



УДК 502:519.2

© 2010

**Г. М. Жолобак, В. Г. Якимчук, О. І. Сахацький,
К. Ю. Суханов, О. Д. Рябоконеко,
член-кореспондент НАН України О. Д. Федоровський**

Моделювання і прогнозування в задачах природокористування (на прикладі оцінки врожайності озимої пшениці)

Розглядається рівняння, що зв'язує врожайність озимої пшениці зі значеннями температури й опадів. Рівняння має форму множинної регресії, його коефіцієнти обчислюються за архівними даними ряду років врожайності і подекадних значень сумарних опадів та середньої температури. Прогнозування врожайності виконується шляхом обчислення значення функції регресії при підстановці в неї подекадних значень сумарних опадів і середньої температури. Наведено приклад прогнозування врожайності озимої пшениці для Київської області в 2002 р. і оцінка його точності залежно від часу прогнозу.

При розгляді природного середовища як складної геосистеми, структурні складові якої генетично об'єднані взаємозалежними геофізичними гідрометеорологічними та біологічними процесами, на основі статистичних та архівних даних попередніх спостережень отримують інформацію про взаємний вплив цих процесів у складних природних системах (ПС). Це дає можливість формалізувати причинно-наслідкові зв'язки в ПС, обґрунтувати найраціональніші методи побудови математичних моделей для моделювання та прогнозування сценаріїв розвитку процесів, що відбуваються в ПС, і адаптувати моделі ПС до поточних спостережень.

Такі дослідження найчастіше виконують методами кореляційного і регресійного аналізу. У даному повідомленні розглянуто використання багатовимірної регресійної моделі для моделювання в задачах природокористування на прикладі прогнозування врожайності озимої пшениці.

Відомі методи прогнозування врожайності [1, 2] вимагають численних контактних вимірів характеристик ґрунту та рослин. Нами розробляються також способи прогнозування врожайності зернових на основі вегетаційних індексів із залученням дистанційних матеріалів — космічних знімків різної просторової розрізненості [3, 4]. У рамках розроблюваної методики для отримання точного прогнозу доцільно дистанційні дані доповнювати наземни-

ми агрометеорологічними. Але тут постає питання, чи можливо розрахувати досить точно значення величини врожайності, використовуючи мінімальну кількість вхідних метеопараметрів — таких найбільш вживаних, як середня декадна температура повітря і сума опадів за декаду. Наша мета полягала в дослідженні можливості прогнозування середньорайонної врожайності озимої пшениці шляхом визначення множинної регресії, що описує взаємозв'язок двох ознак-факторів (температура й опади) та результативної ознаки (врожайність) для ряду районів Київської області.

Для достовірного відображення об'єктивно існуючих процесів вегетації зернових культур необхідно виявити істотні взаємозв'язки між цими процесами і метеорологічними умовами. Припускається, що між зазначеними змінними величинами існує залежність, яка є стохастичною та включає кореляційні і регресійні зв'язки. Крім того, при дослідженні таких процесів необхідно не тільки виявити зв'язок між змінними величинами, але й вивчити та встановити його форму, що і є основною задачею регресійного аналізу.

Необхідно знайти оптимальний варіант моделі, що відображає основні закономірності досліджуваного явища з достатнім ступенем статистичної надійності. З одного боку, у модель повинні бути включені всі фактори, що впливають на залежну змінну (у нашому випадку — врожайність). При невиконанні цієї вимоги модель може виявитися неадекватною внаслідок недообліку істотних факторів. З іншого боку, кількість факторів, що включаються в модель, не повинна бути занадто великою. Невиконання цієї вимоги призводить до неможливості використання досить складних залежностей, до зниження точності оцінок, до складності інтерпретації моделі і до труднощів її практичного застосування.

