

КІЛЬКІСНЕ ПРОГНОЗУВАННЯ ВРОЖАЙНОСТІ З ВИКОРИСТАННЯМ СУПУТНИКОВИХ СПОСТЕРЕЖЕНЬ

Розроблено методика використання знімків з геостаціонарних супутників для оцінки стану рослин, проведено статистичний аналіз залежності середньої урожайності по областях України від декадного НДВІ, побудована апроксимуюча функція для прогнозування урожайності, на прикладі озимої пшениці показано, що використання НДВІ є найбільш інформативним для раннього (травневого) прогнозування.

Ключові слова: *вегетаційний індекс, супутникові знімки, регресійний аналіз, урожайність.*

Вступ. Починаючи з 80-х років ХХ сторіччя для оцінки майбутнього урожаю все ширше стали застосовуватися супутникові спостереження за розвитком рослинності на заданій території. Найбільшого поширення у цих дослідженнях набули операції з різними спектральними діапазонами (каналами) супутникових фотографій, які мають відношення до параметрів рослинності у кожному пікселі знімка – так звані Вегетаційні Індекси (ВІ). Ці індекси виведені головним чином на базі емпіричних досліджень і залежать від специфіки відбиття різних частин спектра сонячного світла рослинністю.

У даний час існує більше ніж 100 різних методів розрахунків ВІ, які базуються на особливостях відбивної здатності рослин, ґрунтів та водних поверхонь. Найбільш поширеним ВІ в агрометеорології є Нормований Диференційний Вегетаційний Індекс (НДВІ), який розраховується всього у двох спектральних каналах: червоній зоні спектра *Pred* (0,62 - 0,75 мкм), де знаходиться максимум поглинання сонячної радіації хлорофілом, та в ближній інфрачервоній зоні *Pnir* (0,75 – 1,3 мкм) максимального відбиття енергії клітинною структурою листя.

НДВІ розраховується за формулою:
$$\frac{Pnir - Pred}{Pnir + Pred}$$
 і може набувати значень від -1 до +1 в залежності від поверхні, яка проектується на кожен піксель знімка. Для здорової зеленої рослинності значення НДВІ наближаються до +1, а для водних поверхонь (річок, озер, морів) – до -1. Всі інші поверхні (ґрунти без рослинності, хмари тощо) майже однаково відбивають *Pnir* та *Pred* частини спектра і тому значення НДВІ для них коливаються біля нуля.

Аналіз останніх досліджень і публікацій. Існує достатньо велика кількість робіт, де досліджується прогноз урожайності за допомогою вимірювання НДВІ. Найбільш детальні дослідження для України були зроблені у роботі [1], де проведена оцінка урожайності за допомогою супутників високої розподільної здатності. Але такі дослідження, окрім високої вартості знімків, мають декілька додаткових недоліків, спричинених як технологічними, так і суто метеорологічними особливостями супутникових спостережень. Бажання дослідити стан рослин на конкретній ділянці чи полі з однаковою сільськогосподарською культурою призводить до необхідності підвищувати роздільну здатність супутникових знімків таким чином, щоб в один піксель знімку потрапляла ділянка площею не більше ніж 10x10 м.

Знімки з високою роздільною здатністю можна одержати тільки з низькоорбітальних супутників, які пролітають над заданою ділянкою два-три рази на добу і до того ж “бачать” цю ділянку під різними кутами, що призводить до похибок у обчисленні НДВІ за рахунок різної товщини шару атмосфери між супутником і об’єктом дослідження. Корируючі коефіцієнти, обчислені за методикою, наведеною у роботі [2], все ж не надають суттєвого покращення, оскільки не враховують стан прозорості атмосфери при конкретній зйомці.

