

Ідентифікація сортів гречки їстівної *Fagopyrum esculentum* Moench за морфологічними ознаками з використанням алгоритму найближчих сусідів

Н. С. Орленко*, С. М. Гринів, С. П. Лікар, М. С. Юшкевич

Український інститут експертизи сортів рослин, вул. Генерала Родимцева, 15, м. Київ, 03041, Україна,
*e-mail: n.s.orlenko@gmail.com

Мета. Оцінити сорти гречки їстівної *Fagopyrum esculentum* Moench за морфологічними ознаками з використанням алгоритму найближчих сусідів за допомогою кластеризації сортів рослин та ідентифікувати групи схожих сортів. **Методи.** Аналітичний, математичний, статистичний. Як вхідну інформацію для статистичного оброблення було використано результатні дані експертизи на відмінність, однорідність та стабільність (ВОС) з бази даних автоматизованої інформаційної системи Українського інституту експертизи сортів рослин. Моделювання було проведено з використанням алгоритму найближчих сусідів, методом Machine Learning у середовищі статистичного пакета IBM SPSS Statistics «Statistical Package for the Social Sciences». У ролі параметрів моделі було використано такі типи змінних: цільова змінна – ознака «Рослина: тип росту», фокусна змінна – «Рослина: плоідність», мітка спостережень – «Назва сорту». Перелік показників (features) склали такі ознаки: «Рослина: плоідність», «Стебло: антоціанове забарвлення», «Суцвіття: антоціанове забарвлення бруньки», «Час початку цвітіння», «Рослина: тип росту», «Рослина: за висотою», «Квітка: забарвлення пелюсток», «Стебло: за довжиною», «Стебло: кількість вузлів», «Стебло: діаметр», «Час досягання», «Насінина: за довжиною», «Насінина: форма», «Насінина: забарвлення шкірки», «Насіння: маса 1000 шт.». До складу моделі було узято 25 сортів гречки їстівної вітчизняного та іноземного походження, які включені до Державного реєстру сортів рослин, придатних для поширення в Україні на 2020 рік. **Результати.** Унаслідок комп'ютерного моделювання було сформовано модель подібних сортів гречки їстівної за сімнадцятьма морфологічними ознаками, які відібрано із застосуванням частотного аналізу проявлення морфологічних ознак. Сформована модель містила 17 тренувальних об'єктів (сортів) та дев'ять контрольних об'єктів (сортів). Було ідентифіковано 22 групи подібних сортів гречки їстівної. **Висновки.** Сорти гречки їстівної *F. esculentum* було оцінено за морфологічними ознаками з використанням алгоритму найближчих сусідів за допомогою кластеризації подібних сортів рослин. Було ідентифіковано такі найподібніші групи сортів: перша – 'Крупнозелена', 'Дикуль', 'Дев'ятка' та 'Ювілейна 100'; друга – 'Рута', 'Мальва', 'Надійна' та 'Володар'; третя – 'Ксенія', 'Сімка', 'Селяночка' та 'Мальва'.

Ключові слова: сорт; ознака; модель; кластер; аналіз; статистичний метод.

Вступ

Гречка їстівна *Fagopyrum esculentum* Moench – важлива круп'яна і медоносна сільськогосподарська культура, яка широко культивується в багатьох країнах світу (зокрема у Китаї, Росії, Україні) та має цінні харчові і лікарські властивості [1]. Це зумовлює широкий інтерес закордонних та вітчизняних дослідників до споживчих характеристик й способів вирощування гречки, біологічних особливостей, морфологічних характеристик та господарсько-цінних показників нових сортів. Результати досліджень вітчиз-

няних і зарубіжних учених розкривають потенціал урожайності сортів гречки їстівної, біохімічний склад товарної продукції, взаємозв'язок параметрів продуктивності рослин з кількісними, якісними та псевдо-якісними морфологічними ознаками фенотипу [2]. Сучасні дослідження морфологічних ознак гречки їстівної розкриті в наукових працях щодо карликової лінії гречки їстівної, G410, її стійкість при різних умовах росту та режимі успадкування [3].

Під час впровадження у виробництво нових сортів сільськогосподарських культур з поліпшеними властивостями, селекціонери потребують охорони прав інтелектуальної власності селекційних досягнень на сорти рослин, які перебувають у комерційному обігу. Обов'язковою вимогою поширення сорту на території України є відповідність його критеріям відмінності, однорідності та стабільності [4]. Цей метод ідентифікації нових сортів відбувається через морфологічний опис проявлення ознак рослин під час росту та розвитку [5].