Для проведення аналізу були взяті дані про врожайність і метеоумови в Баришівському, Білоцерківському, Миронівському і Яготинському районах Київської області в 1992–2002 рр. Вибір цих районів обумовлений тим, що кожен з них репрезентує агрометеорологічні особливості основних чотирьох з наявних п'яти агрометеорологічних груп районів Київської області [5]. Задача дослідження — виявлення і вивчення залежності врожайності озимої пшениці від температури і кількості опадів у зазначених районах.

Відповідно до процедури регресійного аналізу визначимо врожайність озимої пшениці (u , центнери з гектара) по району як *результативну ознаку*, а середню температуру повітря (t) і суму опадів (d) за декаду — як *факторні ознаки*. Для спрощення сприйняття розрахунків і всього дослідження введемо в позначення індекс місяця і декади місяця:

t_{091} — послідовність вимірів середньої температури повітря ($^{\circ}$ C) за першу декаду вересня;

d_{091} — сума опадів (мм) за першу декаду вересня;

t_{092} — послідовність вимірів середньої температури повітря за другу декаду вересня;

d_{092} — сума опадів за другу декаду вересня;

...;

t_{072} — послідовність вимірів середньої температури повітря за другу декаду липня;

d_{072} — сума опадів за другу декаду липня;

t_{073} — послідовність вимірів середньої температури повітря за третю декаду липня;

d_{073} — сума опадів за третю декаду липня.

Для прогнозування врожайності в певний момент часу між посівом і збиранням врожаю необхідно мати рівняння, що виражає залежність врожайності від декадних значень температури й кількості опадів за період від посіву до часу прогнозу, за винятком зимових місяців. Наприклад, прогноз на початок травня повинен враховувати значення температури й кількість опадів після посіву для всіх декад до кінця квітня.

Для спрощення математичних викладень змінимо позначення факторних ознак $t_{091}, d_{091}, \dots, d_{073}$ на x_1, x_2, \dots, x_k . Тоді багатовимірною лінійною регресійною моделлю залежності u від факторних змінних x_1, x_2, \dots, x_k , що впливають на врожайність, матиме вигляд [6]

$$u = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k. \quad (1)$$

Для оцінки невідомих коефіцієнтів b_j беруться архівні дані за n років врожайності u_1, u_2, \dots, u_n і подекадні значення метеоданих x_1, x_2, \dots, x_k .

У матричній формі модель має вигляд

$$U = X \cdot B, \quad (2)$$

де

$$U = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \quad X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix}, \quad B = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{pmatrix},$$

U — вектор-стовпець фактичних значень залежної змінної (врожайності) розмірності n ; X — матриця значень метеоданих розмірності $n \times k$; B — вектор-стовпець коефіцієнтів параметрів, що підлягають оцінці, розмірності $(k + 1)$.

Значення коефіцієнтів b_j обчислюються методом найменших квадратів за показниками врожаю, опадів і температури з виразу

$$B = (X^T X)^{-1} (X^T Y). \quad (3)$$

Для перевірки значущості рівняння регресії використовувався F -критерій дисперсійного аналізу, в основу якого покладено розкладання загальної суми квадратів відхилень на складові частини [7]:

$$Q_Z = Q_R \cdot Q_V, \quad (4)$$

де Q_R — сума квадратів відхилень (від нуля), обумовлена регресією; Q_V — сума квадратів відхилень фактичних значень, обумовлена впливом випадкових і неврахованих у моделі факторів.

Процедура реалізації описаного методу прогнозування врожайності складається з двох етапів. На першому етапі обчислюються коефіцієнти b_0, b_1, \dots, b_n рівняння регресії за архівними даними врожайності й опадів, на другому — в рівняння з отриманими коефіцієнтами вводяться значення кількості опадів і температури попередніх місяців вегетації року й обчислюється прогнозоване значення врожайності. Програмна реалізація виконана мовою Python з використанням бібліотеки `scipy`.

Для перевірки методу проведено обчислювальний експеримент. Багатовимірною лінійною регресійною моделлю залежності врожайності від факторних змінних x_1, x_2, \dots, x_k має вигляд (1). Коефіцієнти моделі були отримані відповідно до виразу (3) і визначалися за даними 1992–2000 рр. для вказаних вище районів Київської області.

