

стання сонячної енергії спостерігалось в другій половині вегетаційного періоду, коли листкова поверхня рослин досягає своєї найбільшої величини.

Зміни у фотосинтетичній діяльності рослин, які викликані різними умовами світлового режиму і кореневого живлення, в кінцевому підсумку визначають продуктивність цикорію кореневого як інтегрованого результату всіх фізіологічних процесів, що підтверджується даними табл. 3, які свідчать про те, що посів із квадратною формою площі живлення (35×35 см) і густрою рослин 81,6 тис./га забезпечив збір полісахариду (інуліну) 5,6 т/га, або на 0,7 т/га більше, ніж на контролі (45×22,5 см з густрою рослин 98,8 тис./га). Така ж закономірність спостерігалась і за збільшених доз добрив.

**Висновки і пропозиції.** Отже, одержані дані підтверджують припущення, що для формування високої продуктивності посівів цикорію кореневого рівномірність розподілу рослин на площі має більше значення, ніж їх загальна кількість на одиниці площі. До того ж чим більше площа живлення відхиляється від оптимуму (квадрата), тим більше спостерігається зниження врожайності.

#### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ:

1. Вильчук В.А. Цикорий. Ярославль: Верхневолжское книжное издательство, 1982. С. 80–90.
2. Манько А.Е. и др. Цикорий корнеплодный. Сахарная свекла. 1995. № 6. С. 24.
3. Роїк М.В., Борисик В.О., Зуев М.М., Курило В.Л., Мазуренко А.М., Пачевський І.А. Технологія вирощування і збирання цукрових буряків при комбінованій ширині міжрядь. Київ, 2006. 62 с.
4. Ткач О.В. Цикорій і особливості його вирощування. Наукові праці Інституту біоенергетичних культур і цукрових буряків: зб. наук. пр. К.: ФОП Корзун Д.Ю., 2012. Вип. № 15. С. 343–348.
5. Ткач О.В. Вплив площі живлення на урожайність цикорію кореневого. Наукові праці Інституту біоенергетичних культур і цукрових буряків: зб. наук. пр. К.: ФОП Корзун Д.Ю., 2015. Вип. № 23. 176 с.

УДК 631.53.01:631.526.3:633.18

### АНАЛІЗ І КЛАСИФІКАЦІЯ ХАРАКТЕРИСТИК СОРТОВОГО СКЛАДУ РИСУ ІЗ ЗАСТОСУВАННЯМ ІС STATISTICA

**Вожегов С.Г.** – д.с.-г.н., с.н.с.,

*Інститут рису Національної академії аграрних наук України*

**Цілінко М.І.** – к.с.-г.н., с.н.с.,

*Інститут рису Національної академії аграрних наук України*

**Зоріна Г.Г.** – аспірант,

*ДВНЗ «Херсонський державний аграрний університет»*

*У статті продемонстровано результати аналізу та класифікації кількісних та якісних показників сортів рису Преміум, Україна–96, Віконт за допомогою програмно-інформаційного комплексу Statistica, що призначений для статистичної обробки інформації. Як експериментальні дані було використано результати науково-дослідної роботи з удосконалення технологічних процесів вирощування насіння сучасних сортів рису, що проводилась протягом 2016–2018 рр. в Інституті рису НААН. За допомогою методу клас-*

теризації *k*-середніх засобами дисперсійного аналізу нами було знайдено міжгрупові дисперсії та статистичні рівні значущості досліджуваних характеристик кожного із сортів рису. У більшості аналізів за досліджувані роки домінували показники маси 1000 зерен, продуктивної куцистості, висоти рослин і пустозерності, кластером із найбільш оптимальними характеристиками за структурою врожаю виявився кластер «Розсадник розмноження і Супереліта».

