

## **POZYSKIWANIE I WYKORZYSTANIE DANYCH POGODOWYCH W KRAJOWYM SYSTEMIE WSPOMAGANIA DECYZJI W PRODUKCJI ROŚLINNEJ**

Andrzej S. Zaliwski

*Zakład Agrometeorologii i Zastosowań Informatyki,*

*Instytut Uprawy Nawożenia i Gleboznawstwa-Państwowy Instytut Badawczy w Pulawach*

**Streszczenie.** Uzyskanie nowej informacji w Krajowym Systemie Wspomagania Decyzji w Produkcji Roślinnej wiąże się z automatyzacją pozyskiwania danych pogodowych. W systemie dane pogodowe są wykorzystywane do prezentacji aktualnego stanu pogody oraz do generowania nowych informacji w modelach. Przedstawiono zasadę agregacji danych i analizę modelu „Prognoza negatywna” pod względem przetwarzania danych. Model generuje nowe informacje przez porównanie informacji zawartych w uporządkowanych łańcuchach danych pogodowych z informacjami w łańcuchu danych wzorcowych.

**Słowa kluczowe:** dane pogodowe, jakościowa teoria informacji, produkcja roślinna, system wspomagania decyzji

### **Wprowadzenie**

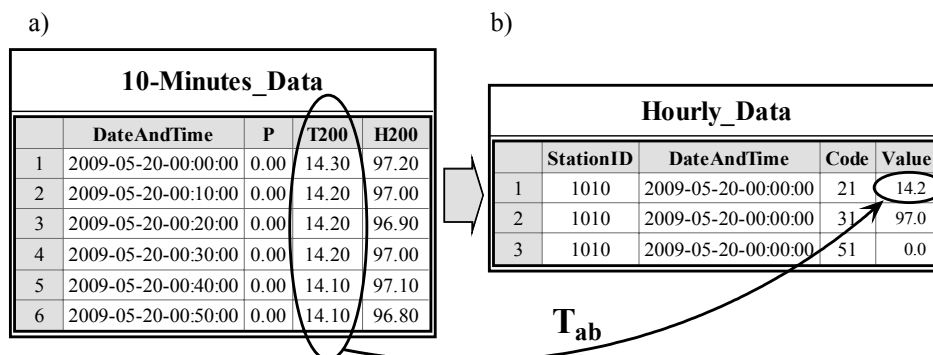
Celem niniejszej pracy jest analiza procesów informacyjnych zachodzących przy pozyskiwaniu i przetwarzaniu danych w „Krajowym systemie wspomagania decyzji w zakresie produkcji roślinnej” (w skrócie SWDPR), opracowywanym w IUNG-PIB.

Z punktu widzenia przetwarzania informacji najbardziej interesującym przedmiotem analizy są niewątpliwie wszelkie modele, ze względu na możliwość generowania nowych informacji. W systemie SWDPR wykorzystywane są m.in. meteorologiczne modele ochrony roślin. Są one najkorzystniejsze z punktu widzenia użytkownika, bowiem nie wymagają ręcznego wprowadzania wielu danych. Warunkiem ich praktycznej przydatności jest sieć automatycznych stacji agrometeorologicznych wyposażonych w czujniki do pozyskiwania wszystkich potrzebnych danych.

### **Pozyskiwanie danych pogodowych**

Jednym z meteorologicznych modeli ochrony roślin w systemie SWDPR jest model ochrony ziemniaka przed zarazą ziemniaka o nazwie „Prognoza negatywna” (PN) [California PestCast 2005]. Model ten jest od kilku lat wykorzystywany w skali produkcyjnej w Polsce z dobrymi rezultatami [Rysak 2009]. Dane zasilające model PN to temperatura i wilgotność względna powietrza oraz suma opadów z pomiarów godzinowych. Ponieważ

pomiary na stacjach agrometeorologicznych IUNG-PIB wykonywane są co 10 minut, pozyskanie danych dla modelu PN obejmuje agregację danych pomiarowych. Konieczna jest także zmiana ich formatu, co wynika z różnych struktur tabel, w których gromadzone są dane ze stacji i dane zasilające model PN (rys. 1).