Для даного періоду за допомогою отриманої моделі було виконано прогнозування середньорайонної врожайності станом на травень місяць. Значення k для рівняння регресії

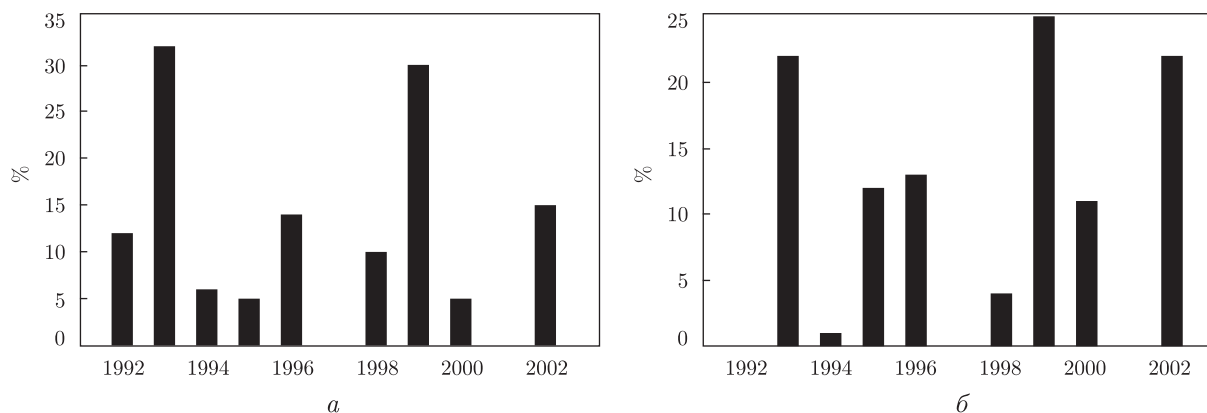


Рис. 1. Похибка прогнозування середньої врожайності озимої пшениці за роками станом на травень місяць для Київської області (а) та Білоцерківського району (б)

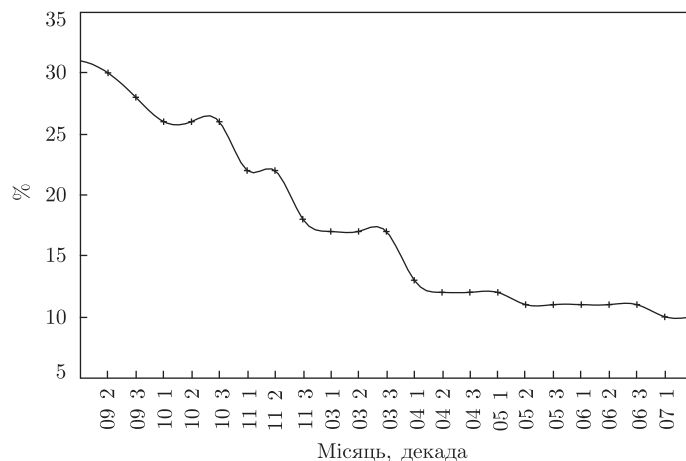


Рис. 2. Похибка прогнозування середньорайонної врожайності озимої пшениці в Київській області для декад сезону 2001–2002 рр.

вибирається залежно від дати, на яку необхідно одержати прогноз врожайності. Наприклад, у рівняння, за яким прогнозується врожайність на початок травня, повинні входити коефіцієнти, що враховують показники усіх попередніх місяців вегетації озимої пшениці. Процентні відхилення величини прогнозу від фактичного рівня врожайності наведені на рис. 1. Найбільша похибка отримана для 1993 і 1999 рр., коли через погодні умови врожаї озимої пшениці найбільше відрізнялись від середніх значень. Зокрема, у 1993 р. був вирощений один з найвищих врожаїв зерна за досліджуваний період. У 1999 р., навпаки, у зв'язку з раннім весняним потеплінням, що змінилося тривалими травневими заморозками, а потім і посухою з кінця травня до середини липня [8], був зафіксований значний недобір врожаю. Інші похибки варіюють у невеликому інтервалі значень 5–15% і відповідають природним коливанням врожаю поблизу лінії тренда середньої врожайності.