Natalia Orlenko
<http://orcid.org/0000-0003-0494-2065>
Svitlana Hryniv
<http://orcid.org/0000-0002-2044-4528>
Svitlana Likar
<http://orcid.org/0000-0002-2590-7376>
Marina Yushkevich
<http://orcid.org/0000-0002-9538-2616>

Тестування на відмінність, однорідність та стабільність – це спосіб визначення, чи відрізняється сорт, на який подано заявку для включення до Державного реєстру сортів рослин, придатних для поширення в Україні, від існуючих сортів того ж виду (відмінність), чи виражені характеристики, що використовують для встановлення відмінності, рівномірно (однорідність) і що ці характеристики не змінюються протягом наступних поколінь (стабільність) [6].

Тестування на ВОС, відповідно до міжнародних вимог (Guidelines for the conduct of tests for distinctness, uniformity and stability TG/278/1, UPOV) та Методики проведення експертизи сортів гречки їстівної (*Fagopyrum esculentum* Moench) на відмінність, однорідність і стабільність, виконують у польових умовах протягом двох послідовних вегетаційних сезонів [7]. Варто зауважити, що на сьогодні не існує жодного статистичного методу для однозначного встановлення відмінності одного сорту від іншого. На думку авторів, спростити технологічну процедуру з визначення відмінності сорту можна за допомогою автоматизованого групування найподібніших сортів за кодами проявлення ідентифікаційної ознаки з подальшим візуальним визначенням відмінних сортів у межах групи [6, 7].

Цей метод було апробовано під час аналізу колекції сортів салату посівного різновидів *var. capitata* і *var. longifolia* та люпину жовтого [8, 9]. Варто зазначити, що ідентифікація груп подібних сортів можлива із застосуванням різних статистичних методів, які відносять до інтелектуального аналізу даних. Одним з таких методів є Machine Learning, а одним з алгоритмів є метод найближчих сусідів [10]. У теорії інтелектуального аналізу близькі один до одного спостереження називаються «сусідами», а сам метод має назву машинного навчання і був розроблений як спосіб розпізнавання структури даних, що мають неточну відповідність та складну структуру спостережень [11, 12, 13]. Саме таким є проявлення морфологічних ознак сортів рослин гречки їстівної.

Мета досліджень – оцінити сорти гречки їстівної *F. esculentum* за морфологічними ознаками з використанням алгоритму найближчих сусідів за допомогою кластеризації сортів рослин за ідентифікації схожих сортів.

Методика та методи досліджень

Під час дослідження використано дані результатів експертизи сортів гречки їстівної на відмінність, однорідність і стабільність, які було включено до Державного реєстру

сортів рослин, придатних для поширення в Україні (Реєстр сортів рослин України). Станом на 25.04.2020 у Реєстр сортів рослин України включено 25 сортів *F. esculentum*, з яких 23 – вітчизняного походження [14].

Джерелом вихідних даних слугували морфологічні кодові формули сортів гречки їстівної, які складаються з відповідних кодів проявлення ідентифікаційних ознак вегетативних і генеративних органів рослин. Оброблення результатів експертизи здійснювали методом машинного навчання із застосуванням алгоритму найближчих сусідів. Розрахунки виконували у середовищі статистичного пакета IBM SPSS Statistics «Statistical Package for the Social Sciences» із застосуванням таких параметрів класифікаційного методу Nearest Neighbor Analysis [11, 15]. У ролі цільової змінної (target optional) було обрано ознаку «Рослина: тип росту», фокусна змінна (focal case identifier) – «Рослина плоїдність», мітка спостережень (case label) – «Назва сорту». Перелік показників (features) склали такі ознаки: «Рослина: плоїдність», «Стебло: антоціанове забарвлення», «Суцвіття: антоціанове забарвлення бруньки», «Час початку цвітіння», «Рослина: тип росту», «Рослина: за висотою», «Квітка: забарвлення пелюсток», «Стебло: за довжиною», «Стебло: кількість вузлів», «Стебло: діаметр», «Час досягання», «Насінина: за довжиною», «Насінина: форма», «Насінина: забарвлення шкірки», «Насіння: маса 1000 шт.».

