

Włodzimierz Kęska
Politechnika Poznańska
Instytut Maszyn Roboczych i Pojazdów Samochodowych

METODA STATYSTYCZNEJ OBRÓBK DANYCH POMIAROWYCH W MODELOWANIU I WIZUALIZACJI MAP PŁONU

Streszczenie

Mapy rozmieszczenia plonu ziarna na powierzchni pola sporządza się na podstawie zarejestrowanych w trakcie zbioru danych o pozycji geograficznej kombajnu i natężeniu strumienia ziarna dopływającego do jego zbiornika. Dane te są obarczone licznymi losowymi błędami i wymagają intensywnej obróbki statystycznej, polegającej na ich transpozycji, filtrowaniu, uśrednianiu i interpolacji. W pracy zaproponowano procedurę interpolacyjno-uśredniającą z wykorzystaniem wykładniczej funkcji wagowej. Procedurę tę zaimplementowano w wykonanym w ramach niniejszej pracy programie komputerowym do opracowywania oraz graficznej prezentacji cyfrowych map plonu. Porównano zaproponowaną metodę z powszechnie stosowaną metodą odwrotnych odległości. Na podstawie eksperymentów cyfrowych stwierdzono, że metoda wykorzystująca wykładniczą funkcję wagową pozwala na uzyskanie lepszego przybliżenia rzeczywistego rozkładu plonu niż metoda odwrotnych odległości.

Słowa kluczowe: rolnictwo precyzyjne, mapy plonu, geostatystyka

Wykaz oznaczeń

- $z(\mathbf{x}_0), z(\mathbf{x}_i)$ – wysokość plonu w wybranym punkcie [kg/m²],
 \mathbf{x}_i – wektor współrzędnych punktu pomiarowego [m]
 w_i – współczynnik wagowy [-]
 b – czynnik skalujący [-]
 a – współczynnik skali zasięgu uśredniania [m]
 m – współczynnik kształtu funkcji wagowej [-]
 h_i – odległość i-tego punktu pomiarowego od punktu obliczeniowego \mathbf{x}_0 . [m]
 $p_{i1}, p_{i2}, i=1..n$ – ciąg parametrów położenia funkcji plonu [m]
 $a_i i=1..n$ – ciąg parametrów skali funkcji plonu

Wprowadzenie

Mapy plonu są zwykle sporządzane na podstawie pomiarów natężenia strumienia ziarna dopływającego do zbiornika kombajnu z jednoczesną rejestracją pozycji kombajnu w czasie tych pomiarów. W efekcie dysponuje się dyskretnym zbiorem danych, określających wysokość plonu w wybranych punktach pola. Pojedynczy rekord z pliku uzyskiwanego na polu podczas współbieżnej rejestracji pozycji kombajnu za pomocą GPS oraz sygnału określającego lokalny plon ziarna przedstawiono poniżej. Linia rozpoczynająca się znakiem \$ zawiera standardowy meldunek odbiornika GPS, w którym m. in. jest podany dokładny czas, współrzędne geograficzne w systemie WGS84, liczba widocznych satelitów, a także oszacowanie DOP.

4303, 0, 0

\$GPGGA,113113,5208.8085,N,01714.3624,E,2,08,1.0,54.6,M,40.0,M,,*71

Dane te pochodzą z pomiarów przeprowadzonych przez autora w roku 2003 przy zbiorze kukurydzy w Gowarzewie k/Poznania. Graficzna reprezentacja tego zbioru przedstawiona jest na rysunku 1. Sporządzenie mapy plonu w postaci ciągłej funkcji wymaga inteligentnego filtrowania tych danych oraz ich interpolacji [Beck i in. 2001, Noack i in. 2003]. Spośród wielu możliwych metod interpolacyjnych, w rolnictwie precyzyjnym szerokie rozpowszechnienie uzyskały metody polegające na obliczaniu nieznannej wartości funkcji w punkcie \mathbf{x}_0 jako średniej ważonej z wartości w kilku najbliższych punktach pomiarowych [Lodovicy i in. 2001].

$$z(x_0) = \sum_{i=1}^n w_i z(x_i) \quad (1)$$

Zakłada się że współczynniki wagowe są tak wyskalowane, by ich suma była równa jedności:

