

УДК 551.445 : 630.0

Ю.Ф. Кобченко, О.Ю. Кобченко, В.О. Резуненко, З.А. Ковалевська

*Харківський національний університет імені В.Н. Каразіна*

## **БАГАТОВИМІРНИЙ АНАЛІЗ ФІТОПОГОДНИХ КОМПЛЕКСІВ НА МОВІ J**

**Постановка проблеми.** Важливим завданням гідрометеорологічної служби є прогнозування і порівняння сезонного ходу метеорологічних величин з метою оцінки їхнього впливу на урожай сільськогосподарських культур. Таке порівняння допомагає прогнозувати погодні умови, які активно впливають на приріст, розвиток і формування біомаси рослин. Це складна і трудомістка робота. Нами запропонований метод такого порівняння як для окремих метеоелементів, так і для їх співвідношення, що утворює погодні комплекси, у цілому можна говорити про багатовимірний аналіз фітопогодного комплексу.

**Вихідні передумови.** Широкого розвитку при обробці експериментальних гідрометеорологічних даних одержали методи статистичного аналізу. Емпіричне вирішення питання зв'язку урожаю з погодними умовами, а також моделювання на ЕОМ фізіологічних процесів, фотосинтезу і тепло-волого забезпечення рослин знайшло втілення у роботах В.П. Дмитренка [1], О.Р. Константинова [5], А.Н. Попєвого [6,7], О.Д. Сиротенко [8], В.Б. Сочави [9] та ін. Указані методи являють собою класичну методику дослідження у сільськогосподарській кліматології. Вони дозволили встановити основні закономірності впливу погодних факторів на процес формування врожаю. Локальність емпіричних зв'язків і відсутність загальної схеми вирішення питань, що враховують взаємний вплив визначальних факторів, не дозволяють повною мірою розраховувати на широкий підхід у дослідженнях, які повинні відповідати досягненням сучасного рівня науки.

Нами була зроблена спроба розглянути фізіологічні основи утворення фітомаси, залежність її від тепло-волого забезпечення культур [3-5]. До цього напряму відносяться дослідження механізму продукційного процесу, фотосинтетично активної радіації, тепло-волого споживання рослин, дефіцит вологості листа культури, фотометричні характеристики і методи статистичного аналізу з використанням програмного забезпечення ЕОМ.

Для реалізації цієї схеми створюється спеціальне проблемно-орієнтоване програмне забезпечення. Воно має модульну структуру і використовує алгоритмічну мову "J". Мова програмування J є інтерпретованою системою для розробки програмного забезпечення загального використання. Розроблена наприкінці 80-х ХХ ст. за участю Кенета Айверсона [10], мова J є сучасним втіленням своєї попередниці — мови АРЕ того

ж автора [11,12]. Мова J успадковує загальність та простоту APL, водночас застосовуючи сучасний набір функціональних символів, що робить її зручною для роботи на будь яких сучасних комп'ютерних платформах.

**Постановка завдання. Основна мета** роботи полягає у вивченні фітопогодних комплексів як об'єкта дослідження агрометеорології. Це міждисциплінарна галузь наукових досліджень, що розвивається на стику природничих, сільськогосподарських, технічних дисциплін і носить системний характер. Нові наукові галузі потребують і впровадження нових методів дослідження, серед яких запропоновані оціночно-моніторингові методи, моделювання природних процесів, метод системно-структурного аналізу, статистичні методи обробки вихідної інформації.