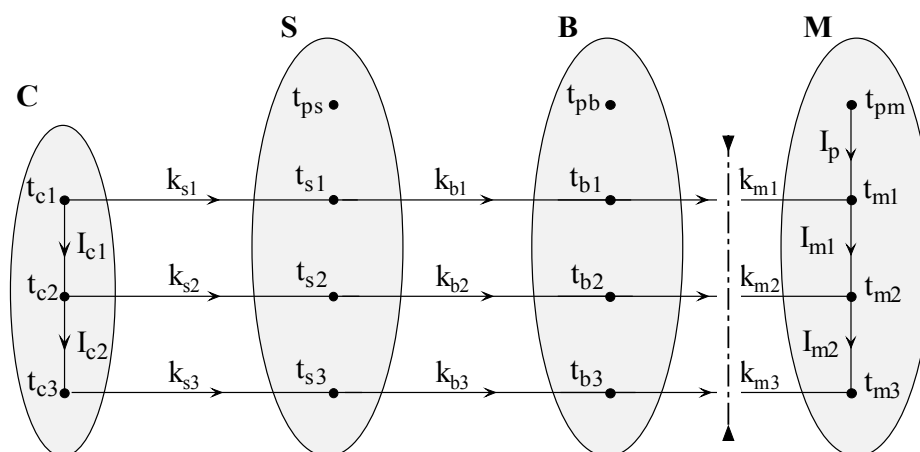
Rys. 1. Agregacja danych ze stacji agrometeorologicznej dla modelu „Prognoza negatywna”. Oznaczenia: a) dane 10-minutowe, b) dane godzinowe; P, T200, H200, Value - kolumny z wartościami elementów pogodowych, DateAndTime - data i czas pomiaru, StationID - indeksy stacji, Code - kody elementów pogodowych,  $T_{ab}$  - transformacja danych przez agregację

Fig. 1 Aggregation of data from an agrometeorological station for the “Negative prognosis” model. Significance: a) 10-minutes data, b) hourly data; P, T200, H200, Value - columns with weather element values, DateAndTime - date and time of measurements, StationID - station indices, Code - weather element codes,  $T_{ab}$  - data transformation by aggregation.

Między tabelami 10-Minutes\_Data i Hourly\_Data (rys. 1) występują następujące różnice strukturalne. W tabeli 10-Minutes\_Data wartości elementów pogodowych są przechowywane w osobnych kolumnach *P*, *T200* i *H200*. Identyfikacja danych w przestrzeni geograficznej jest możliwa dzięki przypisaniu tabeli do konkretnej stacji. W tabeli Hourly\_Data dane przechowywane są w jednej kolumnie *Value*, identyfikację danych w przestrzeni geograficznej zapewnia indeks stacji *StationID*, kod elementu pogodowego *Code* wskazuje, co oznacza wartość w kolumnie *Value*. Sposób identyfikacji danych na osi czasu jest podobny w obu przypadkach (jedyna różnica to inna częstotliwość pomiarów).

W analizie procesów informacyjnych konieczne jest rozróżnienie pojęć „dane” i „informacja”. Dostępna literatura podaje różne określenia tych pojęć [np. Forlicz 2008; Jaros i Pabis 2007]. W niniejszej pracy analiza procesów informacyjnych jest przeprowadzona zgodnie z Jakościową teorią informacji [Mazur 1970], dlatego przez wzgląd na spójność wykładu przyjęto definicje analogiczne do podanych przez M. Mazura, zastępując termin „komunikat” terminem „dana” oraz „tor sterujący” terminem „tor informacyjny”. Daną można więc rozumieć jako „stan fizyczny różniący się w określony sposób od innego stanu fizycznego”, informacja natomiast jest to „transformacja jednej danej w drugą daną w torze informacyjnym”. Mazur rozróżnia dwa rodzaje transformacji, operacyjne (np. operacje arytmetyczne) i asocjacyjne (np. relacja jeden-do-jeden w bazie danych).