Наявність серпанку, смогу, пилу та навіть приземного вітру, який нахилиє листя під різними кутами до напрямку зйомки, зменшує значення НДВІ на деяку випадкову

величину, що додає суттєву похибку у дослідженнях і потребує осереднення результатів за досить тривалий період (2 – 3 декади), за який може змінитися і сам стан рослинності. Крім того, на початку травня, коли ефективність використання НДВІ найбільш інформативна (як показано у роботі [3]), у межах країни чи навіть області ще нема остаточних даних про те, на якому полі чи ділянці посіяна та чи інша культура, а замовником прогнозу врожайності є, як правило, уряд та інші державні структури, які цікавляться сумарним врожаєм по країні чи регіону.

Використані методи вимірювання НДВІ. У даній роботі для розрахунків НДВІ було використано знімки геостационарних супутників МЕТЕОСАТ-8 та МЕТЕОСАТ-9, які для території України мають роздільну здатність 5x7 км, тому в один піксель знімка будуть попадати декілька ділянок з різними культурами. При цьому з карти НДВІ виключаються ті території, що вкриті лісами чи водоймами, а також території міст та селищ міського типу. Однак за кожен день спостережень з'являється можливість отримувати до 50 якісних знімків всієї території України, зроблених під тим самим кутом та через однаковий шар атмосфери.

Опрацювання первинних знімків для обчислення НДВІ виконувалося за спеціальним алгоритмом осереднення, а саме:

- оскільки будь-яка похибка (наявність серпанку, пилу чи невеликої хмарності) тільки погіршує НДВІ, а похибок (фізичних чи оптичних), які покращують НДВІ, не існує, то для кожного пікселя знімків, зроблених у *Pnir* та *Pred* спектральних каналах, розраховувався НДВІ та порівнювався з тим самим пікселем за попередні знімки періоду опрацювання, та для НДВІ > 0 вибирався найбільший показник, а для НДВІ < 0 (водна поверхня) вибирався найменший. Таке опрацювання близько 500 знімків за декаду дозволяло виділити для кожного пікселя території найкращий НДВІ, який не залежав від повної чи часткової хмарності, серпанку, пилу, смогу та вітру за цей період;
- кожен обрахований піксель порівнювався з “маскою” території для визначення відношення цього пікселя до сільськогосподарських земель та вилучалися пікселі, які відносяться до садів, лісів, водних об'єктів та населених пунктів;
- значення НДВІ осереднювалися по площі складного контуру кожної області;

Таким чином, після первинного опрацювання знімків було визначено значення середнього НДВІ по кожній області за кожен декаду незалежно від сільськогосподарських культур, які в цих областях вирощувалися.

Такі дослідження виконувалися в Українському Гідрометцентрі впродовж п'яти років, що дозволило у першому наближенні провести статистичний аналіз залежності врожайності від НДВІ за кожен декаду. Для такого аналізу було використано дані фактичної урожайності озимої пшениці по кожній області України за останні 15 років. Звісно, що статистичний аналіз по п'яти значеннях подекадних НДВІ для кожної області, які отримано за п'ять років спостережень, буде мати дуже велику похибку апроксимації. Тому для підвищення достовірності аналізу випадкового процесу його було штучно приведено до стаціонарного та ергодичного процесу непрямих параметрів для кожної області, а саме:

- урожайність в різних областях України за один і той же рік відрізняється за рахунок різної якості земель та різних погодних умов, тому в якості непрямого параметра аналізу було використано відносне відхилення врожайності від середнього багаторічного значення по тій же області (у відсотках);
- так само і значення НДВІ посушливої Миколаївської або Одеської областей завжди менші, ніж значення НДВІ у Карпатах або Поліссі. Тому у якості непрямого параметра аналізу було використано відсоток відхилення кожної декадної величини НДВІ по кожній області від середньої декадної величини НДВІ за п'ять років спостережень.

Таким чином, аналізуючи відносно відхилення врожайності як функцію відносного відхилення НДВІ, можна (з деякими припущеннями) для статистичного аналізу використовувати не п'ять значень урожайності та декадної НДВІ для кожної області, а 5 (років) x 25 (областей) = 125 значень і відхилень, що мають приблизно однакову залежність для будь-якої області.