Для побудови моделі групи сортів в параметрі «кількість найближчих сусідів (k)» задавали: визначити три найбільше схожих сорти. Так як у ролі цільової змінної обрано «Рослина: тип росту», то як альтернативу (forced entry) задавали ознаки, що рекомендовані методикою для групування сортів-кандидатів із подібними загальновідомими сортами на групи для полегшення оцінювання відмінності [16]. Для групування використовували ознаки, які не варіюють або дуже слабо варіюють у межах сорту. Ці ознаки можуть бути використані окремо або в комбінаціях з іншими, а саме «Час початку цвітіння», «Рослина: за висотою», «Квітка: забарвлення пелюсток», «Стебло: кількість вузлів», «Час досягання», «Насінина: забарвлення шкірки».

Для комп'ютерного оброблення даних морфологічних ознак було використано номінальну і порядкову шкалу [17, 18]. Номінальна шкала застосовувалась для групування подібних сортів за такими морфологічними ознаками: «Рослина: плоїдність», «Стебло:

антоціанове забарвлення», «Суцвіття: антоціанове забарвлення бруньки», «Рослина: тип росту», «Квітка: забарвлення пелюсток», «Насінина: форма», «Насінина: забарвлення шкірки», а порядкова шкала – для групування за ознаками: «Час початку цвітіння», «Рослина: за висотою», «Стебло: за довжиною», «Стебло: кількість вузлів», «Стебло: діаметр», «Час досягання», «Насінина: за довжиною», «Насіння: маса 1000 шт.».

Під час ідентифікації сортів гречки їстівної використовували метод морфологічного опису ознак із застосуванням алгоритму найближчих сусідів. Для індивідуального добору

високопродуктивних рослин гречки їстівної використовували кластерний аналіз [19, 20].

Результати досліджень

Національні сортові ресурси гречки їстівної представлено на сьогодні загальновідомими сортами вітчизняної та іноземної селекції. Джерелознавча база даних автоматизованої інформаційної системи УІЕСР містить інформацію щодо 172-х сортів гречки їстівної, заявки на які подавали починаючи з 1936 по теперішній час. Щорічно база аналітичних даних поповнюється у середньому 3–4-ма сортами (рис. 1).

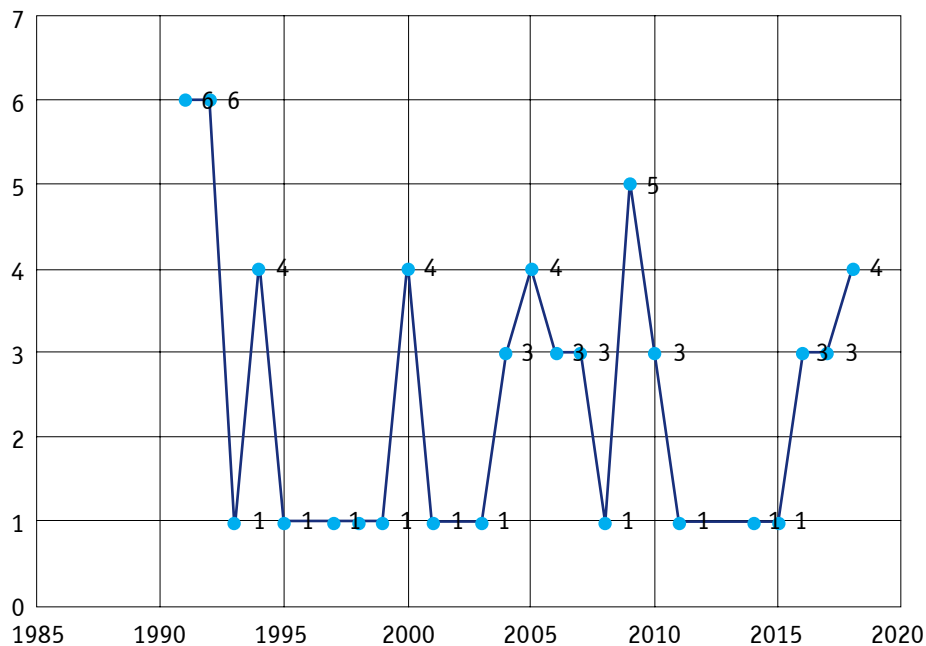


Рис. 1. Динаміка формування національних сортових ресурсів гречки їстівної за період 1991–2020 рр.