Na rys. 2 przedstawiono schemat przepływu danych między czujnikiem temperatury stacji agrometeorologicznej a systemem SWDPR w jednym torze informacyjnym (w rzeczywistości liczba torów informacyjnych odpowiada liczbie czujników). Schemat ten wyjaśnia pochodzenie danych w kolumnie *T200* w tabeli *10-Minutes\_Data* (rys.1). Stany fizyczne czujnika (zmiany oporu elektrycznego pod wpływem temperatury otoczenia) są transformowane przy pomocy kodów  $k_{s1}, \dots, k_{s3}$  na wartości liczbowe  $t_{s1}, \dots, t_{s3}$ . Kody  $k_{s1}, \dots, k_{s3}$  (podobnie jak  $k_{b1}, \dots, k_{b3}$ ) są transformacjami danych wzdłuż toru informacyjnego (tzn. nie w ramach jednego zbioru, ale pomiędzy zbiorami). Kody muszą być tak dobrane, aby informacje  $I_{m1}$  i  $I_{m2}$  uzyskane z danych w zbiorze M nadawały się do sterowania modelem zamiast informacji  $I_{c1}$  i  $I_{c2}$ , które w chwili pracy modelu już nie istnieją. Zbiory S, B i M (rys.2) zawierają także dane  $t_{ps}$ ,  $t_{pb}$ ,  $t_{pm}$ , których brak w zbiorze oryginałów - są to parametry stacji, bazy danych i modelu, dotyczące wielkości mierzonych. Rozpatrując np. wartości  $t_{s1}, \dots, t_{s3}$  w połączeniu z danymi  $t_{ps}$  uzyskujemy informację, że chodzi o temperatury chwilowe powietrza na wysokości 2m mierzone w °C. Parametry, jako właściwości stałe wymienionych obiektów, nie muszą być każdorazowo przekazywane wzdłuż toru informacyjnego.

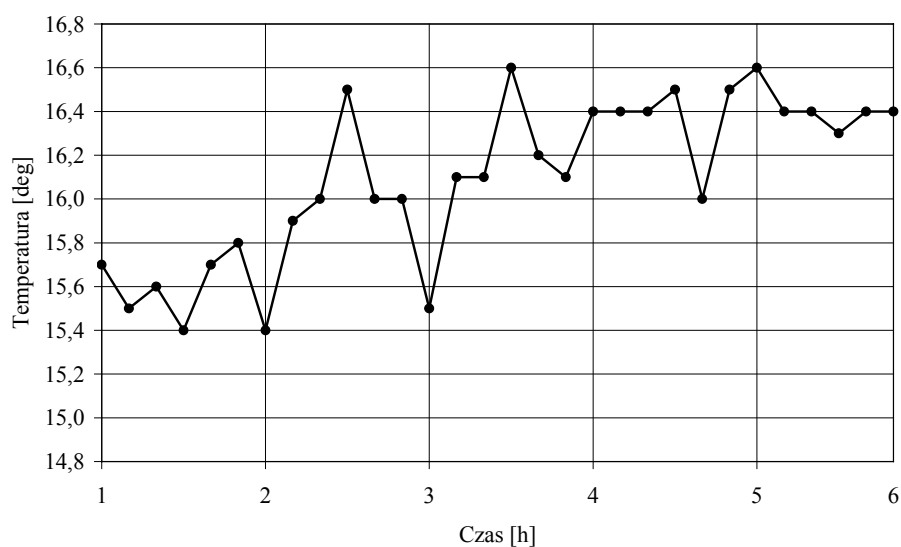


Rys. 2. Przepływ danych w torze informacyjnym między czujnikiem stacji a modelem bez utraty informacji. Oznaczenia: C - zbiór oryginałów (temperatur czujnika), S - zbiór danych na wyjściu stacji, B - zbiór danych w bazie danych, M - zbiór danych na wejściu modelu,  $t_{ps}$ ,  $t_{pb}$ ,  $t_{pm}$  - parametry stacji, bazy danych i modelu,  $\{t_{c1}, \dots, t_{c3}\}$ ,  $\{t_{s1}, \dots, t_{s3}\}$ ,  $\{t_{b1}, \dots, t_{b3}\}$ ,  $\{t_{m1}, \dots, t_{m3}\}$  - temperatury,  $I_p$ ,  $I_{c1}$ ,  $I_{c2}$ ,  $I_{m1}$ ,  $I_{m2}$  - informacje,  $\{k_{s1}, \dots, k_{m3}\}$  - kody

Fig. 2. Information-lossless data flow in sensor-model communication channel. Significance: C - originals set (sensor temperatures), S - data set at station output, B - data set in database, M - data set at model input,  $t_{ps}$ ,  $t_{pb}$ ,  $t_{pm}$  - station, database and model parameters,  $\{t_{c1}, \dots, t_{c3}\}$ ,  $\{t_{s1}, \dots, t_{s3}\}$ ,  $\{t_{b1}, \dots, t_{b3}\}$ ,  $\{t_{m1}, \dots, t_{m3}\}$  - temperatures,  $I_p$ ,  $I_{c1}$ ,  $I_{c2}$ ,  $I_{m1}$ ,  $I_{m2}$  - information,  $\{k_{s1}, \dots, k_{m3}\}$  - codes