Методи обрахунку апроксимуючого виразу. Для апроксимації статистичних спостережень використовувався апарат регресійного аналізу, а саме, пошук апроксимуючої функції методом найменших квадратів [4]. Суть його полягає у знаходженні функції з певного класу за наступним критерієм: для кожної функції вводимо до розгляду наступну величину

$$\sum_{i=1}^k (y_i - f(x_i))^2 \quad (1)$$

де $(x_1, y_1), \dots, (x_k, y_k)$ - отримані пари параметр-значення. Тоді для шуканої функції значення виразу (1) має бути найменшим серед усіх інших функцій даного класу. Оскільки середньоквадратичне відхилення залежить лише від цієї величини, то і воно теж буде мінімальним.

Розглянемо випадок, коли будь-яку функцію із даного класу можна однозначно задати набором чисел. Таким є, наприклад, випадок, коли вибраний клас є множиною поліномів степенем не більше n - кожен багаточлен задається однозначно впорядкованим набором своїх коефіцієнтів. Тоді задача пошуку апроксимуючої функції таким методом зводиться до пошуку мінімуму функції

$$g(a_0, \dots, a_n) = \sum_{i=1}^k (y_i - f_{a_0, \dots, a_n}(x_i))^2 \quad (2)$$

на всьому \mathbb{R}^{n+1} . Як відомо [5], даний мінімум досягається тоді і тільки тоді, коли виконується система рівностей

$$\frac{\partial g}{\partial a_j} = 0; \quad j = 0, 1, \dots, n. \quad (3)$$

У випадку апроксимації набору значень поліноміальною функцією

$$f_{a_0, a_1, \dots, a_n}(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n \quad (4)$$

Підставляємо (4) у (2). Тоді для (3) отримуємо:

$$\begin{aligned} \frac{\partial g}{\partial a_j} &= \sum_{i=1}^k \frac{\partial}{\partial a_j} (y_i - (a_0 + a_1 x_i + \dots + a_j x_i^j + \dots + a_n x_i^n))^2 = \\ &= \sum_{i=1}^k 2(y_i - a_0 - a_1 x_i - \dots - a_j x_i^j - \dots - a_n x_i^n) x_i^j = 0 \end{aligned} \quad (5)$$

або, переносячи у ліву частину доданки із a_0, a_1, \dots, a_n

$$\left(\sum_{i=1}^k x_i^j \right) a_0 + \left(\sum_{i=1}^k x_i^{j+1} \right) a_1 + \dots + \left(\sum_{i=1}^k x_i^{j+n} \right) a_n = \sum_{i=1}^k y_i x_i^j; \quad j = 0, 1, \dots, n \quad (6)$$

Як бачимо (6) - система лінійних рівнянь відносно a_0, a_1, \dots, a_n .

Можна показати, що визначник головної матриці системи (6) відмінний від нуля при умові, що $k > n$, тобто кількість точок більша за порядок шуканого багаточлена. Така система однозначно розв'язана, і шукані коефіцієнти полінома можна знайти, наприклад, за допомогою методу Крамера [6].

Отже, маючи певну кількість пар значень НДВІ – Урожайність, можна зробити прогноз урожайності для всіх проміжних значень НДВІ. Для цієї задачі доцільно робити такий прогноз у вигляді полінома третього степеня, оскільки так досягається допустимий рівень відхилення і

у той же час частково зменшується похибка, спричинена відносно малою кількістю точок. Тобто, шукана залежність прогнозованого відхилення урожайності (P) від відхилення НДВІ (η) буде мати вигляд

$$P(\eta) = a_0 + a_1\eta + a_2\eta^2 + a_3\eta^3 \quad (7)$$

де коефіцієнти a_0, a_1, a_2, a_3 обчислюються з вищенаведеної системи рівнянь.