**Ключові слова:** сорти рису, якість насіння, кластерний аналіз, маса 1000 насінин, висота рослин, енергія проростання.

***Vozhegov S.G., Tsilinko N.I., Zorina A.G. Analysis and classification of characteristics of rice varietal composition using Statistica***

В статті продемонстровані результати аналізу і класифікації кількісних і якісних показателів сортів рису Премиум, Україна–96, Виконт з допомогою програмно-інформаційного комплексу Statistica, який призначений для статистичної обробки інформації. В якості експериментальних даних були використані результати науково-дослідницької роботи по удосконаленню технологічних процесів вирощування семянь сучасних сортів рису, які проводились в період 2016–2018 рр. в Інституті рису НААН. З допомогою методу кластеризації *k*-середніх засобами дисперсійного аналізу даних нами були знайдені міжгрупові дисперсії і статистичні рівні значущості досліджуваних характеристик кожного із сортів рису. В більшості випадків аналізів за досліджувані роки домінували показники маси 1000 зерен, продуктивної куцистості, висоти рослин і пустозерності, кластером з найбільш оптимальними характеристиками по структурі врожаю виявився кластер «Розсадник розмноження і Супереліта».

**Ключевые слова:** сорта риса, качество семян, кластерный анализ, масса 1000 зерен, высота растений, энергия прорастания.

***Vozhegov S.G., Tsilinko N.I., Zorina A.G. Analysis and classification of characteristics of rice varietal composition using Statistica IP***

The article demonstrates the results of analysis and classification of quantitative and qualitative indicators of rice varieties Premium, Ukraine–96, Viscount using the Statistica software, which is intended for statistical processing of information. As experimental data, we used the results of research work on the improvement of technological processes of growing seeds of modern rice varieties that was conducted at the Institute of Rice NAAS in 2016–2018. Using the *k*-means method of clustering of the dispersion analysis of data, we found the intergroup dispersion and statistical significance levels of the studied characteristics of each of the rice varieties. In most cases of analysis in the studied years, the weight of 1000 grains, productive tillage capacity, plant height and emptiness of grain dominated; the cluster “Rozsadyk rozmnozhenia and Superelita” turned out to have with the most optimal characteristics of the yield structure.

**Key words:** rice varieties, seed quality, cluster analysis, weight of 1000 grains, plant height, germination energy.

**Постановка проблеми.** Сучасна селекційна робота з рису ведеться з використанням генетичного потенціалу сортозразків світової колекції та місцевих сортів культурного рису, які мають величезне різноманіття ознак і властивостей. Кращі з них після ретельного вивчення використовуються в гібридизації з метою отримання ліній, які поєднують усі необхідні параметри моделі високопродуктивного сорту [1, с. 15].

У системі різних агротехнічних та інших контрольованих людиною заходів, спрямованих на підвищення врожайності та забезпечення її стабільності, велика роль належить сортовому насінню [2, с. 212].

Кількісні та якісні характеристики сортового складу рису потребують ретельного їх вивчення та систематизації, тому застосування різноманітних новітніх методів статистичного аналізу для обробки й аналізу досліджуваних даних, що реалізовані в сучасних програмних комплексах, сьогодні є дуже актуальним.

**Аналіз останніх досліджень і публікацій.** Для виконання завдання із створення нових сортів і вдосконалення наявних необхідно виділити та знайти генетичні закономірності з добору вихідного матеріалу для селекційної роботи. У селекції

використовуваних сортів рису для півдня України важливо володіти інформацією про вияв кількісних ознак, особливостей їх успадковування і мінливості, ефективність доборів за ознаками, які пов'язані з адаптивним і продуктивним потенціалами у регіоні, тощо [3, с. 345].

Установлено, що на півдні України взаємодії «генотип – середовище» незначні за тривалістю вегетаційного періоду і довжиною стебла; більш високі – за довжиною головної волоті та кількістю колосків у головній волоті; неоднозначні – за кількістю зерен у головній волоті, масою зерна у волоті та рослині, масою 1000 зерен та врожайністю [4, с. 161; 5, с. 231; 6, с. 16]. Показано, що наявний генофонд рису дає змогу створювати цінний синтетичний селекційний матеріал із комплексом ознак продуктивності та якості зерна [7, с. 175].