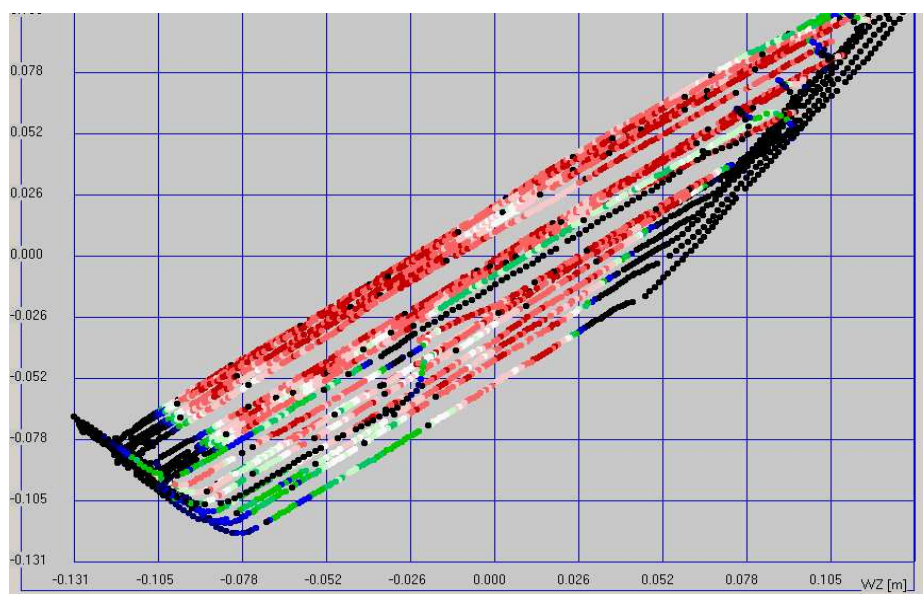
$$\sum_{i=1}^n w_i = 1 \quad (2)$$

Najprostszą, stosowaną powszechnie metodą wyznaczania współczynników wagowych, jest metoda odwrotnych odległości. W metodzie tej współczynniki wagowe są odwrotnie proporcjonalne do odległości h_i pomiędzy punktami \mathbf{x}_0 oraz \mathbf{x}_i . Procedurą posiadającą ściśle uzasadnienie statystyczne jest tzw. kriging (interpolacja Kriego'a). W interpolacji Kriego'a współczynniki wagowe wyznacza się w oparciu o tzw. wariogram, tj funkcję określającą stopień zdeterminowania wartości funkcji plonu w danym punkcie przez wartości w punktach otaczających. Funkcja ta rośnie tym szybciej, im bardziej niezależne od siebie są dane z sąsiadujących punktów pomiarowych. Szczegółowy opis tej metody można znaleźć m.in. w pracy [Wackernagel'a 1998]. Obydwie metody dają funkcje interpolujące przechodzące dokładnie przez punkty pomiarowe. Jeśli jednak, jak to ma miejsce

podczas sporządzania map plonu na kombajnie zbożowym, dysponuje się stosunkowo dużą liczbą danych pomiarowych, to przy ograniczonej dokładności pojedynczego pomiaru, dokładna interpolacja prowadzi do uzyskania funkcji o bardzo złożonym kształcie, która niepotrzebnie odwzorowuje wszystkie losowe błędy w punktach pomiarowych. Proponuje się w tym przypadku zastosowanie interpolacji przybliżonej, ze współczynnikami wagowymi wyznaczanymi ze wzoru:

$$w_i = b \exp \left[- \left(\frac{h_i}{a_i} \right)^m \right] \quad (3)$$

Funkcja ta dąży do zera przy zwiększaniu h , jednak przyjmuje ona skończoną wartość dla $h=0$, co oznacza, że nawet dla $h=0$, tj. w węźle interpolacji, sąsiednie punkty pomiarowe będą wpływały na wyznaczoną wartość funkcji interpolującej. Powstała hipoteza, że w szczególnej, wyżej opisanej sytuacji, jaką spotyka się przy sporządzaniu map plonu ziarna, zastosowanie tej funkcji wagowej może dawać lepsze przybliżenie funkcji plonu $z(\mathbf{x})$ - w porównaniu z pozostałymi, wyżej omówionymi metodami.



