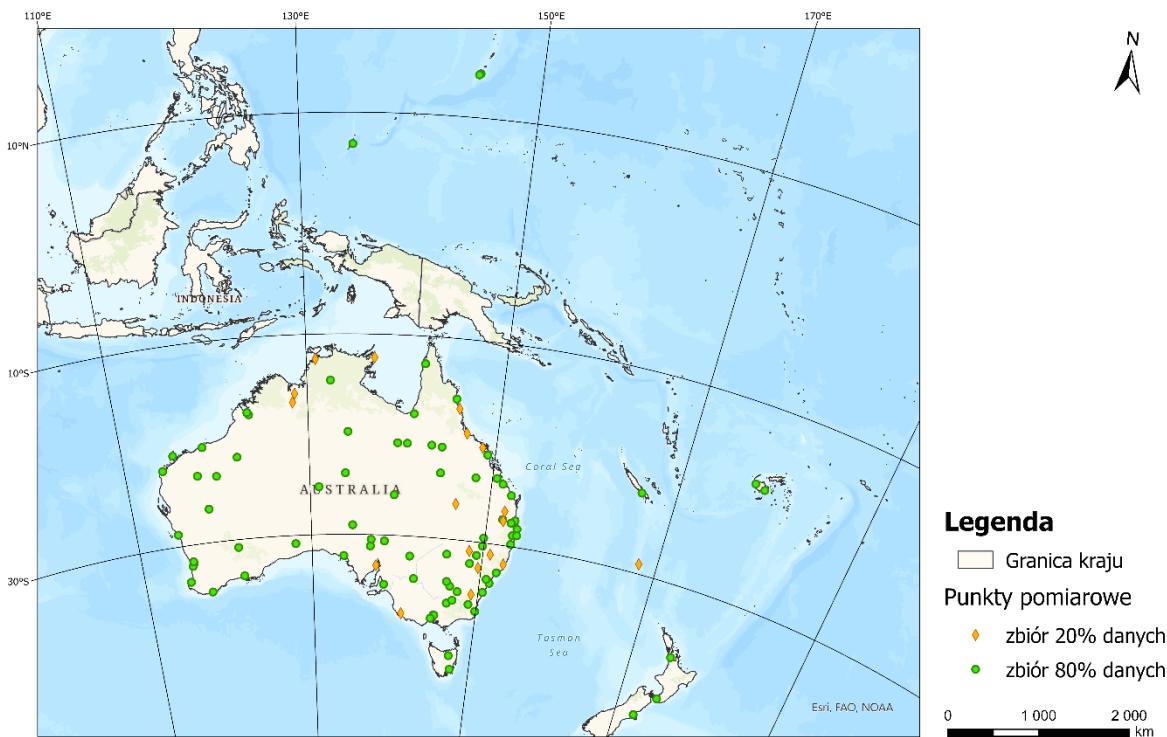
Analizując wszystkie interpolacje przestrzenne można zauważyc, że tam gdzie sieć punktów jest gęstsza (obszar Australii, w szczególności południowo-zachodniej) otrzymany obraz jest bardziej zróżnicowany, w obszarze tym występuje wiele klas, a przejście między nimi są wyraźnie widoczne, natomiast mniejsza ilość punktów bazowych (obszar morskie) charakteryzuje się zmniejszoną liczbą klas.

3.3. Walidacja z wykorzystaniem podzbioru 20%

Liczba punktów pomiarowych wynosi 101. Dane podzielono na dwa zbiory – 80% i 20%. Zbiór 80% zawiera 81 punktów pomiarowych, natomiast zbiór 20% zawiera 18 punktów pomiarowych, z czego 17 położone jest na obszarze Australii. Z uwagi na wyjście poza obszar analizy dwóch punktów pomiarowych znajdujących się w podzbiorze 20% usunięto je.

Położenie wszystkich bazowych punktów pomiarowych temperatury powietrza w Australii i Oceanii ukazane jest na rysunku poniżej. Ich rozkład jest nierównomierny, największa liczba punktów pomiarowych znajduje się w Australii, pomiędzy równoleżnikami 10°S - 40°S. W obrębie Australii większość punktów położona jest wzdłuż wschodniego wybrzeża. Na obszarze Nowej Zelandii zlokalizowane są 3 punkty pomiarowe, pozostałe punkty położone są na wyspach Oceanu Spokojnego na wschód od Australii (Mikronezja, Malezja, Polinezja).

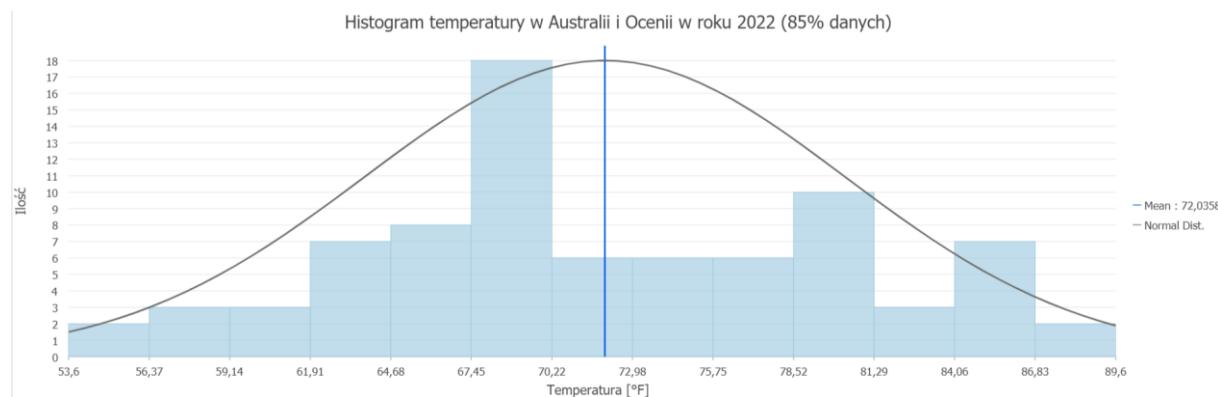
Położenie punktów pomiarowych temperatury powietrza w Australii i Oceanii (80% danych)



Rysunek 28 Położenie punktów pomiarowych temperatury powietrza w Australii i Oceanii dla zbiorów 80% i 20%

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

W przeprowadzonym badaniu oś pozioma oznacza przedział temperatury wyrażony w stopniach Farenheita, natomiast oś pionowa oznacza liczebność zbioru. Rozkład częstotliwości oraz statystyki podsumowujące przedstawiono na rysunkach poniżej. Rozkład częściowy będący wykresem słupkowym ukazuje liczebność przyporządkowanych zmiennych do przedziału wartości. Szerokości przedziałów są równe i wynoszą 2,77 stopnia. Analizując histogram przestawiony na rysunku poniżej można zauważyć, że największą liczebnością zbioru – dominantą – odznacza się przedział 67,45–70,22°F stopni, natomiast najmniejszą liczebność posiada pierwszy przedział 53,60–56,37°F oraz ostatni przedział 86,83–89,6°F. Na podstawie histogramu nie zaobserwowano asymetrii (skośności) rozkładu, co potwierdzają dane statystyczne wyliczone w oprogramowaniu ArcGIS Pro. Skośność wynosi 0,09, a kurtoza 2,3. W rozkładzie normalnym parametry te wynoszą odpowiednio 0 i 3.



Rysunek 29 Histogram temperatury w Australii i Oceanii w 2022 roku – 80 % danych

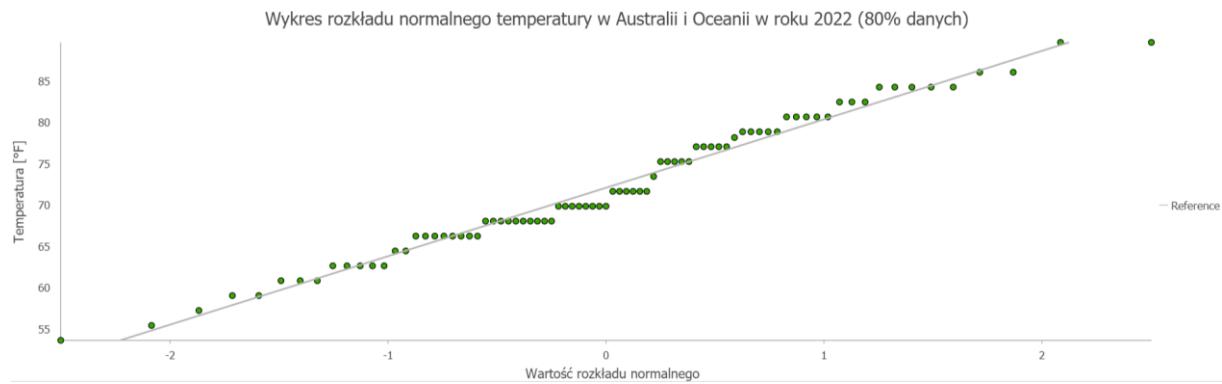
Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

Statistics	
Dataset	
<input checked="" type="checkbox"/> Mean	72,0
<input type="checkbox"/> Median	69,8
<input type="checkbox"/> Std. Dev.	8,3
Rows	81
Count	81
Nulls	0
Min	53,6
Max	89,6
Sum	5 834,9
Skewness	0,09
Kurtosis	2,3

