

## Streszczenie

Inspiracją badań zaprezentowanych w niniejszej monografii były obserwacje poczynione w trakcie prac w ramach projektu dotyczącego przetwarzania danych z lotniczego skaningu laserowego (Hejmanowska i in. 2008a). Produktem końcowym był NMT (numeryczny model terenu), którego jakość próbowano oszacować. Dobierano różne oczka siatki i analizowano różnice pomiędzy NMT a chmurą punktów ze skaningu laserowego. Obserwowano między innymi histogramy odchyłek pomiędzy NMT a punktami pomiarowymi ALS. Okazało się, że kształt wszystkich histogramów tych odchyłek znacznie odbiega od kształtu funkcji rozkładu normalnego. Histogramy były bowiem silnie skoncentrowane wokół zera i miały „długie ogony”, co kojarzyło się na pierwszy rzut oka z rozkładem Laplace’a.

W wyniku późniejszych studiów literaturowych znaleziono potwierdzenie powyższego zjawiska oraz sformułowano propozycję modyfikacji rozkładu normalnego na potrzeby analiz dokładności NMT tworzonych z wykorzystaniem pomiarów ALS i lotniczej fotogrametrii (Höhle, Höhle 2009). W publikacji tej została zaproponowana tzw. metoda odporna (*robust method*), jako rozwiązanie alternatywne w stosunku do powszechnie wykorzystywanego podejścia, czyli charakteryzowania dokładności za pomocą odchylenia standardowego czy błędu średniego kwadratowego bazującego na rozkładzie normalnym. Po przeanalizowaniu wyników prezentowanych w publikacji, czyli histogramów odchyłek, kształtu odpowiadającej im krzywej Gaussa i przebiegów krzywych ilustrujących proponowaną odporną funkcję gęstości prawdopodobieństwa, nasuwa się wniosek, że krzywa Gaussa znacznie przeszacowuje błąd NMT, ale metoda odporna go nie doszacowuje. Wniosek ten potwierdził sensowność poszukiwań innej funkcji rozkładu gęstości prawdopodobieństwa. Zaproponowany został rozkład Laplace’a, początkowo jedynie na podstawie analizy wizualnej.

Badania jakości NMT, głównie w kontekście wpływu jego dokładności na wynik analiz przestrzennych, prowadzone były od pewnego czasu (Hejmanowska 2005, Hejmanowska i in. 2008a i b). Wpływ ten jest możliwy do uwzględnienia w analizach przestrzennych, jednak żeby o nim mówić, trzeba wiedzieć, w jaki sposób wiarygodnie określić jakość NMT. Badaniom dotyczącym określania jakości modelu NMT poświęcona jest publikacja autorki (Hejmanowska i in. 2008b), w której przedyskutowano różne definicje i podejścia do określania jakości, wiarygodności, dokładności, niedokładności, błędu i niepewności NMT. Za interesującą można uznać klasyfikację zaproponowaną przez Petera Fishera (Fisher 1999), wychodzącą od pojęcia „**niepewność**”, którą w przypadku dobrze zdefiniowanego NMT można byłoby określić jako **błąd**. Jeśli NMT nie został dobrze zdefiniowany, to niepewność mogłaby być opisana takimi pojęciami, jak „**nieokreśloność**” i „**niejednoznaczność**”.

Równocześnie w środowisku geodezyjnym zaczęły się pojawiać informacje na temat zaleceń dotyczących nieużywania pojęcia „**błąd**”, zamiast którego zaproponowano właśnie określenie „**niepewność**”. W roku 1993 opublikowano w wersji angielskiej Przewodnik (polska wersja wydana przez Główny Urząd Miar – GUM 1999), w którym zawarto założenia dotyczące wyrażania niepewności pomiaru. Główny nacisk w Przewodniku położono na warunki pomiaru, które bardzo rzadko bywają w pełni kontrolowane. Błąd jest rozpatrywany przy założeniu idealnych warunków pomiaru oraz tego, że obiekt jest „mierzalny” i znamy wartość referencyjną cechy obiektu, która jest mierzona. Wydaje się, że w różnych sytuacjach pomiar geodezyjny nie jest wykonywany w kontrolowanych warunkach laboratoryjnych, a obiekt bywa czasem bardzo trudny do zdefiniowania i pomiaru referencyjnego. Dobrym przykładem może być NMT z jego naturalnym mikro- i makroreliefem, zmieniającym się pod wpływem pór roku czy zabiegów agrotechnicznych.

W obu wymienionych materiałach źródłowych (Fisher 1999, 2005, GUM 1999) pojawia się pojęcie niepewności, przy czym jest ono nieco inaczej rozumiane. W związku z tym podjęto próbę połączenia tych dwóch podejść w model koncepcyjny szacowania jakości danych przestrzennych, w którym przyjęto następujące założenia:

- Jeśli wszystkie warunki pomiaru i obiekt są możliwe do poznania i opisu, to stosujemy pojęcie „**błąd**”.
- Jeśli warunki pomiaru i obiekt są dobrze określone, ale nie wiemy, czy jest to sytuacja idealna, moglibyśmy używać pojęcia „**niepewność**”.
- Gdy obiekt i / lub warunki pomiaru nie są wystarczająco zdefiniowane, mamy do czynienia z **nieokreślonością**; w takiej sytuacji stosuje się zwykle analizę z wykorzystaniem zbiorów rozmytych. Jako parametr jakości danych można by przyjąć szerokość przedziału wyostrzenia.
- Ostatnim przypadkiem jest sytuacja, w której obiekt i / lub warunki pomiaru są niejednoznacznie zdefiniowane. Nie oznacza to, że takie dane są całkowicie bezużyteczne, a jedynie tyle, że nie znamy ich parametrów jakościowych i taka informacja powinna być zamieszczona jako metadane.