Попередні дослідження показали, що наявний генетичний потенціал зразків рису може забезпечити створення сортів з урожайністю 8,5–10,0 т/га і більше, але для його реалізації необхідна теоретична база для селекції [8, с. 66]. Оптимізація умов вирощування рослин із метою отримання високоякісного насіння стимулює вчених до застосування більш поглибленого вивчення таких вегетативних ознак сортового складу рису, як продуктивна куцистість, висота рослин, довжина головної волоті тощо за допомогою сучасних методів аналізу. Одним із статистичних методів, які застосовуються в різноманітних сферах наукових досліджень, є метод кластерного аналізу, що впроваджений у сучасному програмно-інформаційному комплексі (ПК) Statistica [9].

**Постановка завдання.** Завданням дослідження було проаналізувати та отримати результати класифікації кількісних та якісних показників сортів рису Преміум, Україна–96, Віконт за допомогою ПК Statistica, що призначений для статистичного аналізу, візуалізації, прогнозування даних.

Як вхідні дані було взято результати науково-дослідної роботи з удосконалення технологічних процесів вирощування насіння сучасних сортів рису з метою підвищення посівних і врожайних властивостей, що проводилась протягом 2016–2018 рр. в Інституті рису НААН. Вирощування насіння рису в розсадниках первинних ланок супереліти й еліти проводили за прийнятим «Положенням про виробництво насіння первинних ланок та еліти зернових, зернобобових і круп'яних культур в Україні» (Методичні рекомендації, Київ, 1998 р.) на основі законів України «Про насіння» та «Про охорону прав на сорти рослин».

Використання кластер-аналізу для вирішення нашого завдання найбільш ефективне, тому що він призначений для об'єднання деяких об'єктів у класи (кластери) так, щоб в один клас потрапляли максимально схожі, а об'єкти різних класів максимально відрізнялися один від одного. Техніка кластеризації застосовується в найрізноманітніших галузях, наприклад у медицині, біології, економіці тощо [10]. Кластерний аналіз містить широкий набір методів аналізу даних, нами був використаний ієрархічний аналіз і метод к-середніх.

Для формування ієрархічної діаграми кластерного аналізу нами був вибраний метод повного зв'язку, що визначає відстань між кластерами як найбільшу відстань між будь-якими двома об'єктами в різних кластерах (тобто самими «видаленими сусідами»), на виході було отримано ієрархічне дерево для аналізу. Кожний вузол діаграми являє собою об'єднання двох або більше кластерів, положення вузлів на осі визначає відстань, на якій були з'єднані відповідні кластери [11, с. 32]. У кластеризації методом середніх програма переміщає об'єкти (тобто нагляди) з одних груп (кластерів) в інші для того, щоб одержати найбільш значимий результат відповідно до критерію дисперсійного аналізу.

Алгоритм кластеризації полягає в тому, що обчислення починаються з випадково вибраних спостережень, які стають центрами груп, після чого об'єктний склад кластерів міняється з метою мінімізації мінливості в середині кластерів.

Коли результати класифікації одержані, можна розрахувати середнє значення показників з кожного кластера, щоб оцінити, наскільки вони відрізняються між собою [12, с. 123].

**Виклад основного матеріалу дослідження.** У дослідженні всі дані з кожного року були внесені до таблиць Excel, а потім сконвертовані до таблиць Statistica залежно від категорій насіння: Розсадник розмноження (PP) – добазове насіння, Супереліта, Еліта – базове насіння, Сертифіковане насіння I – репродукції (СН-1) та стандартизовані відповідно до вимог системи для кожної категорії сортів насіння в розрізі років.

Використовуючи параметри методу об'єднання кластерів за правилом «повного зв'язку» та мірою близькості евклідової відстані, були сформовані горизонтальні діаграми ієрархічної кластеризації для кожного сорту рису за 2016, 2017, 2018 рр. досліджень.

На прикладі аналізу сформованих програмою ієрархічних дерев кластеризації для сорту Преміум (рис. 1) можна спостерігати, що за 2016 р. показники поділені на 3 кластери: 1 кластер – СН-1, 2 кластер – Супереліта та Еліта та 3 кластер – PP. За 2017 р. кластери перерозподілилися інакше: найбільш близькими за значеннями були елементи в 1 кластері – СН-1 та Еліта, в 2 кластері – PP та Супереліта; у 2018 р. ієрархічний розподіл за кластерами відбувся так само, як і у 2017 р.