Rysunek 30 Wartości statystyczne histogramu temperatury powietrza dla zbioru 80%

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

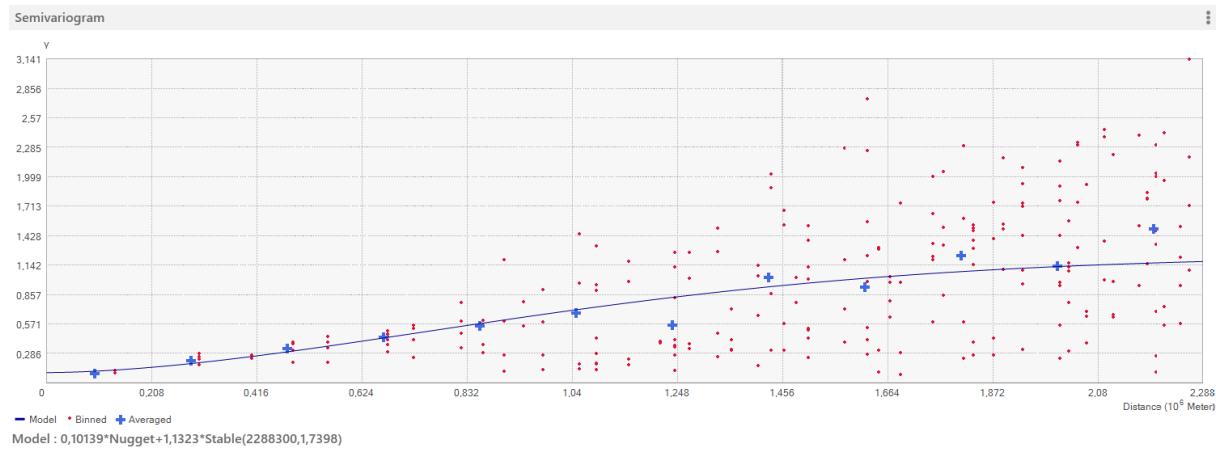
Analizując wykres poniżej można zauważyc, że widać drobne odstępstwa w przypadku punktów na początku i końcu wykresu – punkty te leżą dalej od prostej niż pozostałe punkty, jednak odstępstwo to jest na tyle małe, że uznać można, iż dane pochodzą z rozkładu normalnego.



Rysunek 31 Wykres rozkładu normalnego temperatury powietrza w Australii i Oceanii w 2022 roku dla 80% danych

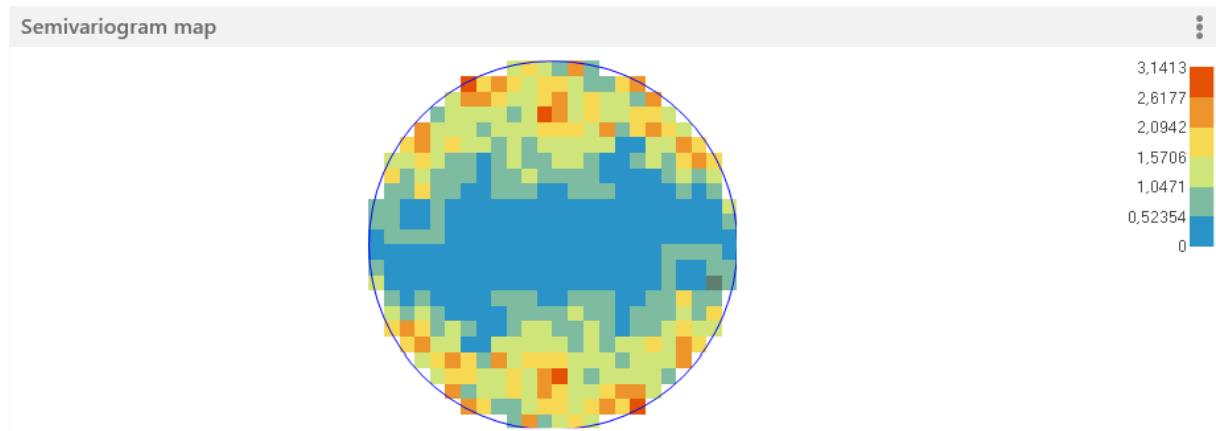
Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

Wykres poniżej przedstawia semiwariancje dla danych pomiarowych temperatury powietrza w Australii i Oceanii. Semiwariancja powstała poprzez utworzenie punktów na wykresie, gdzie oś rzędnych oznacza różnice w wysokości, a oś odciętych różnice w odległości. Otrzymany wykres ma charakter wykresu rozrzutu punktów. Na wykresie można zauważyc wzrost zmienności wraz ze zwiększeniem się odległości pomiędzy pomiarami, aż do osiągnięcia pewnego stałego poziomu. Analizując kształt wykresu można zauważyc, że otrzymana funkcja zbliżona jest wykresem do typowego modelu gaussowskiego. Nie zauważalna jest losowość danych, efekt samorodków (nugget effect) jest niewielki. Wyniki analizy semiwariancji pokazują, że pozyskane dane są wystarczające do wykonania analiz przestrzennych. Analizując mapę semiwariogramu można zauważyc, że badana powierzchnia wykazuje najmniejsze zróżnicowanie z zachodu na wschód.



Rysunek 32 Semivariogram dla analizowanych 80% danych

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro



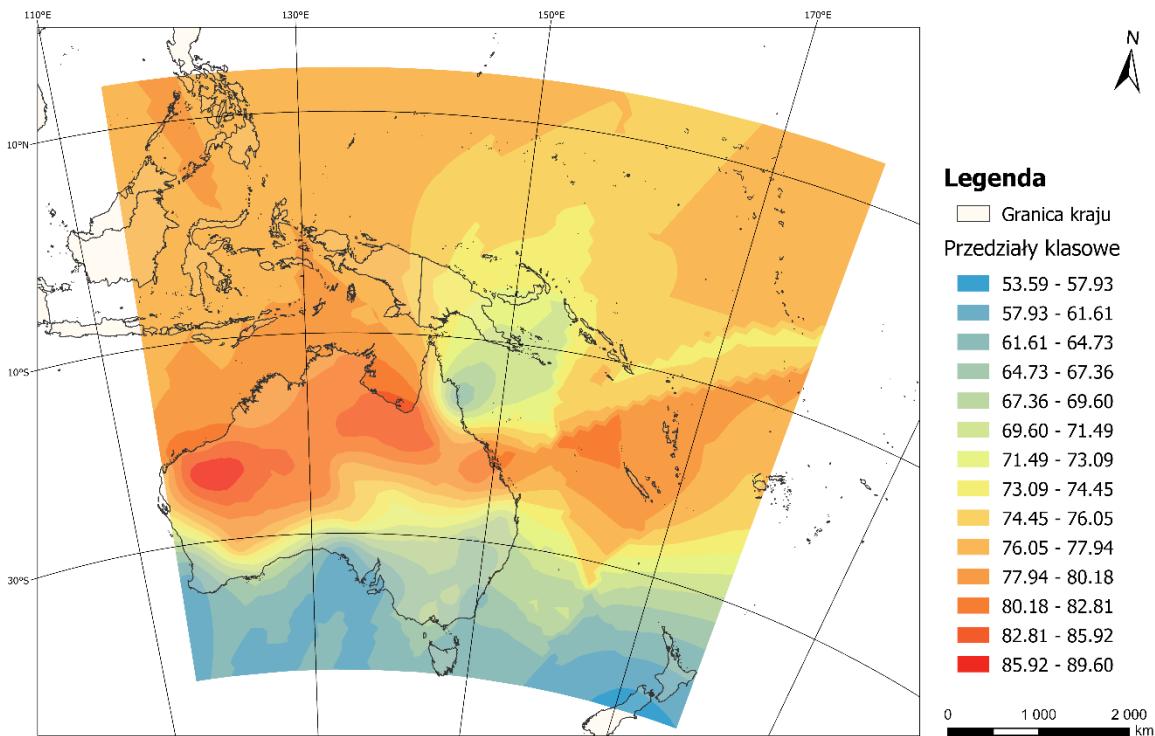
Rysunek 33 Mapa semiwariancji dla analizowanych 80% danych

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

a) Empirical Bayesian Kriging

Rysunek 34 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą EBK

Interpolacja przestrzenna temperatury powietrza metodą EBK (80% danych)