Результати досліджень. За даною методикою авторами було розроблено спеціальне програмне забезпечення, де за вхідні дані брались фактичні значення декадних НДВІ та фактичної урожайності по областях України у вигляді Excel-файлів, і розраховувались вищезазначені коефіцієнти апроксимації і середньоквадратичне відхилення фактичних значень від апроксимуючої кривої, окремо для усіх декад вегетації. Програма сама розраховувала середні значення НДВІ і урожайності по областях та їх відносні відхилення. Для прикладу були використані дані по озимій пшениці, хоча ця методика може застосовуватись для будь-якої культури.

Як показують розрахунки, найкращими з міркувань мінімізації середньоквадратичної похибки прогнозування виявилися перша та третя декади травня, що збігається з даними, отриманими у роботі [1]. Тобто середньоквадратичне відхилення фактичних значень урожайності за ці декади від апроксимуючої кривої забезпечує найменшу похибку, яка на 25% менша за відхилення від середнього значення урожайності. Ефективність лише раннього прогнозування таким методом зумовлена зокрема і тим, що як було зазначено раніше, роздільна здатність геостаціонарного супутника не забезпечує при обчисленні НДВІ відокремлення посівів пшениці від інших культур, подальша вегетація яких проходить по-різному.

Нижче наведені два графіки, на яких зображені фактичні значення відносної урожайності (точки) і значення апроксимації (лінія) як функції від відносного відхилення НДВІ, для третьої декади травня (рис. 1) і для другої декади липня (рис. 2). Як видно, травень місяць є більш інформативним. Досить велика мінливість значень урожайності в зоні малих відхилень НДВІ зумовлена тим, що раннє прогнозування (травень) не враховує впливу погодних умов наступних місяців поточного року на розвиток рослин (посуха, холодне літо тощо).

Для прогнозування майбутньої урожайності треба обчислити НДВІ за травень поточного року по кожній області, розрахувати його відносне відхилення від середніх показників та підставити ці значення у формулу апроксимації. Отримані дані і будуть відносним відхиленням прогнозованої урожайності від середніх багаторічних показників.

Варто зазначити, що з кожним додатковим роком спостережень за НДВІ зростає точність такого методу прогнозування, оскільки збільшується кількість статистичних даних для апроксимації.

Висновки.

Досліджено ще один метод прогнозування врожайності, який використовує спостереження з геостаціонарних супутників. Алгоритм обчислення прогнозу врожайності використовує наявні статистичні дані врожайності по областях та регресійний аналіз для знаходження апроксимуючої функції.

Показано, що використання знімків з геостаціонарних супутників, які не дозволяють виділити окремі поля та ділянки, ефективно для раннього прогнозування початкового періоду вегетації сільськогосподарських культур. Зазначений метод був перевірений при прогнозуванні врожайності озимої пшениці, але може застосовуватись для прогнозування будь-якої іншої культури.

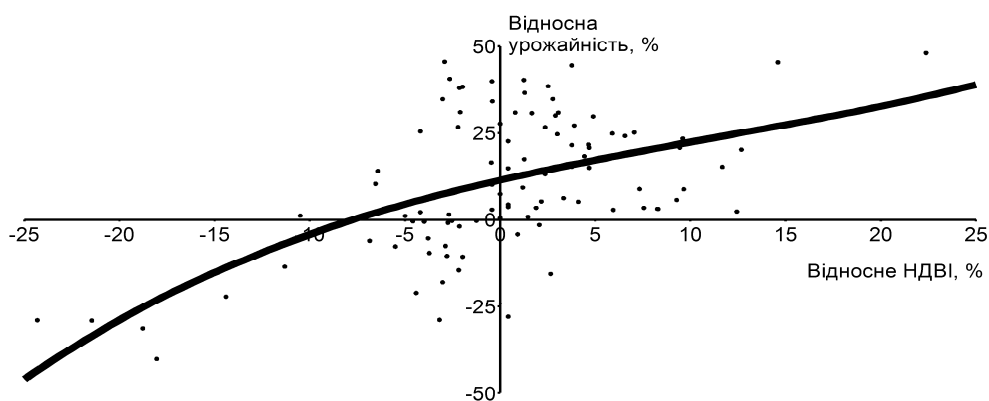


Рис. 1- Залежність відносної урожайності від НДВІ в третій декаді травня.