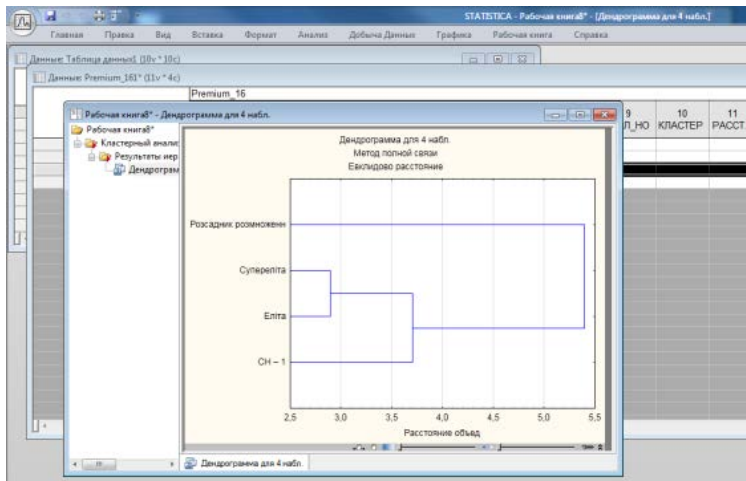


Рис. 1. Діаграма ієрархічного дерева кластеризації для сорту Преміум за 2016 р.

Використовуючи як середні або центри кластерів (початкової конфігурації) кількість кластерів, визначену попереднім аналізом, за допомогою методу k-середніх, нами було знайдено міжгрупові дисперсії показників структури врожаю кожного із сортів окремо з кожного досліджуваного року, які порівнюються з внутрішньогруповими дисперсіями для прийняття рішення, чи є середні для окремих змінних у різних сукупностях значимо різними. Ми можемо брати до уваги результати дисперсійного аналізу, порівнюючи для кожного вимірювання середні

(тобто вимірювання характеристик) між сумами (кластерами). На прикладі результатів дисперсійного аналізу показників із сорту Україна–96 за 2018 р. (рис. 2) можна зазначити, що, зважаючи на амплітуду (і рівні значущості) F-статистики, змінні «Висота рослин» і «Пустозерність» є головними у вирішенні питання про розподіл об'єктів за кластерами для цього аналізу.

Дисперсійний аналіз (Ukraina_18)						
	Між – SS	сс	Всередині – SS	сс	F	Значим. – p
Висота рослин, см	4,938365	2	1,042812	1	2,36781	0,417551
Маса гол. вол., г	0,990771	2	0,100557	1	4,92643	0,303548
К-сть з. гол. вол., шт	0,708802	2	0,146897	1	2,41258	0,414329
Пустозер., %	2,793037	2	0,016116	1	86,65633	0,075742
Прод. куш.	2,625082	2	2,542788	1	0,51618	0,701454

Рис. 2. Копія екрану з результатами дисперсійного аналізу показників сорту Україна–96 за 2018 р.

Іншим способом визначення природи кластерів була перевірка середніх значень для кожного кластера, що надається за допомогою графіку середніх. Зазвичай цей графік дає найкраще уявлення про результати.

Якщо поглянути на лінію кластера 1 графіку середніх сорту Віконт за 2016 р. дослідження (рис. 3), до якого належить категорія СН-1, то можна спостерігати у членів цього кластера (у нашому випадку – характеристик структури врожаю) найбільш високий показник пустозерності і найменш низькі показники інших характеристик. Члени кластера 3 (РР) характеризуються найнижчим показником пустозерності і мають більш високі показники вегетативних ознак сортового складу рису продуктивної кущистості, довжини головної волоті, кількості зерен із головної волоті, енергії проростання, схожості та маси 1000 зерен. Кластер 2 (Супереліта та Еліта) має переважно середні за величинами між 1 та 3 кластерами показники, за винятком найвищого показника висоти рослини та найнижчої довжини головної волоті.