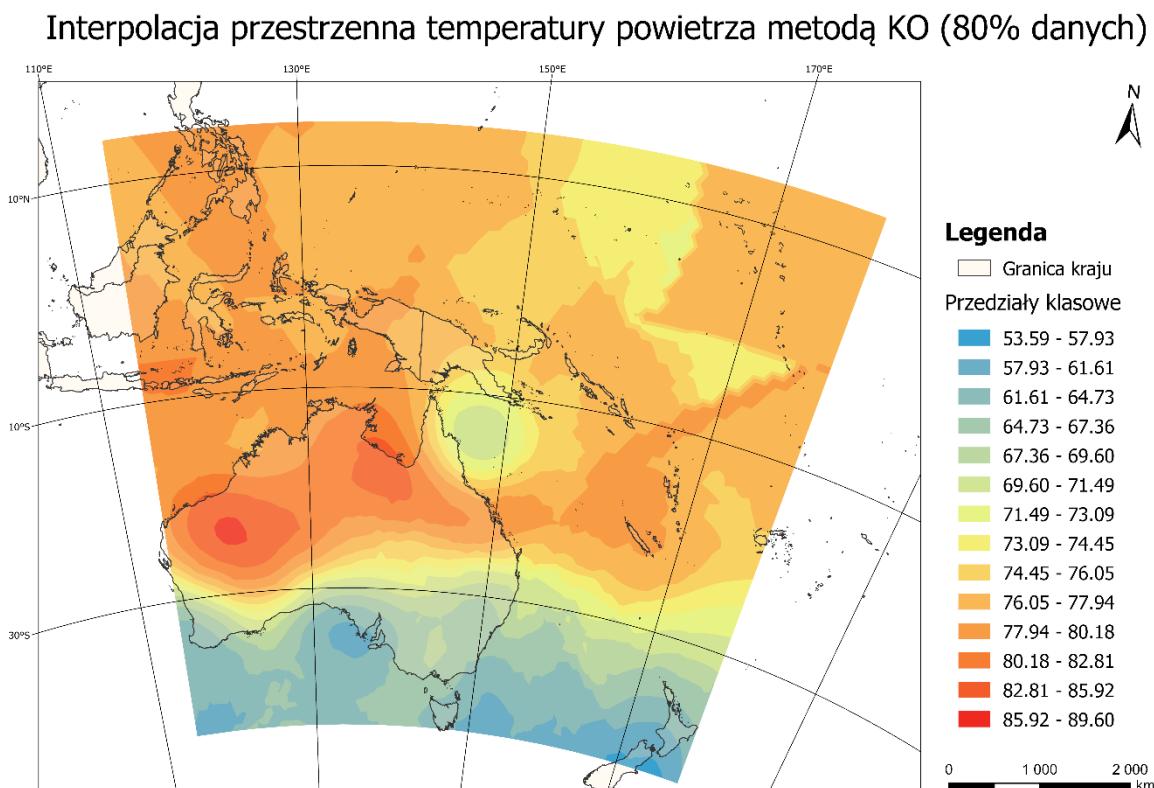


Źródło: opracowanie własne

Pierwszą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda EBK (Empiryczny Kriging Bayesowski). Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najniższe wartości znajdują się pod równoleżnikiem 30°S , najwyższe w okolicy punktu $20^{\circ}\text{S}, 120^{\circ}\text{E}$. Zauważalna jest anomalia wartości na północno-wschodnim wybrzeżu Australii, która może być spowodowana występowaniem łańcucha górskiego Wielkich Górz Wododziałowych, biegnących wzduż wschodniej granicy Australii). Wysoka wartość temperatury w północno-zachodniej części Australii spowodowana może być występowaniem na tym obszarze rozległych obszarów pustynnych. Analizując rysunek powyżej można zauważać skoki pomiędzy przedziałami. Wartość RMSE EBK wynosi 3,82.

b) Kriging Ordinary

Rysunek 35 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Ordinary

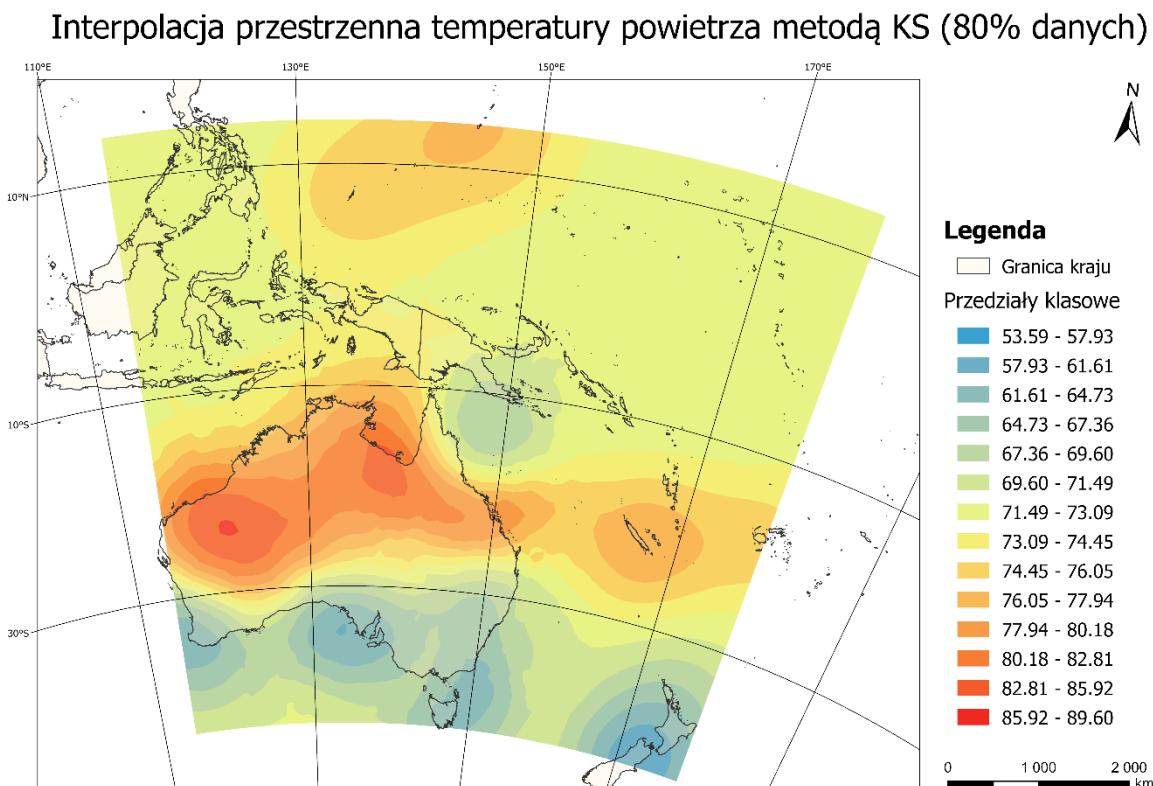


Źródło: opracowanie własne

Drugą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Kriging Ordinary (Kriging zwykły). Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najwyższe wartości temperatur znajdują się w północnej części Australii. Najniższe wartości znajdują się na obszarze Australii. Duże nagromadzenie i nachodzenie się obszarów z innych przedziałów klasowych zauważalne jest w północno-wschodniej części analizowanego obszaru. Granice pomiędzy przedziałami na tym obszarze są ostre. Wartość RMSE Kriging Ordinary wynosi 3,40. W przypadku tej metody anomalia nie jest tak silna w porównaniu do metody EBK, dodatkowo cały obszar badawczy jest silnie zróżnicowany, a w przypadku metody EBK północna część należy głównie do jednego przedziału klasowego.

c) Kriging Simple

Rysunek 36 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Simple

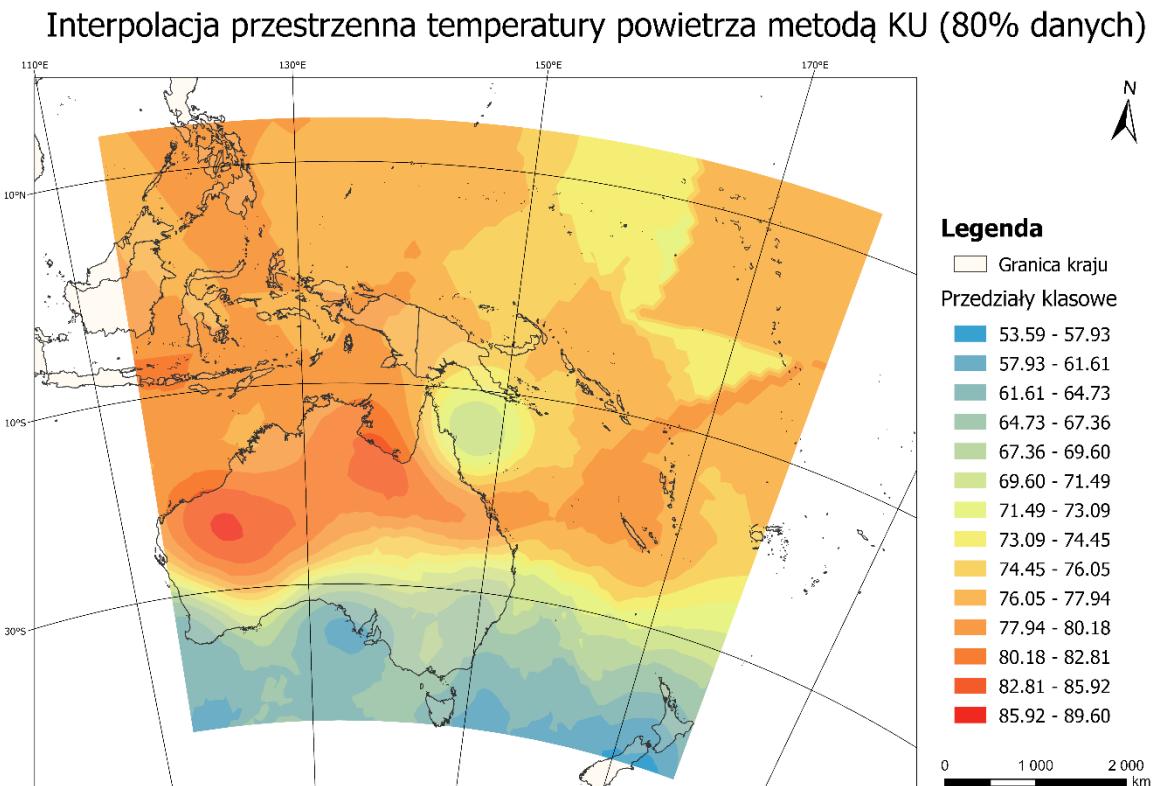


Źródło: opracowanie własne

Trzecią metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Kriging Simple. Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. W przypadku tej metody rozpiętość barw na mapie jest mniejsza w porównaniu do innych metod wykorzystujących ten zbiór danych. Większość obszaru należy do 3 – 4 przedziałów klasowych. Wartości w powyższych analizach zostały sklasyfikowane do innych przedziałów niż w powyższych analizowanych metodach. Analizując mapę można zauważać duże obszary znajdujące się w obrębie jednego przedziału klasowego. Największe zróżnicowanie temperaturowe zauważalne jest w centralnej części analizowanego obszaru. Najwyższe wartości znajdują się na obszarze Australii północno-zachodniej. Wartość RMSE Kriging Simple wynosi 3,66.