Для 2002 р. виконано прогнозування для всіх декад від посіву до збирання врожаю, динаміка зміни процентної похибки прогнозу від посіву до збирання врожаю показана на рис. 2. Крива похибки має спадний характер, що пояснюється додаванням нових метеода-

них, але з початку травня відбувається незначне підвищення точності прогнозу до мінімальної похибки $\approx 11\%$.

Таким чином, нами отримано рівняння, що зв'язують врожайність озимої пшениці зі значеннями температури й кількості опадів, у формі множинної регресії з коефіцієнтами, обчисленими за архівними даними врожайності, опадів і температури. Прогнозування врожайності виконано шляхом обчислення значення функції множинної регресії при підстановці в неї значень кількості опадів і температури, отриманих у попередні місяці вегетації пшениці. Розрахунок прогнозованих величин середньорайонної врожайності озимої пшениці Київської області для декад сезону 2001–2002 рр. свідчить про можливість прогнозування з точністю, яка у рамках проведеного експерименту не нижча від результатів прогнозування за контактними вимірами стану рослин і ґрунту.

1. Платонова Т. Ф. Прогнозирование динамики урожайности сельскохозяйственных культур. – Кишинев: Штиинца, 1983. – 88 с.
2. Страшная А. И., Андреева В. Н. О прогнозе урожайности зерновых и зернобобовых культур в Центральномчерноземном экономическом районе // Методы агрометеорологических прогнозов. – Ленинград: Гидрометеоиздат, 1991. – С. 3–22.
3. Лялько В. И., Сахацький А. И., Жолобак Г. М. и др. Опыт комплексирования многозональных снимков NOAA/AVHRR и “Landsat – 7” для прогноза урожайности озимой пшеницы (на примере районов Киевской области) // Косміч. наука і технологія. – 2003. – 9, № 4. – С. 99–103.
4. Лялько В. И., Сахацький О. И., Жолобак Г. М. Особливості прогнозування врожайності зернових культур за багатоспектральними даними // Багатоспектральні методи дистанційного зондування Землі в задачах природокористування. – Київ: Наук. думка, 2006. – С. 276–191.
5. Методические указания по составлению прогноза среднерайонного урожая озимой пшеницы по Киевской области. – Киев: УкрУГКС, 1987. – 21 с.
6. Кендал М., Стьюарт А. Многомерный статистический анализ и временные ряды. – Москва: Наука, 1976. – 736 с.
7. Королюк В. С., Портенко Н. И., Скороход А. В., Турбин А. Ф. Справочник по теории вероятности и математической статистике. – Киев: Наук. думка, 1978. – 584 с.
8. Дмитренко В. П. Оценка влияния температуры воздуха и осадков на формирование урожая основных зерновых культур. Методическое пособие. – Ленинград: Гидрометеоиздат, 1976. – 48 с.

Державна установа “Науковий центр
аерокосмічних досліджень Землі
ІГН НАН України”, Київ

Надійшло до редакції 23.07.2009

**G. M. Zholobak, V. G. Yakimtchuk, A. I. Sakhatsky,
K. Yu. Sukhanov, O. D. Ryabokononko,
Corresponding Member of the NAS of Ukraine O. D. Fedorovsky**

Simulation and prediction in problems of nature management (by the example of the winter wheat yield)

The relation for the winter wheat yield as a function of the temperature and precipitation is considered. The relation has the form of multiple regression. Its coefficients are calculated on the archival data on yields for a number of years and the decadal total precipitation and the average temperature. Yield forecast is carried out by calculating the values of the regression function. The example of the yield forecasting of winter wheat for the Kiev region in 2002 and the estimation of its accuracy depending on the forecast time are given.