Заявниками цих сортів є установи Національної академії аграрних наук України – Інститут рису, Національний науковий центр «Інститут землеробства», Сумський інститут агропромислового виробництва та інші установи, такі як ТОВ «Агрофірма Суми-Насіння», ТОВ науково-виробнича агрофірма «Землеробець», ТОВ науково-виробниче мале підприємство «Антарія», ВНДІ зернобобових і круп'яних культур, Подільський державний аграрно-технічний університет.

Унаслідок моделювання з використанням методу найближчих сусідів було сформовано модель подібних сортів *F. esculentum*, візуалізацію якої подано на діаграмі «Predictor Space» (рис. 2).

Ця модель містить 25 сортів гречки їстівної, з яких 21 сорт ('Арата', 'Крупнозелена', 'Амазонка', 'Оранта', 'Ювілейна 100', 'Малинка', 'Квітнева', 'Ярославна', 'Дикуль',

'Син-3/02', 'Рута', 'Руслана', 'Селяночка', 'Ксенія', 'Надійна', 'Ольга', 'Мальва', 'Сімка', 'Кам'янчанка', 'Медова', 'Володар') увійшли у навчальну вибірку та чотири сорти ('Перлина Поділля', 'Дев'ятка', 'Софія', 'Воля') – у контрольну.

Повний перелік груп сортів гречки їстівної за результатами моделювання низки (17) морфологічних ознак наведено у таблиці 1.

Аналіз за діапазонами відстаней даних таблиці свідчить, що у діапазоні відстаней від 0 до 2 виокремлено сім груп сортів. Сорти 'Крупнозелена', 'Дикуль', 'Дев'ятка' та 'Ювілейна 100' є найподібнішими між собою. Зауважимо, що нульові коефіцієнти відстаней, які розраховано для сортів, як у сорту 'Крупнозелена' свідчать про те, що ці сорти є подібними за 17 ознаками, які відібрані з 21 ознаки, що визначені методикою проведення експертизи на ВОС.

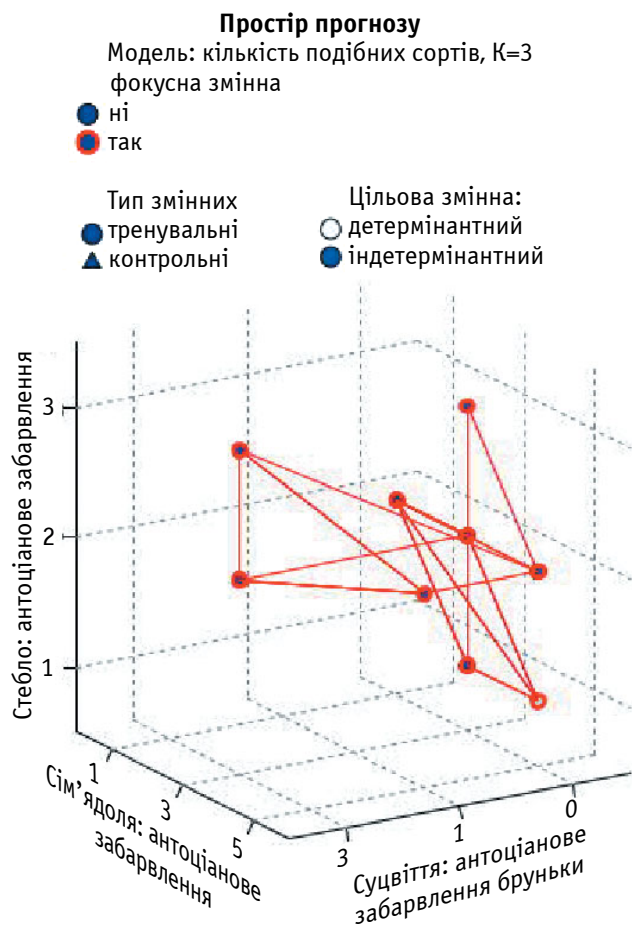


Рис. 2. Головна діаграма моделі найбільшої подібності

Антоціанове забарвлення сім'ядольних листочків у сорту 'Крупнозелена' відсутнє, а у сортів 'Дикуль', 'Дев'ятка' та 'Ювілейна 100' – слабе. Форма основи листової пластинки у сортів 'Крупнозелена' та 'Ювілейна 100' зрізана, а у сортів 'Дикуль', 'Дев'ятка' – стріло-подібна.