Rys. 1. Przykład rozkładu punktów pomiarowych oraz wartości surowego sygnału wysokości lokalnego plonu zarejestrowane podczas kombajnowego zbioru kukurydzy. Źródło: badania własne

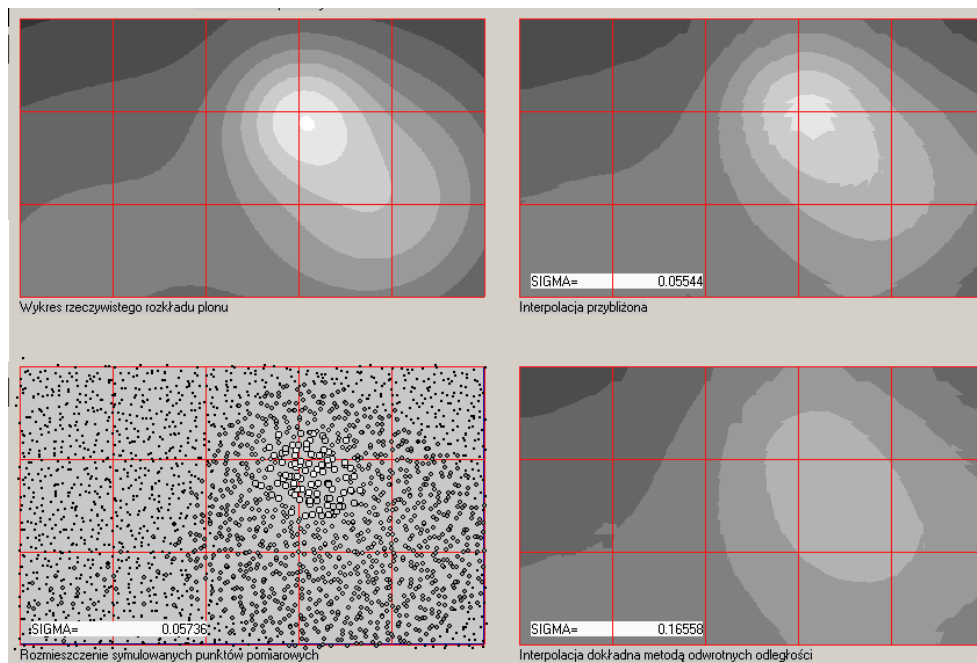
Fig. 1. An example of on field distribution of measuring points and raw yield data logged while corn harvesting

Materiał i metody

Dla zweryfikowania powyższej hipotezy zastosowano symulacje komputerową. Badanie symulacyjne polegało na tym, że ze znanej, danej analitycznie funkcji $z(\mathbf{x})$, generowano zbiór symulowanych danych pomiarowych, rozmieszczonych wzdłuż trasy kombajnu na polu o powierzchni 15ha. Do dokładnie obliczonych na podstawie funkcji $z(\mathbf{x})$ wartości plonu, dodawano losowo generowane błędy pomiaru plonu ziarna, a do wektora położenia \mathbf{x} błędy pomiaru pozycji maszyny. Tak uzyskany zbiór sztucznych danych pomiarowych poddawano obróbce statystycznej za pomocą przedstawionych wyżej procedur, uzyskując aproksymowany obraz rozmieszczenia plonu $f(\mathbf{x})$, który łatwo można było porównać z wyjściową funkcją analityczną $z(\mathbf{x})$. Miarą dokładności odwzorowania była wartość odchylenia standardowego funkcji aproksymującej $f(\mathbf{x})$ od funkcji $z(\mathbf{x})$. Stosowne oprogramowanie do manipulacji zbiorami danych, ich statystycznej obróbki i wizualizacji wyników obliczeń w formie map bitowych na ekranie monitora i drukarce wykonano we własnym zakresie, w środowisku Delphi 3.