Na rys. 2 nie uwzględniono daty i czasu pomiaru, generowanych przez zegar stacji, ani danych z innych czujników, a przedstawionych w tabelach na rys. 1a i 1b. W rzeczywistości wszystkie rodzaje danych płyną początkowo własnymi torami informacyjnymi i są dołączane do zbioru S. Następnie jednym cyfrowym torem informacyjnym przesyłane są do zbioru B, gdzie tworzone są rekordy danych pogodowych w tabeli 10-Minutes\_Data. Data i czas pomiaru umożliwiają wybieranie i sortowanie zbioru rekordów przy tworzeniu łańcuchów danych zasilających model. Tor informacyjny pomiędzy bazą danych a modelem jest przerwany w celu zaznaczenia, że stacja nie zasila modelu bezpośrednio. Model NP bowiem nie pracuje w czasie rzeczywistym, ale wykorzystuje dane historyczne (obrazy danych oryginalnych w zbiorze B) od wschodów ziemiaka do ostatnich danych zapisanych w bazie danych.

Z rys. 1 wynika, że zbiór danych 10-minutowych jest transformowany na zbiór danych godzinowych przy pomocy transformacji  $T_{ab}$  redukującej liczebność danych. Redukcję taką można przeprowadzić wg różnych zasad, np. przez wybór jednej danej 10-minutowej z pomiarów z tej samej godziny z tabeli 10-Minutes\_Data.

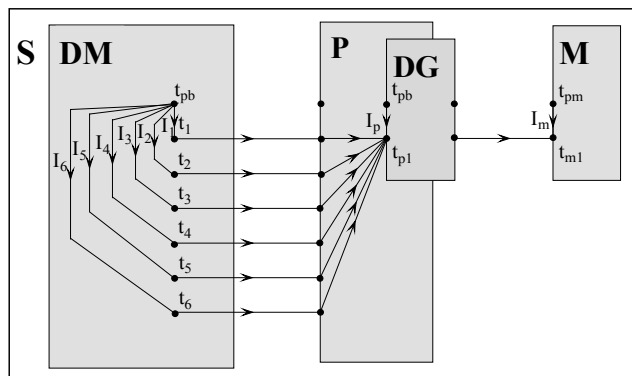


Rys. 3. Przebieg temperatur 10-minutowych w czasie pięciu kolejnych godzin (wybrano odcinek czasowy o wyraźnych wahanach temperatury)

Fig. 3 The course of 10-minutes temperatures in five consecutive hours (period of pronounced temperature variability was selected)

Charakter zmian temperatury w ciągu godziny (rys. 3) wskazuje jednak na konieczność podejścia zapewniającego uwzględnienie wszystkich danych 10-minutowych (DM) wchodzących w zakres czasowy danej godzinowej (DG). W systemie SWDPR zastosowano agregację informacji przez uśrednienie. Zasadę przetwarzania informacji w procesie agregowania wyjaśnia rys. 4. Informacje 10-minutowe  $I_1, \dots, I_6$  zawarte w zbiorze temperatur mierzonych  $\{t_1, \dots, t_6\}$  w ciągu jednej godziny przez stację (zapisane w tabeli DM) zostają

zastąpione wartością średnią  $I_p$  obliczoną przez program P. Mimo, że agregacja powoduje utratę części informacji (oprócz przypadków, gdy wszystkie dane w zbiorze  $\{t_1, \dots, t_6\}$  są takie same), to jednak jedna dana wybrana arbitralnie z sześciu jest rozwiązaniem gorszym, zwłaszcza przy dużych chwilowych wahaniami temperatury (rys. 3).