Кількість вузлів на стеблі у сорту 'Ювілейна 100' середня, а у сортів 'Крупнозелена', 'Дикуль', 'Дев'ятка' – велика.

Наступні дві групи найподібніших сортів складають: 'Рута', 'Мальва', 'Надійна', 'Володар' і 'Ксенія' та сорти 'Сімка', 'Селяночка' і 'Мальва'.

Висновки

Сорти гречки їстівної *F. esculentum* оцінено за морфологічними ознаками з використанням алгоритму найближчих сусідів за допомогою кластеризації подібних сортів рослин. Унаслідок моделювання з використанням методу найближчих сусідів було сформовано модель подібних сортів. Найподібнішими групами сортів були:

- перша: 'Крупнозелена', 'Дикуль', 'Дев'ятка' та 'Ювілейна 100';
- друга: 'Рута', 'Мальва', 'Надійна' та 'Володар';
- третя: 'Ксенія', 'Сімка', 'Селяночка' та 'Мальва'.

Для тесту на відмінність сортів рослин рекомендовано застосовувати аналітичний ме-

Таблиця 1

Групування подібних сортів за відстанню до найближчих сусідів (k – Nearest Neighbor and Distance)

Фокусний сорт	Назва сорту			Відстань		
	Перший подібний	Другий подібний	Третій подібний	до 1-го	до 2-го	до 3-го
'Арата'	'Оранта'	'Ювілейна 100'	'Амазонка'	1,41	2,00	2,45
'Крупнозелена'	'Дикуль'	'Дев'ятка'	'Ювілейна 100'	0,00	0,00	1,41
'Амазонка'	'Ювілейна 100'	'Оранта'	'Софія'	2,00	2,00	2,00
'Оранта'	'Арата'	'Ювілейна 100'	'Амазонка'	1,41	1,41	2,00
'Ювілейна 100'	'Оранта'	'Дикуль'	'Дев'ятка'	1,41	1,41	1,41
'Малинка'	'Володар'	'Мальва'	'Перлина Поділля'	0,00	1,41	2,00
'Перлина Поділля'	'Володар'	'Малинка'	'Надійна'	2,00	2,00	2,00
'Ярославна'	'Софія'	'Медова'	'Дикуль'	2,00	2,00	2,00
'Дев'ятка'	'Дикуль'	'Ювілейна 100'	'Медова'	0,00	1,41	2,00
'Дикуль'	'Дев'ятка'	'Ювілейна 100'	'Медова'	0,00	1,41	2,00
'Син-3/02'	'Мальва'	'Надійна'	'Володар'	2,45	2,83	2,83
'Софія'	'Мальва'	'Ярославна'	'Дикуль'	1,41	2,00	2,00
'Рута'	'Мальва'	'Надійна'	'Володар'	0,00	1,41	1,41
'Руслана'	'Дикуль'	'Дев'ятка'	'Ювілейна 100'	1,41	1,41	2,00
'Селяночка'	'Ксенія'	'Сімка'	'Надійна'	1,41	1,41	1,41
'Ксенія'	'Сімка'	'Селяночка'	'Мальва'	0,00	1,41	1,41
'Надійна'	'Мальва'	'Селяночка'	'Перлина Поділля'	1,41	1,41	2,00
'Мальва'	'Надійна'	'Володар'	'Малинка'	1,41	1,41	1,41
'Сімка'	'Ксенія'	'Селяночка'	'Мальва'	0,00	1,41	1,41
'Кам'янчанка'	'Мальва'	'Надійна'	'Володар'	1,41	2,00	2,00
'Медова'	'Дикуль'	'Дев'ятка'	'Ярославна'	2,00	2,00	2,00
'Володар'	'Малинка'	'Мальва'	'Перлина Поділля'	0,00	1,41	2,00

тод для групування за морфологічними ознаками сортів рослин і для інших ботанічних таксонів. Застосований метод інтелектуального аналізу даних запропоновано протестувати за експертизи на ВОС сортів-кандидатів.