Wyniki badań i ich analiza

W wyniku każdej przeprowadzonej symulacji uzyskiwano wartości liczbowe, określające dokładność przybliżenia funkcji $z(\mathbf{x})$, uzyskanego różnymi metodami. Dla przykładu, na rysunku 2 przedstawiono graficznie wyniki takiej symulacji. Pierwszy wykres odzwierciedla zadaną funkcję plonu $z(\mathbf{x})$. Jako funkcję testową wybrano funkcję o postaci sumy szeregu dwuwymiarowych funkcji Gaussa. Na wykresie drugim zilustrowano rozmieszczenie punktów pomiarowych i wartości plonu odpowiadające tym punktom, uzyskane podczas symulowanego zbioru kombajnowego. Błąd pomiaru pozycji maszyny generowano na podstawie zarejestrowanej w pomiarach doświadczalnych trajektorii błędu GPS Garmin 16 HSV. Błąd pomiaru plonu generowano sztucznie zakładając jego rozkład normalny o odchyleniu standardowym 10% w stosunku do średniej wysokości plonu. Na wykresie prawym górnym przedstawiono funkcję rozmieszczenia plonu za pomocą metody przybliżonej z wykładniczą funkcją wagową, zaś na prawym dolnym przedstawiono tę funkcję wyznaczoną za pomocą metody odwrotnych odległości. Względny błąd przybliżenia dla metody pierwszej wynosi 0,055 zaś dla metody drugiej już 0,165. Jak widać, przybliżenie uzyskane zaproponowaną metodą jest znacznie lepsze od tego uzyskanego drugą metodą, tj. metodą interpolacji dokładnej. Dalsze powtórzenia tego eksperymentu numerycznego dla różnych poziomów zakłóceń i różnych rodzajów funkcji plonu dały podobny wynik.



Rys. 2. Przykład symulacji kombajnowego zbioru i porównania dwóch metod statystycznego opracowania wyników pomiarów lokalnego plonu ziarna
 Fig. 2. An example of grain harvest simulation and comparing of the two methods for statistical processing of the raw yield data

Podsumowanie

1. W wyniku porównania różnych procedur statystycznych stwierdzono, że w warunkach gęstego rozmieszczenia punktów pomiarowych i znacznego losowego zakłócenia wyników pomiarów pozycji maszyny i natężenia strumienia ziarna, jakie może mieć miejsce podczas wykonywania pomiarów na kombajnie zbożowym w czasie zbioru, metoda średnich ważonych z zastosowaniem wykładniczej funkcji wagowej może dawać dokładniejsze przybliżenie rzeczywistego rozkładu plonu niż metoda odwrotnych odległości, przy podobnym nakładzie obliczeniowym. Wniosek ten, na podstawie przedstawionego powyżej rozumowania, można intuicyjnie rozszerzyć na interpolację Kriega, co jednak wymaga jeszcze dodatkowej weryfikacji numerycznej.
2. Wykorzystywany w ramach niniejszej pracy do weryfikacji zaproponowanej metody interpolacji program komputerowy, po dopracowaniu i dostosowaniu do specyficznych potrzeb użytkownika, może być wykorzystany praktycznie

w rolnictwie precyzyjnym do sporządzania numerycznych map plonu, wzbogacając ofertę znanego już oprogramowania, przy czym dostępność wersji źródłowej programu daje możliwość jego łatwej adaptacji i doskonalenia - co ma szczególne znaczenie w pracach badawczych.

Bibliografia

Beck A. D., Searcy S. W., Roades J. P. 2001. Yield data filtering techniques for improved accuracy. *Applied engineering in Agriculture*. 17. 423-431; 1.

Lodovicy Ch., Schwaiberger R., Leithold P. 2001. *Precision Farming*. DLG Verlag.

Noack P.O., Muhr T., Demmel M. 2003. An algorithm for automatic detection of defective yield data. *Precision Agriculture*, Waageningen A. P. 445-449.

Wackernagel H. 1998. *Multivariate Geostatistics*. Springer Verlag.

METHOD FOR STATISTICAL PROCESSING OF RAW YIELD DATA IN MODELING AND VISUALIZATION OF YIELD MAPS

Summary

Yield maps are constructed on the base of raw data (geographical position and grain flow) logged on the combine in the course of harvest. Such data are greatly distorted by errors originating from many different sources. Some of those errors can be eliminated by transposition, filtration and interpolation of the raw data. In this work an averaging - interpolation procedure, based on the weighted averages using the exponential weighting function is described. The special computer program for constructing numerical yield maps and testing the selected statistical procedures for yield data processing was written and tested. The numerical experiments has pointed out, that the newly proposed method has better ability to eliminate the random errors of yield function then inverse distance procedure broadly used in practice.

Key words: precision agriculture, yield map, geostatistics