Rys. 4. Agregacja danych. Oznaczenia: S - system SWDPR, DM - tabela danych 10-minutowych,  $t_{pb}$ ,  $t_{pm}$  - parametry (jednostka temperatury  $1^\circ\text{C}$ ),  $t_1, \dots, t_6$  - temperatury 10-minutowe,  $I_1, \dots, I_6$  - informacje 10-minutowe, P - program przetwarzający dane, DG - tabela danych godzinowych,  $t_{p1}$ ,  $t_{m1}$  - temperatury godzinowe (wynik agregacji),  $I_p$ ,  $I_m$  - informacje utworzone z danych agregowanych, M - model

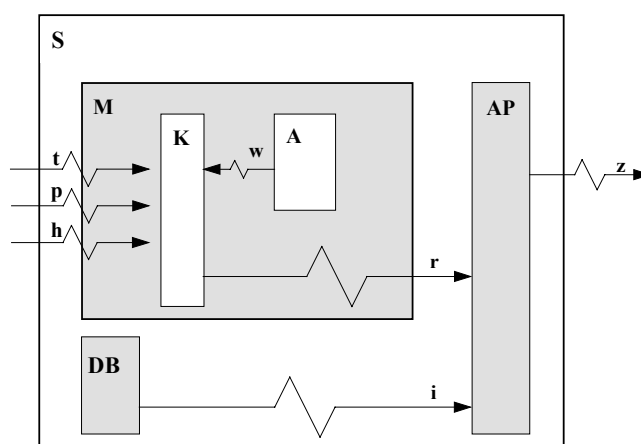
Fig. 4. Data aggregation. Significance: S - PPDSS system, DM - table with 10-minutes data,  $t_{pb}$ ,  $t_{pm}$  - parameters ( $1^\circ\text{C}$  temperature unit),  $t_1, \dots, t_6$  - 10-minutes temperatures,  $I_1, \dots, I_6$  - 10-minutes information, P - data processing program, DG - table with hourly data,  $t_{p1}$ ,  $t_{m1}$  - hourly temperatures (result of aggregation),  $I_p$ ,  $I_m$  - information created from data aggregated, M - model

### Wykorzystanie danych pogodowych

Dane pogodowe w systemie SWDPR ([www.ipm.iung.pulawy.pl](http://www.ipm.iung.pulawy.pl)) są wykorzystywane do prezentacji aktualnego stanu pogody (strona internetowa: Informacje o pogodzie - Stacje agrometeorologiczne) oraz są przetwarzane w modelach w celu uzyskania aktualnych informacji, w formie zaleceń, potrzebnych do podejmowania decyzji.

Uzyskanie nowych informacji w systemie SWDPR wiąże się przede wszystkim z dopływem świeżych danych, bowiem algorytmy transformowania danych w informacje zmieniają się rzadziej. Zasadę uzyskiwania informacji przedstawiono na rys. 5 na przykładzie modelu „Prognoza negatywna” (M) systemu SWDPR (S). Dane „świeże” są przesyłane do modelu w postaci łańcuchów danych  $t, p, h$ , złożonych z godzinowych wartości elementów pogodowych. Początek każdego łańcucha musi odpowiadać dacie wschodów ziemniaka, koniec natomiast ostatniej danej w tabeli danych godzinowych Hourly\_Data (przedstawionej na rys.1b). Pojedynczy łańcuch danych tworzony jest przez wybór danych odnoszących się do określonej stacji agrometeorologicznej (identyfikacja danych w przestrzeni geograficznej) oraz ich uporządkowanie wg daty w porządku rosnącym. Model

transformuje dane zawarte w łańcuchach i dane w łańcuchu wzorców ( $w$  na rys. 5) na informacje dotyczące dobowych wartości ryzyka wystąpienia zarazy ziemniaka. Wartości te są dodawane do siebie i informacja ta jest porównywana ze wzorcem. Wynik porównania (łańcuch  $r$  na rys. 5) jest przesyłany do aplikacji internetowej, której zadaniem jest nadanie interpretacji słownej (utworzenie tekstu zalecenia w postaci łańcucha danych  $z$ ) w zależności od informacji zawartej w łańcuchu  $r$ . Zalecenie to jest wyświetlane przez przeglądarkę użytkownika systemu SWDPR i może być wykorzystane w praktyce.