**Виклад основного матеріалу.** Великий обсяг фактичного матеріалу при дослідженні фітопогодних комплексів робить необхідним перехід до автоматизації процесу отримання кліматологічної інформації. Розроблені автоматизовані системи збору, обробки, контролю, зберігання і використання інформації для оперативних кліматологічних цілей. Важливим завданням при їх створенні є організація архівів, що слугує базою гідрометеорологічного банку даних (ГМБД). Це система інформаційних масивів великого обсягу з технічними і програмними засобами інформаційного обміну, призначених для розв'язання визначеного кола наукових і прикладних проблем. Побудова інформаційної бази фітопогодних комплексів заснована на об'єктивному поданні інформації, модель якої включає: назву архіву, кількість метеоелементів, назви і коди елементів, координати в часі і просторі. Таким чином, нами був створений ГМБД, що містив щоденні дані про температуру, відносну вологість повітря, радіацію, тепловий баланс, кількість посушливих днів за відповідний період. Система обробки інформації передбачала такі процедури: утворення банку даних, статистичну обробку даних, видачу даних за запитом. Статистична обробка первинної метеорологічної інформації для розв'язання кліматологічних завдань виконана за допомогою стандартної програми Stat Graf із використанням програмування на мові J, що складається зі стандартного статистичного пакету та спеціальної програми кластерного аналізу. Дані вводяться у систему у вигляді зручних для приготування текстових файлів. Система має можливість вибору способів обробки та задані параметрів. Результати можуть бути отримані у текстовому і графічному вигляді [12]. Особливості мови J складають багатий набір структурних та математичних операцій, що застосовуються однаково як для скалярних величин, так і для масивів даних, що відносить J до типу векторних мов програмування. В особливості, це стосується матричних операцій — таких, як детермінант, агрегативні операції (сума, максимум і т. ін.) по рядках або стовпцях, сторненість тощо. J також є

функціональною мовою, маючи набір операторів над функціями: вставку функції між елементами масиву, умовне та повторне застосування, оператори, задані користувачем [10],

Мова J особливо придатна для математичного, статистичного та логічного аналізу даних. Це потужний інструмент створення нових та кращих рішень існуючих проблем, але ще більш ефективний для пошуку рішень проблем, які не є добре з'ясованими. Компактність нотації, інтерактивність і потужна ефективність створюють середовище для всебічного аналізу даних і швидкого втілення нових можливих рішень завдяки прямому маніпулюванню ними.

Система J має: інтегровану середу розробки, що складає інтерактивну сесію, налащик тощо; стандартні бібліотеки, інструменти та пакети; графічний інтерфейс з управлінням подіями; дизайнер екранних форм для графічних програм користувача; інтерфейси до інших мов програмування, бібліотек та програм; інтегровану дво- та тривимірну графіку відображення файлів у пам'яті для високоефективних програм.

Для елементів мови J використовуються назви частин мови: дані називаються іменниками, функції дієсловами, а оператори над однією або двома функціями прислівниками і сполученнями відповідно [10]. Порядок виконання дієслів визначається дужками або, за їх відсутності, з права наліво, як для математичних функцій  $\log \sin x$ . Прислівники і сполучення виконуються раніше за дієслів. Від'ємні числа визначаються підкресленням а десятинна крапка у дробовому числі повинна мати перед собою нуль за відсутності цілої частини, наприклад, число 0.4 є -0,4.

Майже всі примітивні функції позначаються одним символом, за яким може слідувати крапка або двокрапка. Наприклад, діадичний  $>$  означає більше.  $>.$  повертає більший з двох аргументів,  $>:$  означає більше або дорівнює. Привласнення до локальної змінної у функції позначається  $x =$ . у, а до глобальної змінної через  $x =$ : у. Іменники можуть бути одиночними елементами, або атомами; одномірним масивом, або списком; двомірним масивом, або таблицею чи матрицею, а також багатомірним масивом. У випадку аргументів, що є масивами, дієслова застосовуються поелементно.

Дієслова можуть визначатися функціонально, тобто з невисловленими аргументами, наприклад,  $\text{mean} =$ :  $+ / \% \#$ , що визначає середнє як суму, поділену на кількість; а також висловлено, наприклад, монадично  $f =$ :  $3 : '2 + 4 * u'$  і діадично  $g =$ :  $4 : 'x + 2 * u'$ , де аргументи позначаються змінними x для лівого і у для правого.

Для прикладу розглянемо порівняння комплексів метеорологічних показників за допомогою кластерного аналізу. Залежно від поставлених завдань метеокомплекс можуть складатись із однорідних або різнорідних показників. Наприклад, нас цікавлять

метеоумови, за яких виникає пошкодження рослин атмосферною посухою. Це визначається трьома показниками: температурою повітря, вологістю повітря і швидкістю вітру. При цьому вологи у ґрунті повинно бути в необхідній кількості. За інших умов цей стан системи "ґрунт - рослина - повітря" визначався б як ґрунтова посуха. Найбільш зручним і достатнім для характеристики атмосферної посухи є аналіз двох показників — максимальної температури повітря і мінімальної відносної вологості повітря за добу. Існують певні критерії несприятливості співвідношення цих величин. Наприклад, максимальна температура повітря більша 30°C і мінімальна відносна вологість повітря менша 30 %. Якщо взяти добовий температурний максимум, то чим більше його величина і триваліший його період, то, таким чином, ми отримуємо не тільки інтенсивність посухи, але й відносність і тривалість несприятливих погодних умов.