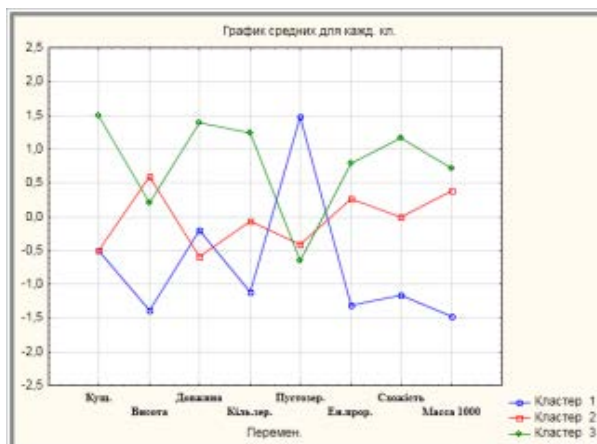


Рис. 3. Графік середніх структури врожаю сорту Віконт за 2016 р.

Динаміка показників графіку середніх для сорту Віконт за 2017–2018 рр. (рис. 4 (а), (б)) мала дещо іншу природу поведінки.

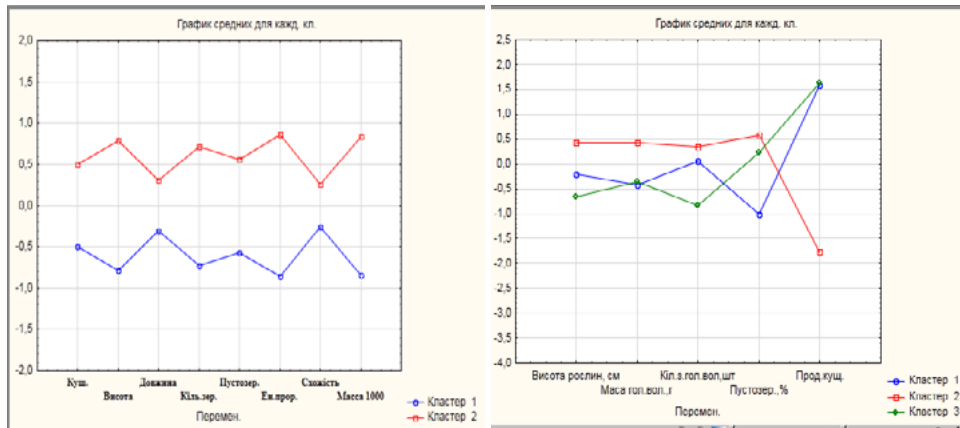


Рис. 4. Графік середніх для сорту Віконт за 2017 р. (а) та за 2018 р. (б)

Як продемонстровано на графіку середніх 2017 р., члени кластера 1 (СН-1 та Еліта) вирізнялись найнижчими показниками за всіма характеристиками структури врожаю, а члени кластера 2 (РР та Супереліта) мали більш високі показники. У 2018 р. елементи кластера 2 (РР) мали найвищі значення за винятком кущистості, найнижча висота рослин, кількість зерен із головної волоті була в кластері 3 (Еліта), а найвищу продуктивну кущистість мали елементи кластера 1 (СН-1 та Супереліта).

Методи ієрархічної кластеризації та к-середніх були проведені також для сортів Преміум і Україна–96 в розрізі сортів за 2016–2018 рр., для кожного випадку мала місце ситуація, де розподіл на кластери здійснювався для кожного аналізу окремо або на 2, або на 3 кластери з різними евклідовими відстанями між кластерами. Результати динаміки кластерів із найбільш оптимальними показниками структури врожаю в розрізі сортів і років надано в таблиці 1.