d) Kriging Universal

Rysunek 37 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Universal

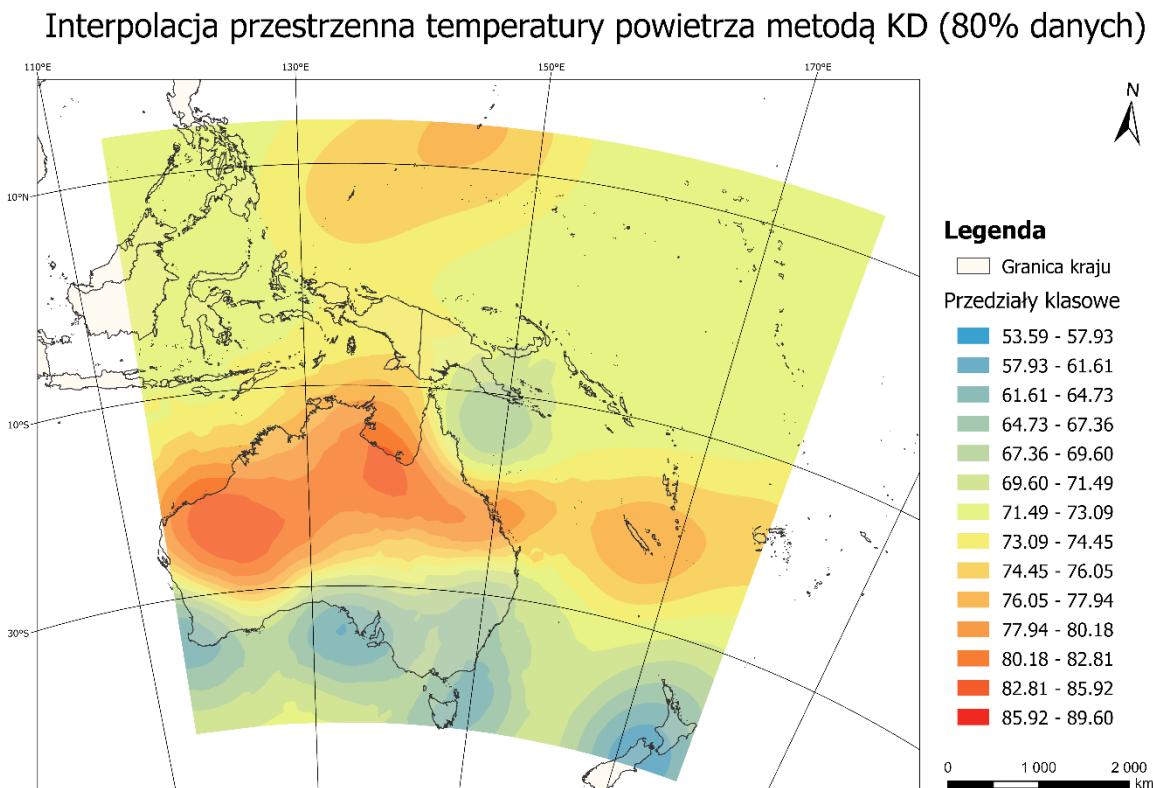


Źródło: opracowanie własne

Czwartą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Kriging Universal. Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najwyższe wartości temperatury znajdują się w Australii północno-zachodniej, stanowiąc duży obszar koncentracji jednego przedziału. Analizując interpolację przestrzenną przedstawioną na rysunku powyżej można zauważyc dużą powierzchnię anomalii występującą na obszarze Papua-Nowa Gwinei. Wartość RMSE Kriging Universal wynosi 3,40 czyli tyle samo co w przypadku metody Kriging Ordinary. Obie metody w przypadku tego zbioru posiadają najniższy wskaźnik RMSE. Wizualizacje mapowe obu metod są bardzo zbliżone.

e) Kriging Disjunctive

Rysunek 38 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Disjunctive

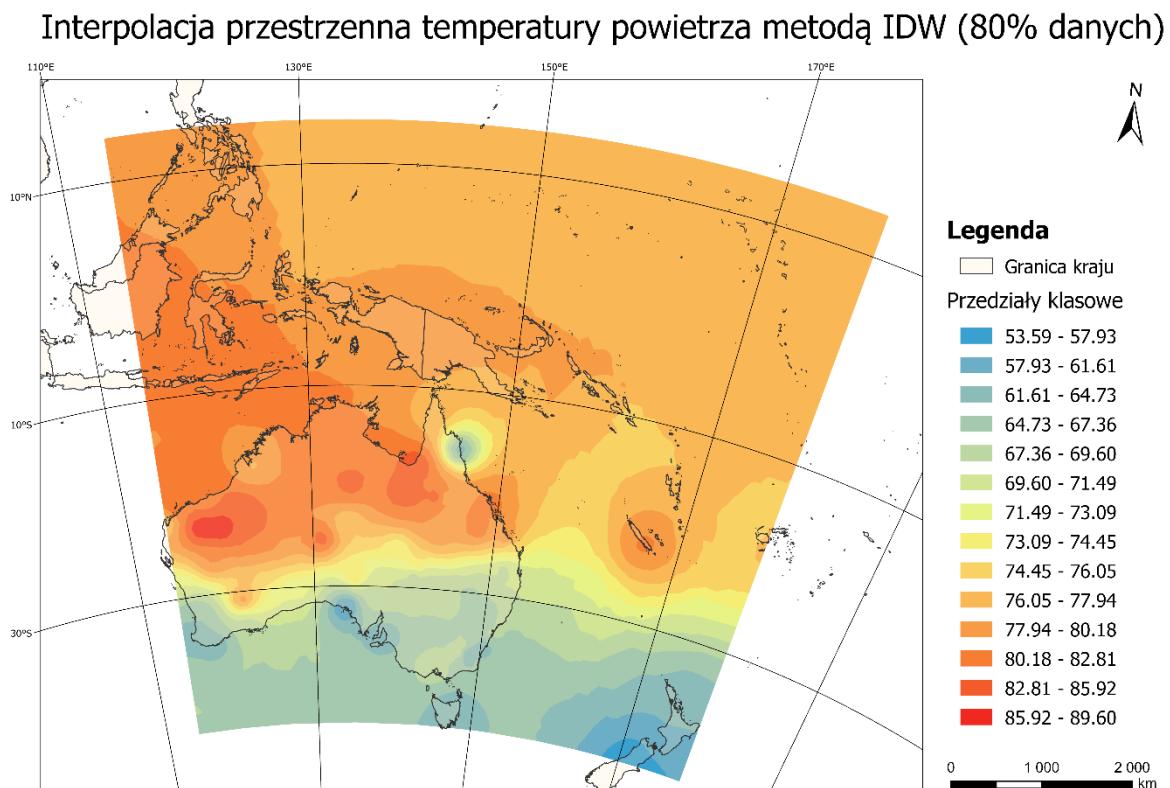


Źródło: opracowanie własne

Piątą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Kriging Disjunctive. Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najniższe wartości znajdują się w północno – wschodniej części analizowanego obszaru, natomiast najwyższe w południowej części analizowanego obszaru. Analizując interpolację przestrzenną przedstawioną na rysunku 12 można zauważyć, występowanie wyspy wartości wysokich temperatur na obszarze Australii północnej. Wartość RMSE Kriging Disjunctive wynosi 4,22, co jest szóstym najwyższym wynikiem wśród analizowanych 7 metod interpolacji przestrzennej.

f) Inverse Distance Weighting

Rysunek 39 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą IDW

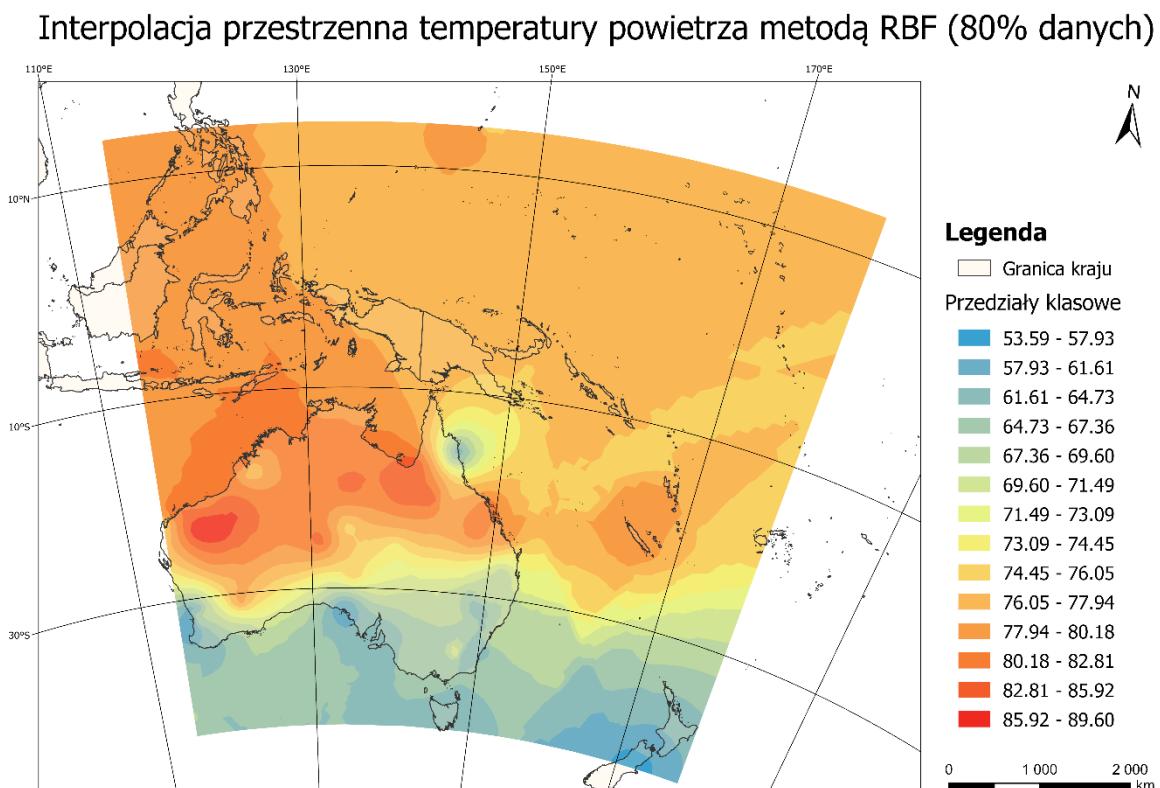


Źródło: opracowanie własne

Szóstą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Inverse Distance Weighting (IDW, metoda odwrotnej odległości). Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najniższe wartości znajdują się na obszarach znajdujących się pod równoleżnikiem 30°S, natomiast najwyższe w północno – zachodniej części analizowanego obszaru. W przypadku tej metody cała część obszaru powyżej równoleżnika 30°S zakwalifikowana została do przedziałów o wysokich temperaturach. Wyjątkiem jest anomalia na wschodnim wybrzeżu Australii, gdzie występuje punktowy obszar o niższej temperaturze w porównaniu do okolicznych obszarów. Negatywną cechą IDW jest tendencja do tworzenia tzw. efektów oczkowych (SZYMANOWSKI 2014). Analizując interpolację przestrzenną przedstawioną na rysunku powyżej można zauważyc, że występują duże obszary należące do jednego przedziału, przerywane kilkoma punktowymi obszarami o wyższej lub niższej klasie temperatury. Wartość RMSE IDW wynosi 3,56.