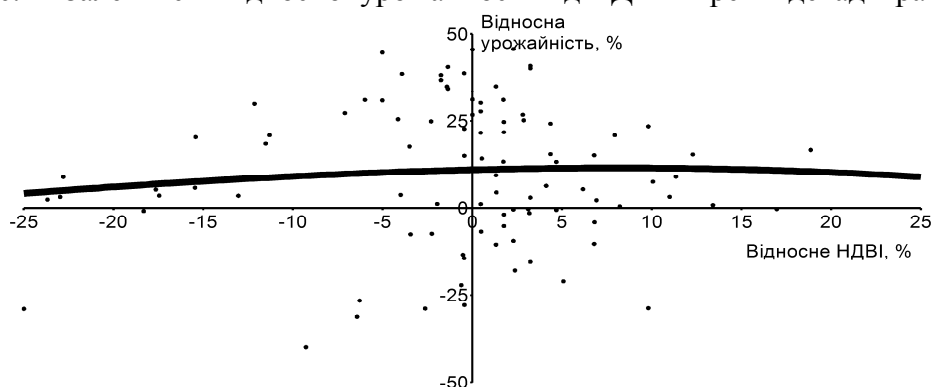


Рис. 2 - Залежність відносної урожайності від НДВІ в другій декаді липня

Список використаної літератури

1. Антоненко В.С., Гаценко Р.В. Оценка состояний посевов и прогноз урожайности озимой пшеницы в Украине по данным многоспектральной космической съёмки // Наук. Праці УкрНДГМІ. – 2005. – Вип. 254. – С. 55-71.
2. Кривобок А.А. Атмосферная коррекция многоспектральных спутниковых данных // Наук. праці УкрНДГМІ. – 1998. – Вип. 246. – С. 140-150.
3. Антоненко В.С., Гаценко Р.В., Бенедичук Л.Э. Временная изменчивость спектральных коэффициентов яркости посевов ярового ячменя и ее учет при аэрофотометрических обследованиях состояния посева // Труды УкрНИГМИ. – М.: Гидрометеиздат, 1992. – Вып. 244. – С.116-128.
4. Бахвалов Н. С., Жидков Н. П., Кобельков Г. Г. Численные методы. — 8-е изд. — М.: Лаборатория Базовых Знаний, 2000, — 460 с.
5. Корн Г., Корн Т. Справочник по математике для научных работников и инженеров. — М.: Наука, 1970. — С. 575-576.
6. Ильин В. А., Позняк Э. Г. Линейная алгебра: Учебник для вузов. — 6-е изд., стер. — М.: ФИЗМАТЛИТ, 2004. — 280 с.

Количественное прогнозирование урожайности с использованием спутниковых наблюдений.

Михайловський В.В., Адаменко Т.А., Кульбіда М.І.

Разработана методика использования снимков с геостационарных спутников для оценки состояния растений. Проведен статистический анализ зависимости средней урожайности по областям Украины от декадного НДВИ. Построена аппроксимирующая функция для прогнозирования урожайности. На примере озимой пшеницы показано, что использование НДВИ является наиболее информативным для раннего (майского) прогнозирования.

Ключевые слова: вегетационный индекс, спутниковые снимки, регрессионный анализ, урожайность.

Quantitative predictions of yield with the usage satellite observations

Mykhaylovskyy V., Adamenko T., Kulbida M.

The method of using images from geostationary satellites to estimate the conditions of plant is developed. A statistical analysis of the dependency of regional mean yield on ten-day NDVI is conducted. An approximating function for predicting yield is constructed. With the example of winter wheat, it is shown that the usage of NDVI is the most informative for the early (May) forecasting.

Keywords: vegetation index, satellite images, regression analysis, yield.