Таблиця 1

**Кластери з найбільш оптимальними показниками структури врожаю в розрізі сортів рису за 2016–2018 рр. досліджень**

Сорт	Найменування кластера		
	2016	2017	2018
Україна–96	РР+Супереліта	РР+Супереліта	РР+Супереліта
Преміум	РР	РР+Супереліта	Еліта
Віконт	РР+Супереліта	РР+Супереліта	РР

**Висновки і пропозиції.** За результатами досліджень і розрахунків доведена ефективність застосування методів кластерного аналізу, що впроваджені в сучасному ППК Statistica, на прикладі кластеризації показників структури врожаю сортів рису Віконт, Преміум, Україна–96 за 2016–2018 рр. досліджень. За допомогою цього програмного забезпечення нами було сформовано ієрархічні дендрограми кластерного аналізу з класифікацією показників сортового складу рису та розпо-

ділу їх за кластерами (категоріями насіння) з визначенням їхньої значущості, відмінностей між одержаними групами.

За допомогою методу кластеризації к-середніх засобами дисперсійного аналізу та наочної візуалізації даних нами було знайдено міжгрупові дисперсії досліджуваних характеристик кожного із сортів рису. Встановлено, що за статистичним рівнем значущості дисперсійного аналізу за 2016–2018 рр. досліджень у більшості аналізів домінували показники маси 1000 зерен, продуктивної кущистості, висоти рослин і пустозерності, найоптимальнішим за якісними та кількісними показниками в більшості аналізів виявився кластер «Розсадник розмноження та Супереліта».

Застосування в наших дослідках вищеописаних методів, упроваджених у системі Statistica, дало змогу зробити чимало корисних висновків для створення теоретично-практичної бази наукових основ для виконання завдання з виробництва насіння в первинних ланках супереліти, еліти і своєчасного проведення сортооновлення насінневого матеріалу в рисосійних господарствах.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ:

1. Вожегова Р.А. Становлення та розвиток селекції сільськогосподарських культур в Україні: історико-науковий аналіз: монографія. Херсон, 2007. С. 15–45.
2. Вожегова Р.А. Результати селекції рису в Україні. Збірник наукових праць Селекційно-генетичного інституту – Національного центру насіннєзнавства та сортовивчення. Одеса, 2007. Вип. 10 (50). С. 212–218.
3. Вожегова Р.А. Теоретичні основи і результати селекції рису в Україні: монографія. Херсон, 2010. 345 с.
4. Цілінко М.І. Ефективність використання факторіальної ознаки «маса головної волоті» на підвищення врожайності сортів рису. Зрошуване землеробство: міжвід. темат. наук. зб. Херсон: «Айлант», 2015. Вип. 64. С. 161–165.
5. Вожегова Р.А. Кущистість рослин рису залежно від біологічних властивостей сортів та умов вирощування. Зрошуване землеробство: міжвід. темат. наук. зб. Херсон: «Айлант», 2009. Вип. 52. С. 231–236.
6. Вожегова Р.А. Висота рослин та облистяність стебел рису залежно від біологічних властивостей сортів та умов вирощування. Таврійський науковий вісник: міжвід. темат. наук. зб. Херсон: «Айлант», 2009. Вип. 66. С. 16–26.
7. Петкевич З.З., Вожегова Р.А., Судін В.М. Генетичний потенціал рису та його використання в селекції. Зрошуване землеробство: міжвід. темат. зб. Херсон: «Айлант», 2008. Вип. 50. С. 175–178.
8. Вожегов С.Г., Еропкін В.А. Перспективи виробництва рису в Україні. Бюлетень Інституту зернового господарства. Дніпропетровськ, 2008. Вип. 33/34. С. 66–70.
9. Офіційний сайт розробника ІС Statistica. URL: [http://statsoft.ru/products/STATISTICA\\_Base](http://statsoft.ru/products/STATISTICA_Base) (дата звернення: 09.08.2018).
10. Francisco Torrens and Gloria Castellano. Elemental Classification of Tea Leaves Infusions: Principal Component, Cluster and Meta-analyses. 2018. URL: <https://www.intechopen.com/online-first/elemental-classification-of-tea-leaves-infusions-principal-component-cluster-and-meta-analyses> (дата звернення: 12.08.2018).
11. Дюран Б., Одел П. Кластерний аналіз / Пер. з англ. Є. Демиденко. М.: «Статистика», 1997. С. 32–38.
12. Мандель І.Д. Кластерний аналіз. М.: «Фінанси и статистика», 1988. С. 123–176.