Наприклад, необхідно проаналізувати хід максимальної температури під час посадки культур і в процесі їх вегетації. Для цього використані дані максимальної температури за квітень, травень і червень 1985-2005 рр., що дало змогу скласти комплексну характеристику кожного поточного року з дев'яти показників — це максимальна температура повітря за декаду. Таким чином, множина має 9 вимірів (декадні температури) і 20 порівняльних елементів (років). Показники в даному випадку будуть однорідні — температура повітря визначається в °C. У першому випадку при різнорідних показниках необхідно виконати нормування з метою їх порівнюваності. У другому випадку можна виконувати порівняння величин, використовуючи 9 величин без перетворення.

Класифікація фітопогодних комплексів виконується методом кластерного аналізу з використанням програмування на мові J. Загальний порядок розв'язання поставленого завдання такий: вибирається міра схожості показників; обчислюється матриця кожної пари показників; виконується об'єднання показників відповідно до вибраної міри схожості та вибраної процедури аналізу; виділяються групи елементів; аналізується однорідність груп та відмінності між групами.

Ієрархічний алгоритм має такий методичний аспект. За міру схожості двох показників прийнято евклідову відстань. На першому етапі роботи визначаються геометричні відстані між усіма 20 елементами. Результати підрахунків занесені у таблицю. Матриця геометричних відстаней симетрична. По діагоналі розміщені відстані елементів від "самих себе", які дорівнюють нулю. Кожна колонка (рядок) є список відстаней кожного елемента від усіх інших.

На другому етапі роботи виконується об'єднання показників відповідно з вибраною мірою схожості та вибраною процедурою аналізу і виділяються групи елементів. Результати об'єднання зводяться у таблицю.

Наступний етап роботи передбачає аналіз на однорідність груп та відмінності між групами. Розглядаючи, наприклад, першу колонку матриці (1985), можна сказати, що хід температури повітря за квітень-червень 1985 р. був найбільш подібний до температури 1993 р., бо йому відповідає найменша геометрична відстань 9.6. Можна визначити і найбільш несхожі елементи - по максимальній геометричній відстані. Таким буде елемент, що відповідає 1999 р. з відстанню 21.9. Можна врахувати і другий мінімум, що відповідає 2003 р. з відстанню 10.4, та другий максимум, що відповідає 1998 р. з відстанню 20.3.

Цей матеріал можна показати і у графічному вигляді — такому, як силует, на якому показано послідовність об'єднання показників або дендрограма зв'язку елементів (рис).

Методом візуалізації дендрограма зв'язку дозволяє скласти уявлення про головні тенденції у формуванні міри схожості чи відмінності спостережень та їх груп.