Використана література

1. Генофонд и селекция крупяных культур. Гречиха / под ред. В. А. Драгавцева. Санкт-Петербург : ВИР, 2006. 196 с.
2. Kalinova J., Vrchotova N. Level of Catechin, Myricetin, Quercetin and Isoquercitrin in Buckwheat (*Fagopyrum esculentum*). The Growth of Selected Weeds. *J. Agr. Food Chem.* 2009. Vol. 57, Iss. 7. P. 2719–2725. doi: 10.1021/jf803633f
3. Minami M., Ujihara A., Campbell C. Morphology and inheritance of dwarfism in common buckwheat line, G410, and its stability under different growth conditions. *Breed. Sci.* 1999. Vol. 49, Iss. 1. P. 27–32. doi: 10.1270/jsbbs.49.27
4. Методика проведення кваліфікаційної експертизи сортів рослин на придатність до поширення в Україні. Загальна частина / уклад.: С. О. Ткачик, О. І. Присяжнюк, Н. В. Лещук. 4-те вид., випр. і допов. Вінниця : ФОП Корзун Д. Ю., 2016. 120 с.
5. Атлас морфологічних ознак сортів рослин групи зернових (наочне доповнення до Методики проведення інспектування насінницьких посівів зернових видів) / підгот.: В. І. Сорока, Н. В. Лещук, А. В. Андрищенко. Київ : Фенікс, 2011. 205 с.
6. Методика проведення експертизи сортів гречки їстівної (*Fagopyrum esculentum* Moench) на відмінність, однорідність і стабільність. *Методика проведення експертизи сортів рослин групи зернобобових та круп'яних на відмінність, однорідність і стабільність* / за ред. С. О. Ткачик. 2-ге вид., випр. і допов. Вінниця : Корзун Д. Ю., 2016. С. 129–140. URL: <https://sops.gov.ua/uploads/page/5a5f1ca54a208.pdf>.
7. Test Guidelines for the conduct of tests for distinctness, uniformity and stability of Buckwheat (*Fagopyrum esculentum* Moench) (TG /278/1) / UPOV. Geneva, 2012. 27 с.
8. Орленко Н. С., Лещук Н. В., Симоненко Н. В., Таганцова М. М., Стадніченко О. А. Особливості використання засобів machine learning під час ідентифікації подібних сортів рослин (на прикладі *Lactuca sativa* L. var. *capitata*). *Вісник ПДАА*. 2019. № 4. С. 233–240. doi: 10.31210/visnyk2019.04.309
9. Орленко Н. С., Мажуга К. М., Душар М. Б., Маслечкін В. В. Порівняльний аналіз ієрархічних методів кластеризації, придатних для оброблення даних морфологічних ознак сортів рослин. *Вісник ПДАА*. 2019. № 2. С. 261–269. doi: 10.31210/visnyk2019.02.35
10. Lantz B. Machine Learning with R. Learn how to use R to apply powerful machine learning methods and gain an insight into real-world applications. Birmingham : Packt Publishing Ltd., 2013. 396 p.
11. Бююль А., Цефель П. SPSS: искусство обработки. Анализ статистических данных и восстановление скрытых закономерностей. Санкт-Петербург : ДиаСофтЮП, 2002. 608 с.
12. Лесковец Ю., Раджараман Ю., Ульман Дж. Анализ больших наборов данных. Москва : ДМК, 2016. 498 с.
13. Наследов А. Д. IMBSPSS Statistics 20 и AMOS: профессиональный статистический анализ данных. Санкт-Петербург : Питер, 2013. 416 с.
14. Державний реєстр сортів рослин, придатних для поширення в Україні на 2020 рік. Київ, 2020. 500 с.
15. Rotarov K. O., Kadyrova L. R., Mukhametshina R. R. Embryological features and seed productivity of tartary buckwheat. *J. Fundam. Appl. Sci.* 2017. Vol. 9, Iss. 15. P. 1462–1471. doi: 10.4314/jfas.v9i1s.796
16. Тищенко В. Н. Эффективность использования нового селекционного индекса в селекции озимой пшеницы. *Факторы экспериментальной эволюции организмов*. 2004. Т. 2. С. 266–270.
17. Compton M. E. Statistical methods suitable for the analysis of plant tissue culture data. *Plant Cell Tiss. Organ Cult.* 1994. Vol. 37, Iss. 3. P. 217–242. doi: 10.1007/BF00042336
18. Марманис Х., Бабенко Д. Алгоритмы интеллектуального Интернета. Переводные методики сбора, анализа и обработки данных. Москва : Символ, 2011. 480 с.
19. Тищенко В. М. Кластерний аналіз, як метод індивідуального добору високопродуктивних рослин озимої пшениці в F₂. *Селекція і насінництво*. 2005. Вип. 89. С. 125–137.
20. Тищенко В. Н., Чекалин Н. М., Зюков М. Е. Использование кластерного анализа для идентификации и отбора высокопродуктивных генотипов озимой пшеницы на ранних этапах селекции. *Факторы экспериментальной эволюции организмов*. 2004. Т. 2. С. 270–278.