W monografii skoncentrowano się na dwóch pierwszych, „mierzalnych” przypadkach, czyli na sytuacji idealnej, kiedy jakość danych możemy określić pojęciem „**błąd**”, oraz w sytuacji, gdy pomiar wykonywany jest w warunkach naturalnych i mamy do czynienia z **niepewnością**. W obu przypadkach jako parametr charakteryzujący jakość danych przyjmuje się zgodnie z Przewodnikiem (GUM 1999) odchylenie standardowe błędu / niepewności. Zakłada się przy tym, że rozkład błędów / niepewności ma charakter rozkładu normalnego. W monografii poddano pod dyskusję zasadność takiego założenia, przeanalizowano publikowaną w literaturze metodę odporną i jednocześnie zaproponowano parametry rozkładu Laplace’a na potrzeby określania niepewności danych charakteryzujących się wąskimi histogramami odchyłek od wartości referencyjnych, z „długimi ogonami”. Przeprowadzono dyskusję na temat sposobów określania jakości danych na podstawie dwóch typów danych przestrzennych: wektorowych obiektów powierzchniowych i NMT w postaci regularnej siatki.

## Abstract

Research presented in the monograph was inspired by project concerned processing of airborne laser scanning data (Hejmanowska et al. 2008a). Quality of Digital Terrain Model (DTM) was analyzed as project final product. Differences between DTM and heights of the points from the points' cloud were analyzed, histograms of the differences were observed. Shape of the histograms significantly differs from normal distribution. Histograms were concentrated around zero and so called „long tiles” appeared. Laplace distribution function was a first association. This phenomenon of specific histograms was confirmed in after performed literature studies. In the paper (Höhle, Höhle 2009) robust method was proposed as an alternative to the Gauss approach. Analysis of the empirical histograms, Gauss distribution and curve based on proposed, robust method one can come to the conclusion that normal distribution actually overestimate the DTM error however robust method underestimate it. The conclusion confirms sense of looking for other distribution function. Author of the monograph proposed Laplace distribution function, at the beginning only basing on the visual analysis.

Researches concerning DTM quality were performed until now mainly focusing the influence of the DTM error on results of spatial analysis (Hejmanowska 2005, Hejmanowska et al. 2008a, b). The error propagation is possible to be considered in spatial analysis. However firstly reliable method for DTM quality assessment is needed. Research in this domain was summarized in the paper (Hejmanowska et al. 2008b), where different approaches to DTM quality, reliability, accuracy, inaccuracy, uncertainty and error was defined and discussed. Interesting approach was proposed by Peter Fisher (Fisher 1999). He started from uncertainty, as general. In the case of well defined DTM uncertainty can be described by error. If DTM is not well defined uncertainty can be expressed by vagueness or ambiguity. Parallel in surveying environment some suggestions appeared concerning not use expression: error but just uncertainty. In 1993 assumptions and recommendations for the expression of uncertainty in measurements were published in English, as important for metrology guideline *Evaluation of measurement data – Guide to the expression of uncertainty in measurement*. Polish version was published in 1999 by Central Office of Measures. Main pressure is placed in the Guideline (GUM 1999) on measurements conditions, rarely fully controlled. Expression: error is reserved for ideal measurement situation, assuming the object is measurable and reference value is available. It seems the surveying is mainly not performed in control laboratory conditions and the object is often difficult to defining and reference measurement. DTM is a good

example with his natural micro- and macro relief varying in a year and according agriculture treatment. In two mentioned source papers (Fisher 1999, 2005, GUM 1999) expression: uncertainty appears but understood in different ways.

Therefore synthetic approach was proposed of the in the monograph, as a conceptual model for expression of spatial data quality where:

- If the measurement conditions and the object are possible to ideal description we can use expression: error.
- If the measurement conditions and the object are well defined, but we are not sure the situation is possible to ideal description we should use expression: uncertainty.
- In the case the object and / or measurement conditions are not well defined, we can talk about vagueness, when usually fuzzy sets are applied. Width of defuzzification range was proposed as quality indicator.
- In last case (ambiguity) object and / or measurement conditions are not unequivocal. It does not mean the data are useless, but only any parameters describing their quality are available and this information should be placed in metadata.

Monograph is focused on the two first, „measurable” cases, ideal situation when data quality can be determined by error and the measurements of natural objects in natural conditions when we talk about uncertainty. In both cases as error / uncertainty standard deviation is taken as quality indicator, according Guideline (GUM 1999). Assumption of normal error / uncertainty distribution is tacit made. Legitimacy of the assumption was in monograph discussed. Known from the literature alternative robust method and Laplace distribution were tested as well. Two groups of the examples were presented. One concerns uncertainty analysis of vector polygon objects and second DTM uncertainty in grid form.