g) Radial Basis Functions

Rysunek 40 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą RBF



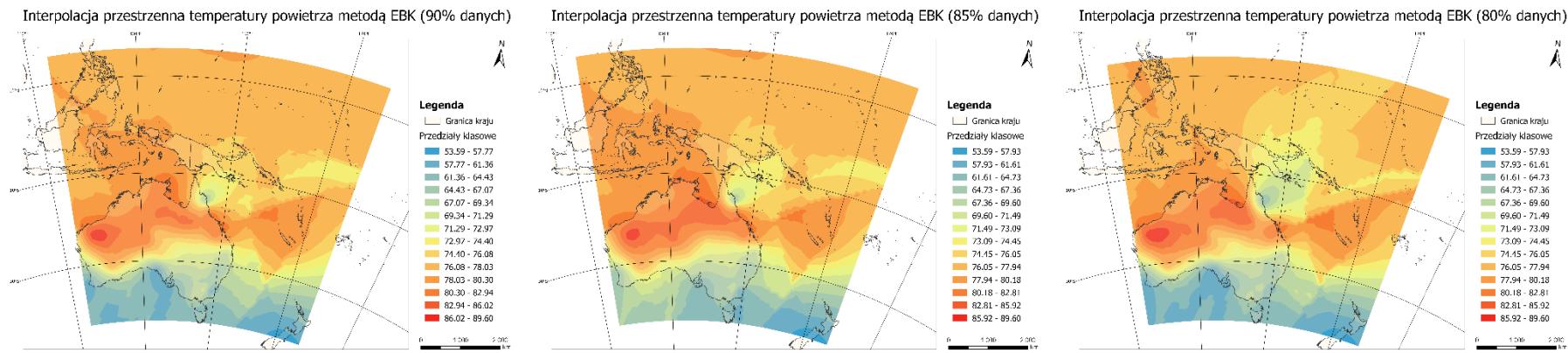
Źródło: opracowanie własne

Siódmą metodą wykorzystaną w badaniu interpolacji przestrzennej była metoda Radial Basis Functions (RBF, Radialna Funkcja Bazowa). Powyższa mapa przedstawia wyniki przeprowadzonej estymacji. Najniższe wartości znajdują się pod równoleżnikiem 30°S. Analizując interpolację przestrzenną przedstawioną na rysunku powyżej można zauważyc, że znaczna część obszaru sklasyfikowana została do przedziału 76,05-77,94, bądź 77,94-80,18. Na mapie widoczne jest wiele obszarów różnych klas oraz występują silne załamania na granicach tych obrębów. Wartość RMSE RBF wynosi 3,73, co jest przedostatnią najwyższą wartością wśród analizowanych metod dla tego zbioru.

Analizując wszystkie interpolacje przestrzenne można zauważyc, że tam gdzie sieć punktów jest gęstsza (obszar Australii, w szczególności południowo-zachodniej) otrzymany obraz jest bardziej zróżnicowany, w obszarze tym występuje wiele klas, a przejście między nimi są wyraźnie widoczne, natomiast mniejsza ilość punktów bazowych (obszar morskie) charakteryzuje się zmniejszoną liczbą klas.

4. Zestawienie metod wykorzystujących podzbiory

4.1. Empirical Bayesian Kriging

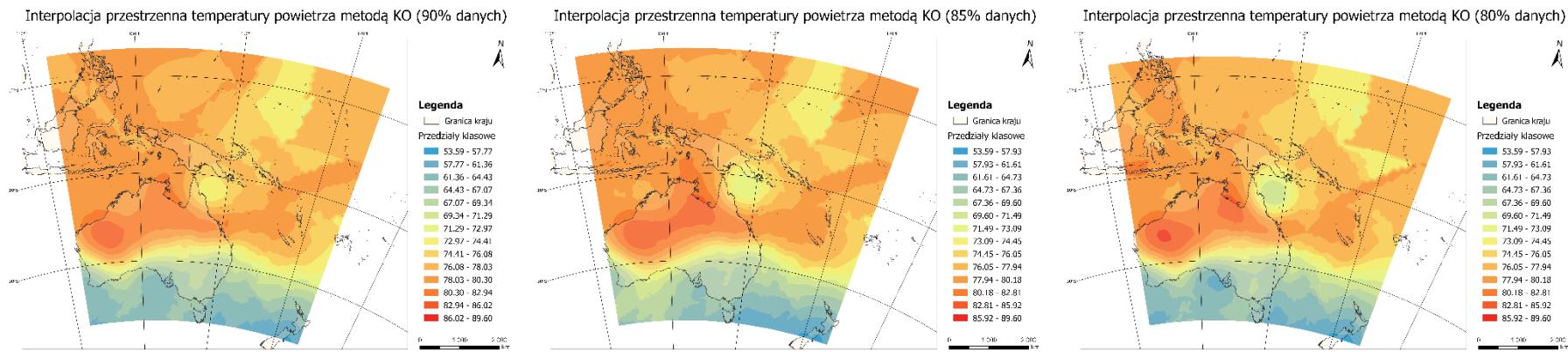


Rysunek 41 Zestawienie trzech interpolacji przestrzennych wykonanych metodą EBK wykorzystujących odpowiednio 90%, 85%, 80% danych

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

Analizując przedstawione mapy interpolacji przestrzennej wykonane w oparciu o trzy zbiory danych można zauważyć, że są one zbliżone, jednak w interpolacji przestrzennej dla 80% danych anomalia występująca na północno-wschodnim wybrzeżu jest większa powierzchniowo. Dodatkowo północno-wschodnia część analizowanego obszaru została sklasyfikowana do niższego przedziału wartości temperatur. W interpolacji wykorzystującej 90% danych metoda EBK stanowi jedną z najlepszych metod, pod względem wartości RMSE, natomiast w interpolacji wykorzystującej 85% danych metoda jest na 4 miejscu, a w interpolacji wykorzystującej 80% danych metoda ta jest najgorsza – wartość RMSE jest najwyższa.

4.2. Kriging Ordinary

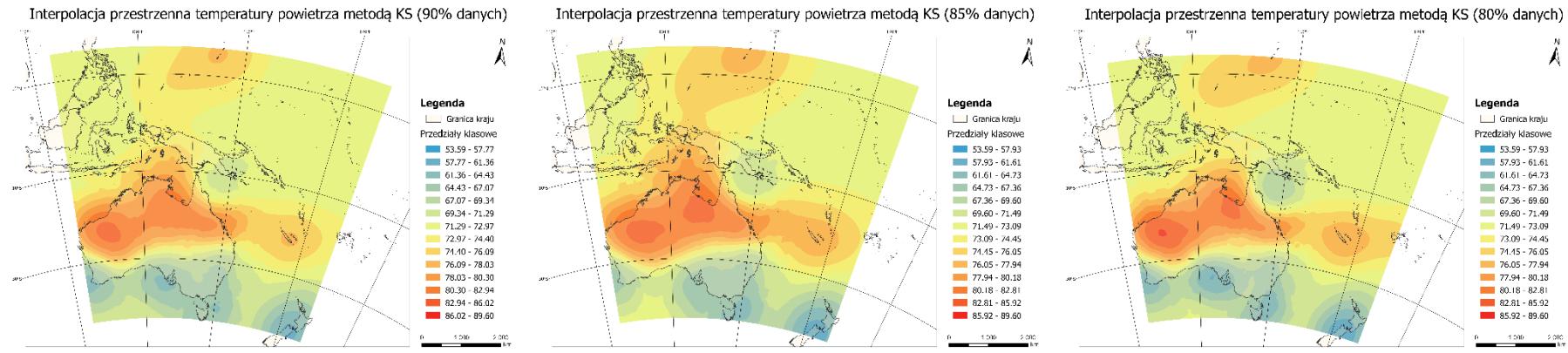


Rysunek 42 Zestawienie trzech interpolacji przestrzennych wykonanych metodą KO wykorzystujących odpowiednio 90%, 85%, 80% danych

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

Analizując przedstawione mapy interpolacji przestrzennej wykonane w oparciu o trzy zbiory danych można zauważyć, że wizualnie są one do siebie zbliżone, jednak w interpolacji przestrzennej dla 80% danych anomalia występująca na północno-wschodnim wybrzeżu posiada wartość temperatury z niższego przedziału. Dodatkowo wizualizacje wykorzystujące 85% i 80% danych posiadają szerszy zakres wykorzystanych przedziałów – skala barw na wizualizacji jest większa. W interpolacji wykorzystującej 90% danych metoda KO stanowi jedną z najmniej odpowiednich metod, natomiast w interpolacji wykorzystującej 85% i 80% danych metoda ta jest jedną z najlepszych – wartość RMSE jest najwyższa.