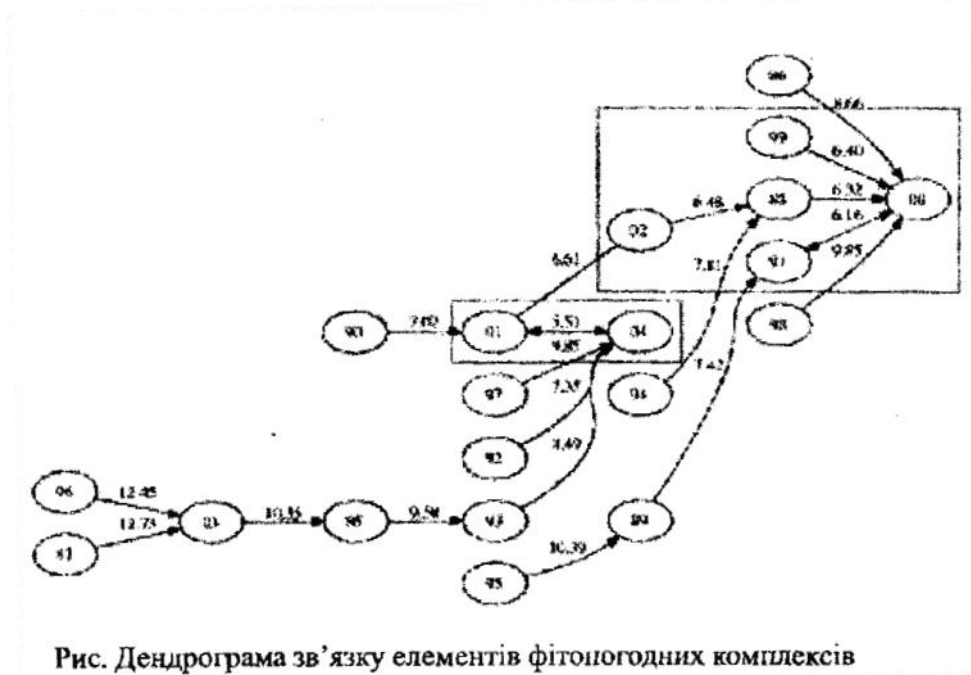


Рис. Дендрограма зв'язку елементів фітопогодних комплексів

**Висновки.** Статистична обробка масивів гідрометеорологічної інформації досліджуваних фітопогодних комплексів за період спостережень дозволила за допомогою аналітичних розрахунків і дендрограми виділити групу елементів, що пов'язує їх у відповідності до величини  $d$ , які не перевищують

10.5. На відповідному рисунку це показано центральною гілкою дендрита. Ці елементи утворюють деяку центральну групу (03, 85, 93, 04, 01, 02, 88, 00) із ходом максимальної декадної температури, близьким до середнього багатого річного ходу. Елементи, між якими геометричні відстані  $d$  не перевищують 6.5, об'єднані у групи з

високою мірою схожості. До першої групи елементів відносяться 01 і 04 де  $d = 5.51$ , а до другої - 00, 91, 88, 02, 99, де середня  $d = 6.32$ . Елементи 1996 і 1987 — найбільш аномальні і навіть від найближчого елементу 2003 вони значно відрізняються ( $d = 12.7$  і  $d = 12.5$  відповідно).

Таким чином, можна провести поглиблений аналіз на предмет схожості або відмінності фітопогодних комплексів за відповідні роки.

#### Література:

1. Дмитренко В.П. Агроклиматическая сценка условий произрастания сельскохозяйственных культур // Труды УкрНИГМИ - 1999. - Вып. 238. - С. 3-18.
2. Кобченко Ю.Ф., Резуненко В.А., Люсин С.В., Солоха Е.А. Моделирование процессов формирования биомассы и урожая сельскохозяйственных культур // Применение персональных компьютеров в научных исследованиях и учебном процессе. - Харьков: ХНУ, 2002. - С. 44-45.
3. Кобченко Ю.Ф., Клименко В.М., Протасов О.І. Еколого-меліоративний моніторинг як метод дослідження складних систем. ~ Харків: ХІСП, 2002. - С. 121-126.
4. Кобченко Ю.Ф., Резуненко В.А., Гвоздь Н.Л. Применение статистического критерия Хи-квадрат для анализа гидрометеорологической информации и прогнозирования развития погодных комплексов // Вісн. Харків, ун-ту. ~ 2003. ~ № 610: Геологія. Географія. Екологія. С. 143-150.
5. Константинов А.Р. Модель и метод расчета урожайности сельскохозяйственных культур//Труды УкрНИГМИ.- 1989. Вып.103. С. 23-38.
6. Полевой А.Н. Теория и расчет продуктивности сельскохозяйственных культур. - Л.: Гидрометеиздат, 1983.- 175 с.
7. Полевой А.Н. Прикладное моделирование и прогнозирование продуктивности посевов. - Л.: Гидрометеиздат, 1998. -319 с.
8. Сиротенко О.Д. Математическое моделирование водно-теплового режима и продуктивности агроэкосистем. - Л.: Гидрометеиздат, 1991. 167 с.
9. Сочава В.Б. Введение в учение о геосистемах. Новосибирск: Наука, 1988. - 427 с.
10. Iverson K. J Introduction & Dictionary. Jsoftware Inc., 2006. - 423 p.
11. JSoftware Inc. System development J, <http://jsoftware.com>. 2005. - 124 p.
12. Keith Smillie. J Companion for Statistical Calculations.
13. <http://www.cs.ualberta.ca/~smillie/Jpage/Jpage.htm>. 2006. - 214 p.