Rys. 5. Generowanie nowych informacji w systemie SWDPR. Oznaczenia: S - system SWDPR, M - model, K - komparator, A - zbiór algorytmów i parametrów modelu, DB - baza danych systemu, AP - aplikacja internetowa,  $t$ ,  $p$ ,  $h$ ,  $w$ ,  $r$ ,  $i$ ,  $z$  - łańcuchy informacyjne odpowiednio: temperatury, opadów, wilgotności,  $w$  - wzorców modelu,  $r$  - ryzyka wystąpienia choroby,  $i$  - interpretacji ryzyka,  $z$  - zaleceń

Fig. 5. New information generation in the SWDPR (the National Decision Support System in Plant Production) system. Significance: S - SWDPR system, M - model, K - comparator, A - model algorithms and parameters set, DB - system database, AP - Internet application,  $t$ ,  $p$ ,  $h$ ,  $w$ ,  $r$ ,  $i$ ,  $z$  - data chains: temperature, precipitation, humidity,  $w$  - model standards,  $r$  - disease occurrence risk,  $i$  - risk interpretation,  $z$  - recommendations

## Podsumowanie

Z punktu widzenia przetwarzania informacji do najbardziej interesujących elementów systemu SWDPR można zaliczyć te, które umożliwiają automatyczne generowanie nowych informacji. Uzyskanie nowych informacji wiąże się przede wszystkim z automatyzacją pozyskiwania danych pogodowych, bowiem wprowadzanie danych przez użytkownika jest ograniczone, a algorytmy przetwarzania danych zmieniają się rzadko. Elementami pozyskiwania danych są zmiana skali czasowej danych (agregacja) i ich gromadzenie. Zmniejszenie skali powoduje utratę części informacji, dlatego ważny jest wybór algorytmów zapewniających możliwie najlepszą jakość informacji zagregowanych.

W systemie SWDPR dane pogodowe są wykorzystywane do prezentacji aktualnego stanu pogody oraz do generowania nowych informacji w modelach. Model „Prognoza negatywna” analizowano pod względem przetwarzania danych. Pracuje on na zasadzie porównania informacji zawartych w łańcuchach danych wejściowych z informacjami wzorcowymi.

## Bibliografia

- Forlicz S.** 2008. Informacja w biznesie. Wyd. PWE. Warszawa. ISBN 978-83-208-1764-5.
- Jaros M., Pabis S.** 2007. Inżynieria systemów. Wyd. SGGW. Warszawa. ISBN 83-7244-832-9.
- Mazur M.** 1970. Jakościowa teoria informacji. WNT. Warszawa. [Dostęp 25-03-2010]. Dostępny w Internecie: [www.autonom.edu.pl/publikacje.html](http://www.autonom.edu.pl/publikacje.html).
- Rysak W.** 2009. Praktyczne zastosowanie systemu wspomagania decyzji w ochronie ziemniaka na terenie województwa lubelskiego. Studia i Raporty IUNG-PIB. z.16: 59-67. [Dostęp 19-07-2010]. Dostępny w Internecie: [www.iung.pulawy.pl/images/wyd/pib/zesz16.pdf](http://www.iung.pulawy.pl/images/wyd/pib/zesz16.pdf).
- California PestCast. 2005. Models: Late Blight of Potato. Statewide IPM Program, Agriculture and Natural Resources. University of California. [Dostęp 25-03-2010]. Dostępny w Internecie: [www.ipm.ucdavis.edu/DISEASE/DATABASE/potatolateblight.html](http://www.ipm.ucdavis.edu/DISEASE/DATABASE/potatolateblight.html).

*Publikacja opracowana w ramach zadania 2.9 programu wieloletniego „Kształtowanie środowiska rolniczego Polski oraz zrównoważony rozwój produkcji rolniczej”*

## ACQUISITION AND USE OF WEATHER DATA IN THE NATIONAL SYSTEM OF DECISION SUPPORT IN PLANT PRODUCTION

**Abstract.** The obtaining of new information in the National System of Decision Support in Plant Production is connected with the automation of obtaining of weather data. Weather data in the system are used for presentation of the current weather and for generation of new information in models. The data aggregation rule and the analysis of the “Negative forecast” model were presented in terms of data processing. The model generates new information by comparing information contained in ordered weather data chains with information in the standard data chain.

**Key words:** weather data, quality theory of information, plant production, decision support system

### Adres do korespondencji:

Andrzej Zaliwski; e-mail: [andrzej.zaliwski@iung.pulawy.pl](mailto:andrzej.zaliwski@iung.pulawy.pl)  
Zakład Agrometeorologii i Zastosowań Informatyki  
Instytut Uprawy Nawożenia i Gleboznawstwa - Państwowy Instytut Badawczy  
ul. Czartoryskich 8  
24-100 Puławy