4.3. Kriging Simple

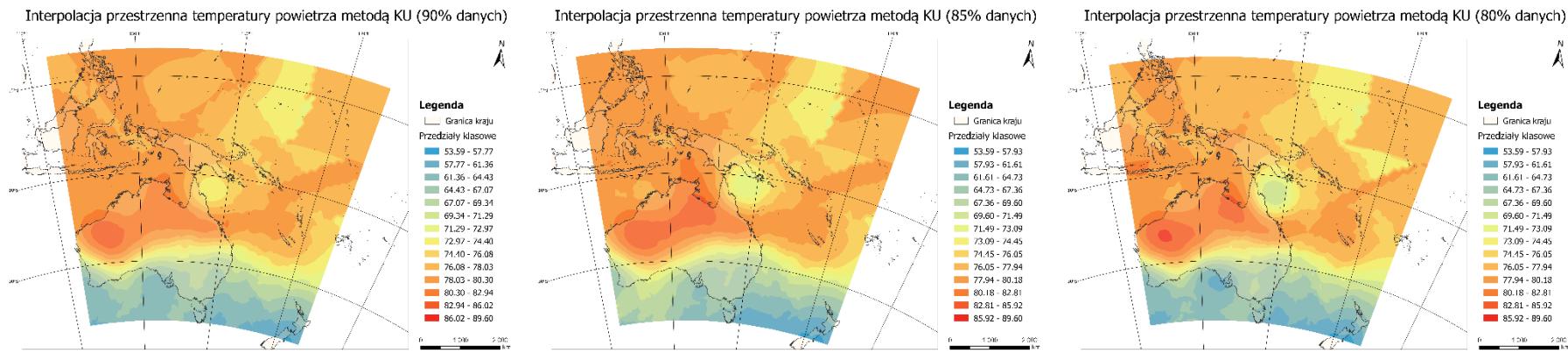


Rysunek 43 Zestawienie trzech interpolacji przestrzennych wykonanych metodą KS wykorzystującymi odpowiednio 90%, 85%, 80% danych

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

Analizując przedstawione mapy interpolacji przestrzennej wykonane w oparciu o trzy zbiory danych można zauważyc, że wizualnie nieco różnią się od siebie – w szczególności w północnej części analizowanego obszaru oraz na obszarze Australii. W ostatniej wizualizacji wykorzystującej 80% danych zastosowanie ma więcej przedziałów klasowych, a granice są bardziej ostre i nierównomierne. W interpolacji wykorzystującej 90% i 80% danych metoda KS znalazła się na 4 miejscu w kolejności pod względem wartości RMSE, natomiast w interpolacji wykorzystującej 80% danych metoda ta jest drugą z najlepszych metod.

4.4. Kriging Universal

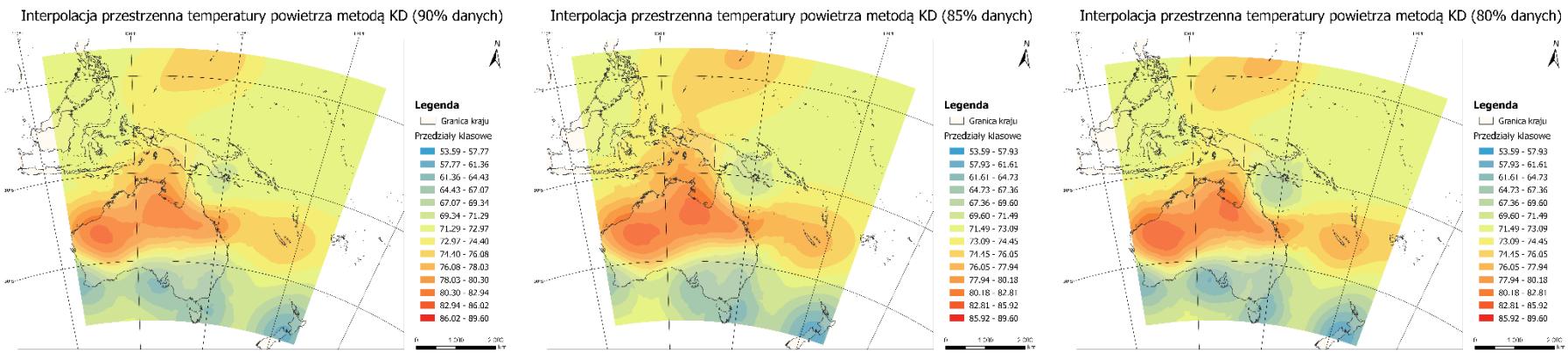


Rysunek 44 Zestawienie trzech interpolacji przestrzennych wykonanych metodą KU wykorzystującymi odpowiednio 90%, 85%, 80% danych

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

Analizując przedstawione mapy interpolacji przestrzennej wykonane w oparciu o trzy zbiory danych można zauważyc, że wizualnie w pewnej mierze różnią się od siebie – w szczególności w północno -wschodniej i centralnej części analizowanego obszaru. W ostatniej wizualizacji wykorzystującej 80% danych zastosowanie ma więcej przedziałów klasowych, a granice są bardziej ostre i nierównomierne. W drugiej i trzeciej wizualizacji zlokalizowanych jest więcej obszarów o małej powierzchni, które ukazują większe zróżnicowanie wartości temperatury na danym obszarze. W interpolacji wykorzystującej 90% danych metoda KU znalazła się na drugim miejscu w kolejności pod względem wartości RMSE, natomiast w interpolacji wykorzystującej 85% i 80% danych metoda ta jest jedną z najlepszych metod – wartość wskaźnika RMSE jest najwyższa.

4.5. Kriging Disjunctive

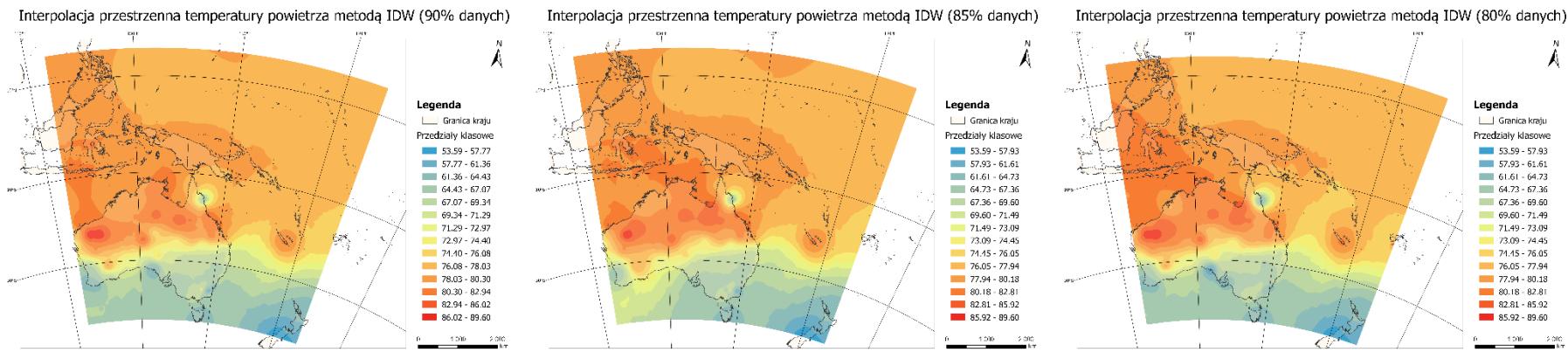


Rysunek 45 Zestawienie trzech interpolacji przestrzennych wykonanych metodą KD wykorzystującymi odpowiednio 90%, 85%, 80% danych

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

Analizując przedstawione mapy interpolacji przestrzennej wykonane w oparciu o trzy zbiory danych można wyraźnie zauważyc, że w porównaniu do powyższych metod, w tej metodzie wykorzystana rozpiętość kolorystyczna nie jest duża, większość z analizowanego obszaru znajduje się w 3-4 środkowych przedziałach wyznaczonych wartości temperatur. Mapy w niewielkim stopniu różnią się od siebie wyłącznie w środkowej części analizowanego obszaru. Ostrość granic we wszystkich wizualizacjach jest niewielka, przejście między poszczególnymi warstwami zachodzi spokojnie. We wszystkich interpolacjach metoda KU znalazła się na trzecim (środkowym) miejscu w kolejności pod względem wartości RMSE, oznacza to, że na metodę tę zmienność danych miała znikomy wpływ.

4.6. Inverse Distance Weighting

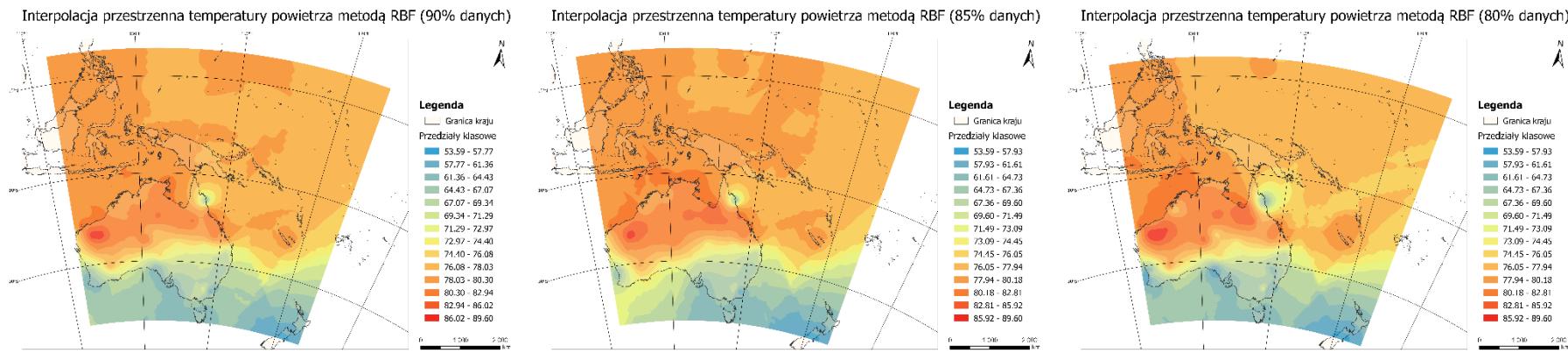


Rysunek 46 Zestawienie trzech interpolacji przestrzennych wykonanych metodą IDW wykorzystujących odpowiednio 90%, 85%, 80% danych

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

Analizując przedstawione mapy interpolacji przestrzennej wykonane w oparciu o trzy zbiory danych można zauważyc, że wizualizacje wykorzystujące 90% i 80% są do siebie silnie zbliżone, co potwierdza pozycje w rankingu uszeregowanym rosnaco według wartosci wskaźnika RMSE. Metoda ta wykorzystując 90% i 80% znalazła się na drugim miejscu. Co oznacza, że jest jedną z lepszym metod w przypadku tego zbioru danych dla tych zbiorów, jednak metoda wykorzystując 85% danych znalazła się na 5 miejscu, czyli jednym z najmniej odpowiednich metod. Wizualizacja wykonana w oparciu o zbiór 85% danych różni się w szczególności w północnej, zachodniej i południowo-zachodniej części analizowanego obszaru. Liczba tzw. „oczek” w tej wizualizacji w porównaniu do dwóch pozostałych jest większa, a ich powierzchnia jest mniejsza.

4.7. Radial Basis Functions



Rysunek 47 Zestawienie trzech interpolacji przestrzennych wykonanych metodą RBF wykorzystujących odpowiednio 90%, 85%, 80% danych

Źródło: opracowanie własne w oprogramowaniu ArcGIS Pro

Analizując przedstawione mapy interpolacji przestrzennej wykonane w oparciu o trzy zbiory danych można zauważyc, że wizualizacje znaczaco się od siebie różnią. Różnica zauważalna jest w szczególności w północnej i centralnej części analizowanego obszaru. W przypadku drugiej metody wyraźnie widoczne są obszary o silnie nieregularnym kształcie. Metoda ta, wykorzystująca 85% danych pod względem wartości wskaźnika RMSE znalazła się na ostatnim miejscu – błąd spośród wszystkich analizowanych metod dla tego zbioru jest największy. Podobnie jest w przypadku wizualizacji wykorzystującej 80% danych, która pod względem wartości wskaźnika RMSE znalazła się na przedostatnim miejscu. Analizując wizualizację interpolacji przestrzenna temperatury powietrza metodą RBF wykorzystująca 90% danych, można zauważyc, że przejścia między przedziałami zachodzą spokojniej. Wartość wskaźnika RMSE dla tej metody i tego zbioru jest jedną z najniższych, co oznacza, że wizualizacja możliwie najlepiej oddaje rzeczywisty rozkład temperatury na obszarze Australii i Oceanii.

5. Wybór optymalnej metody interpolacji

Obecne możliwości modelowania przestrzeni, a co za tym idzie wyższa jakość analiz, możliwe są dzięki rozwojowi geotechnologii integrującej systemy informacji przestrzennej, teledetekcję oraz globalne systemy pozycjonowania (GPS). Zwiększyły dostęp i ilość danych sprawiają, że wykonanie map staje się stosunkowo łatwe, jednak zwiększa się ryzyko nieprecyzyjnego i nieprawidłowego sposobu przedstawiania danych. Modelowanie przestrzenne nie odzwierciedla wiernie rzeczywistości, jednak niezbędne jest dążenie do ograniczenia błędów aproksymacji wraz z właściwą, solidną oceną wyników i przedstawienia stopnia niepewności. Obecnie interpolacja przestrzenna posiada wiele różnych metod, co skutkować może trudnością w wyborze tej odpowiedniej do danego zbioru danych. Wybór optymalnej metody interpolacji powinien być poprzedzony etapem porównawczym, bądź badaniami procedurą selekcji. Wybór optymalnej metody najczęściej dokonywany jest poprzez wskazanie metody, której wyniki obarczone są najmniejszymi błędami (SZYMANOWSKI 2014).

W ramach badania przeprowadzono analizę wyboru optymalnej metody za pomocą wskaźnika RMSE (ang. root mean square error), czyli średniej kwadratowej błędów. RMSE jest wskaźnikiem zalecany do porównania wyników analizy przestrzennej za pomocą różnych metod interpolacji. Jego wartość powinna być jak najmniejsza (Ogryzek, Kurowska, 2016).

Wartość wskaźnika RMSE dla poszczególnych metod obliczono wykorzystując wzór:

$$RMSE = \sqrt{MSE}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - \tilde{x})^2$$

Analizując dane zawarte w tabeli 1 w przypadku zbioru 90% danych najbardziej optymalnymi metodami interpolacji pod względem wartości wskaźnik RMSE są metody EBK oraz RBF. W przypadku zbioru danych 85% i 80% najwyższym błędem RMSE charakteryzuje się metoda Kriging Ordinary oraz Kriging Universal, oznacza to, że metody te w przypadku badania danych pozyskanych z punktów pomiarowych temperatury powietrza dla Australii i Oceanii są najbardziej optymalne. Wartości wskaźnika RMSE dla poszczególnych metod i zbiorów danych przedstawiono w tabeli 1.

Tabela 1 Wskaźniki RMSE analizowanych metod interpolacji przestrzennej

90-10	RMSE
EBK	2,70
KO	2,72
KS	3,14
KU	2,72
KD	3,12
IDW	2,84
RBF	2,70

85-15	RMSE
EBK	5,82
KO	5,58
KS	5,77
KU	5,58
KD	5,80
IDW	6,01
RBF	6,47

80-20	RMSE
EBK	3,82
KO	3,40
KS	3,66
KU	3,40
KD	3,61
IDW	3,56
RBF	3,73

Źródło: opracowanie własne

6. Podsumowanie i wnioski

Celem przeprowadzanej analizy był wybór optymalnej metody interpolacji przestrzennej na podstawie wartości wskaźnika RMSE. W badaniu wykorzystano 7 metod: 2 deterministyczne (IDW, RBF) oraz 5 stochastycznych (KO, KS, KU, KD, EBK). Meteorologicznymi danymi wejściowymi w opracowaniu był zbiór pomiarów temperatury powietrza w Australii i Oceanii ze 101 stacji badawczych. Spośród pełnego zestawu danych wydzielono podzbiory 10%, 15% i 20%, dzięki którym przeprowadzono walidację podzbiorem. Dla zbiorów danych 90%, 85%, 80% dokonano analizy zmienności temperatur z użyciem metod interpolacji przestrzennej.

Każdy przypadek powinien być rozpatrywany indywidualnie, a metoda obarczona najmniejszymi błędami powinna być uznana za optymalną. Ocena wartości RMSE oraz ocena wizualna wyników modelowania zdecydowała o wyborze optymalnej metody. Ważną przestrzenną zależnością, którą należy brać pod uwagę jest pierwsze prawo geografii (reguła Toblera 1970) „wszystkie obiekty są ze sobą powiązane, a siła tych powiązań maleje wraz ze wzrostem odległości między nimi”.

Metody interpolacji przestrzennej zostały wypracowane dla zastosowań w różnych dyscyplinach i z przeznaczeniem dla różnorodnych zbiorów danych i zmiennych. Takie czynniki, jak: wielkość zbioru wejściowego, rozmieszczenie przestrzenne danych, cel interpolacji i charakter interpolowanej zmiennej wpływają na wybór odpowiedniej metody (Szymanowski i in. 2012).

W przypadku rozkładu temperatury wpływ mają czynniki środowiskowe warunkujące jej rozkład. Zależność ta jest statystycznie istotna. Najistotniejszym elementem mogącym wpływać na wartość pomiaru temperatury jest wysokość bezwzględna i położenie geograficzne, wpływ ma również odległość od morza. Miejscowe odchylenia spowodowane są również napływem wilgotnych mas powietrza i pasatami.

Subiektywna interpretacja wizualna uzyskanych wyników jest niezmiernie ważna, bazując na tzw. wiedzy eksperckiej i wizualnym oglądzie wyników możliwa jest detekcja mało realistycznych wyników, jak: zerwanie ciągłości przestrzennej, niespodziewane duże bądź małe wartości modelowanej zmiennej, zwłaszcza w obszarach ekstrapolacji, silne tendencje kierunkowe lub regionalne. Jednak kluczowe znaczenie ma obiektywna, ilościowa ocena wyników, a ścisłe błędy modelowania. Walidacja podzbiorem pozwala na oszacowanie

wielkości błędów. Wielkość błędów CV pozwala na określenie możliwych tendencji regionalnych do przeszacowania lub niedoszacowania wielkości temperatury przez model oraz identyfikacji punktów wielkości odstających. W przypadku pomiarów temperatury powietrza należy wziąć również pod uwagę tendencje regionalne przeszacowania modelu, a więc przypadki, gdy dla grupy stacji sąsiadujących ze sobą obserwowane są wysokie dodatnie wielkości błędu CV. Mogą one występować, gdy stacje te charakteryzują się podobnymi cechami lokalizacji, a model matematyczny nie jest wystarczająco efektywny do prawidłowego określenia wielkości temperatury. Model może wykazywać także regionalną tendencję do zniżenia temperatury w stosunku do wielkości obserwowanych (SZYMANOWSKI I IN. 2012).

W przypadku estymacji geostatystycznej podstawowymi etapami jest:

1. Zebranie danych i ich analiza
2. Wstępna eksploracja danych oraz ocena ich jakości
3. Opracowanie wskaźników geostatystycznych, m.in. histogramu i semiwariogramu
4. Estymacja badanej cechy
5. Porównanie i ocena modeli oraz wybór najbardziej optymalnej

Podsumowując w badaniach interpolacji przestrzennej występuje mnogość metod i ich kombinacji, nie ma więc uniwersalnej metody interpolacji przestrzennej, jednak możliwy jest wybór optymalnej metody. Każdy przypadek powinien być rozpatrywany indywidualnie, a wybór optymalnej metody powinien być dokonany po przeprowadzaniu analizy jakości estymacji.

Literatura

- Cichociński, P. (2011). Porównanie metod interpolacji przestrzennej w odniesieniu do wartości nieruchomości. *Studia i Materiały Towarzystwa Naukowego Nieruchomości*, 19(3), 120-125.
- Gil, M., Frąckiewicz, P., Pukanská, K., & Blišťan, P. (2018). Statystyczne i geostatystyczne metody badania deformacji powierzchni.
- Krasowska, M., & Banaszuk, P. (2013). Wykorzystanie analizy semiwariancji w ocenie zanieczyszczenia wód rzecznych w krajobrazie rolniczym. *Rocznik Ochrona Środowiska*, 15(cz. 2), 1705-1718.
- Kwiatkowska-Malina, J., & Borkowski, A. S. (2014). Wykorzystanie analizy semiwariancji w oszacowaniu stężenia SO₂ w powietrzu atmosferycznym. *Infrastruktura i Ekologia Terenów Wiejskich*, (II/2).
- Marmol, U. (2002). Modelowanie reprezentacji powierzchni topograficznej z wykorzystaniem metody geostatystycznej. *Geodezja*, 8(2), 259-270.
- Ogryzek, M. P. (2018). Parametryczna ocena jakości estymacji map opracowywanych metodami geostatystycznymi. *Studia i Prace WNEiZ US*, (54/3), 319-330.
- Ogryzek, M., & Kurowska, K. (2016). Geostatystyczne metody opracowywania map średnich cen transakcyjnych gruntów rolnych niezabudowanych. *Studia i Prace WNEiZ US*, 45, 397-408.
- Ogryzek, M., & Kurowska, K. (2016). Opracowanie map średnich cen transakcyjnych gruntów rolnych niezabudowanych zbywanych z zasobu ANR OT Olsztyn z wykorzystaniem interpolacji metodą IDW. *Studia i Prace WNEiZ US*, 45, 355-366.
- Szymanowski, M. (2014). Autoreferat
- Szymanowski M., Kryza M., 2011b: Modele regresji i ich rozszerzenia w przestrzennej interpolacji danych [w:] Żyszkowska W., Spallek W. (red.), Główne problemy współczesnej kartografii 2011. Zastosowanie statystyki w GIS i kartografii. Uniwersytet Wrocławski, Wrocław, 174–196.
- Szymanowski, M., & Kryza, M. (2012). Zastosowanie krigingu reszt lokalnego modelu regresji do interpolacji przestrzennej średniej rocznej temperatury powietrza w Polsce. *Przegląd Geofizyczny*, (1), 75-84.
- Szymanowski, M., Kryza, M., & Spallek, W. (2012). Atlas temperatury powietrza w Polsce: studium metodyczne. Rozprawy Naukowe Instytutu Geografii i Rozwoju Regionalnego Uniwersytetu Wrocławskiego 26

- Urbański, J. (2012). GIS w badaniach przyrodniczych. *University of Gdańsk, Gdańsk*.
- Ustrnul, Z. (2004). Metody analizy przestrzennej w badaniach klimatologicznych. A. Bokwa, Z. Ursnul (red.), *Zastos. wybranych Metod w klimatologii*, Inst. Geogr. i Gospod. Przestrz. UJ, 65-88.
- Zasina, D. K. (2019). *Analiza zmienności i niepewności szacowania emisji zanieczyszczeń ze spalania paliw w gospodarstwach domowych metodami geostatystycznymi* (Rozprawa doktorska, Politechnika Warszawska, Wydział Instalacji Budowlanych, Hydrotechniki I Inżynierii Środowiska)

Źródła internetowe

- <https://www.statystyka-zadania.pl/histogram>
- <https://idane.pl/minipodrecznik/testy-rozkadu.html>
- https://www.naukowiec.org/wiedza/statystyka/skosnosc_714.html
- https://www.naukowiec.org/wiedza/statystyka/kurtoza_698.html
- <https://mfiles.pl/pl/index.php/Kurtoza>
- <https://www.wbc.poznan.pl/Content/382516/geostatystyka-prolog.html>
<https://www.wbc.poznan.pl/Content/382516/geostatystyka-prolog.html>
- <https://bookdown.org/nowosad/geostatystyka/metody-interpolacji.html>
- <http://www.wbc.poznan.pl/Content/382516/estymacje-jednozmienne.html>
- <https://desktop.arcgis.com/fr/arcmap/latest/extensions/geostatistical-analyst-understanding-ordinary-kriging.htm>
- <http://www.wbc.poznan.pl/Content/382516/estymacje-jednozmienne.html>
- <https://desktop.arcgis.com/en/arcmap/10.5/tools/geostatistical-analyst-toolbox-empirical-bayesian-kriging.htm>
- <https://www.wbc.poznan.pl/Content/382516/ocena-jakosci-estymacji.html>

Spis tabel

Tabela 1 Wskaźniki RMSE analizowanych metod interpolacji przestrzennej 53

Spis rysunków

Rysunek 1 Podstawowe pojęcia w analizie wariogramu	7
Rysunek 2 Położenie punktów pomiarowych temperatury powietrza w Australii i Oceanii dla zbiorów 90% i 10%	12
Rysunek 3 Histogram temperatury w Australii i Oceanii w 2022 roku – 90 % danych	13
Rysunek 4 Wartości statystyczne histogramu temperatury powietrza dla zbioru 90%	13
Rysunek 5 Wykres rozkładu normalnego temperatury powietrza w Australii i Oceanii w 2022 roku dla 90% danych.....	14
Rysunek 6 Semiwariogram dla analizowanych 90% danych.....	15
Rysunek 7 Mapa semiwariancji dla analizowanych 90% danych.....	15
Rysunek 8 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą EBK (90% danych).....	16
Rysunek 9 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Ordinary (90% danych).....	17
Rysunek 10 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Simple (90% danych)	18
Rysunek 11 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Universal (90% danych).....	19
Rysunek 12 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Disjunctive (90% danych).....	20
Rysunek 13 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą IDW (90% danych).....	21
Rysunek 14 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą RBF.....	22
Rysunek 15 Położenie punktów pomiarowych temperatury powietrza w Australii i Oceanii dla zbiorów 85% i 15%	23
Rysunek 16 Histogram temperatury w Australii i Oceanii w 2022 roku – 85 % danych	24
Rysunek 17 Wartości statystyczne histogramu temperatury powietrza dla zbioru 85%	24
Rysunek 18 Wykres rozkładu normalnego temperatury powietrza w Australii i Oceanii w 2022 roku dla 85% danych.....	25
Rysunek 19 Semiwariogram dla analizowanych 85% danych.....	26
Rysunek 20 Mapa semiwariancji dla analizowanych 85% danych.....	26
Rysunek 21 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą EBK	27
Rysunek 22 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Ordinary	28
Rysunek 23 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Simple	29
Rysunek 24 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Universal	30
Rysunek 25 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Disjunctive	31
Rysunek 26 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą IDW	32

Rysunek 27 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą RBF.....	33
Rysunek 28 Położenie punktów pomiarowych temperatury powietrza w Australii i Oceanii dla zbiorów 80% i 20%	34
Rysunek 29 Histogram temperatury w Australii i Oceanii w 2022 roku – 80 % danych	35
Rysunek 30 Wartości statystyczne histogramu temperatury powietrza dla zbioru 80%	35
Rysunek 31 Wykres rozkładu normalnego temperatury powietrza w Australii i Oceanii w 2022 roku dla 80% danych.....	36
Rysunek 32 Semiwariogram dla analizowanych 80% danych.....	37
Rysunek 33 Mapa semiwariancji dla analizowanych 80% danych.....	37
Rysunek 34 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą EBK	38
Rysunek 35 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Ordinary	39
Rysunek 36 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Simple	40
Rysunek 37 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Universal	41
Rysunek 38 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą Kriging Disjunctive	42
Rysunek 39 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą IDW	43
Rysunek 40 Interpolacja przestrzenna punktów temperatury powietrza w Australii i Oceanii metodą RBF.....	44
Rysunek 41 Zestawienie trzech interpolacji przestrzennych wykonanych metodą EBK wykorzystujących odpowiednio 90%, 85%, 80% danych	45
Rysunek 42 Zestawienie trzech interpolacji przestrzennych wykonanych metodą KO wykorzystujących odpowiednio 90%, 85%, 80% danych	46
Rysunek 43 Zestawienie trzech interpolacji przestrzennych wykonanych metodą KS wykorzystujących odpowiednio 90%, 85%, 80% danych	47
Rysunek 44 Zestawienie trzech interpolacji przestrzennych wykonanych metodą KU wykorzystujących odpowiednio 90%, 85%, 80% danych	48
Rysunek 45 Zestawienie trzech interpolacji przestrzennych wykonanych metodą KD wykorzystujących odpowiednio 90%, 85%, 80% danych	49
Rysunek 46 Zestawienie trzech interpolacji przestrzennych wykonanych metodą IDW wykorzystujących odpowiednio 90%, 85%, 80% danych	50
Rysunek 47 Zestawienie trzech interpolacji przestrzennych wykonanych metodą RBF wykorzystujących odpowiednio 90%, 85%, 80% danych	51