References

1. Dragavtsev, V. A. (Ed). (2006). *Genofond i selektsiya krupyanykh kul'tur. Grechikha* [The Gene bank and breeding of groat crops. Buckwheat]. St. Petersburg: VIR. [in Russian]
2. Kalinova, J., & Vrchotova, N. (2009). Level of Catechin, Myricetin, Quercetin and Isoquercitrin in Buckwheat (*Fagopyrum esculentum* Moench), Changes of Their Levels during Vegetation and Their Effect on The Growth of Selected Weeds. *J. Agr. Food Chem.*, 57(7), 2719–2725. doi: 10.1021/jf803633f
3. Minami, M., Ujihara, A., & Campbell, C. (1999). Morphology and inheritance of dwarfism in common buckwheat line, G410, and its stability under different growth conditions. *Breed. Sci.*, 49(1), 27–32. doi: 10.1270/jsbbs.49.27
4. Tkachyk, S. O., Prysiazhniuk, O. I., & Leshchuk, N. V. (2016). *Metodyka provedennia kvalifikatsiinoi ekspertyzy sortiv roslin na prydatnist do poshyrennia v Ukraini. Zahalna chastyna* [Methodology of conducting qualification examination of plant varieties for suitability for distribution in Ukraine. General part]. (2nd ed., rev. and enl.). Vinnytsia: FOP Korzun D. Yu. [in Ukrainian]
5. Soroka, V. I., Leshchuk, N. V., & Andriushchenko, A. V. (2011). *Atlas morfolohichnykh oznak sortiv roslin hrupy zernovykh (naochne dopovnennia do Metodyky provedennia inspektuvannia nasinnytskykh posiviv zernovykh vydiv)* [Atlas of morphological characteristics of plant varieties of cereals (visual addition to the Methodology of inspection of seed crops of cereals)]. Kyiv: Feniks. [in Ukrainian]
6. Tkachyk, S. O. (2016). Methods of examination of varieties of edible buckwheat (*Fagopyrum esculentum* Moench) for distinctness, uniformity and stability. In *Metodyka provedennia ekspertyzy sortiv roslin hrupy zernobobovykh ta krupianykh na vidminnist, odnoridnist i stabilnist* [Methods of examination of plant varieties of legumes and cereals for distinctness, uniformity and stability]. (2nd ed., rev. and enl.). Vinnytsia: Korzun D. Yu. Retrieved from <https://sops.gov.ua/uploads/page/5a5f1ca54a208.pdf> [in Ukrainian]
7. UPOV. (2012). *Test Guidelines for the conduct of tests for distinctness, uniformity and stability of Buckwheat (Fagopyrum esculentum Moench)* (TG /278/1). Geneva: N.p.
8. Orlenko, N. S., Leshchuk, N. V., Symonenko, N. V., Tahantsova, M. M., & Stadnichenko, O. A. (2019). Peculiarities of using machine learning tools during identification of similar plant varieties (*Lactuca sativa* L. var. *capitata* as example). *Visnik PDAA* [Bulletin of Poltava State Agrarian Academy], 4, 233–240. doi: 10.31210/visnyk2019.04.30 [in Ukrainian]
9. Orlenko, N. S., Mazhuha, K. M., Dushar, M. B., & Maslechkyn, V. V. (2019). Comparative analysis of clustering methods suitable for plant varieties morphological characteristics data processing. *Visnik PDAA* [Bulletin of Poltava State Agrarian Academy], 2, 261–269. doi: 10.31210/visnyk2019.02.35 [in Ukrainian]
10. Lantz, B. (2013). *Machine Learning with R. Learn how to use R to apply powerful machine learning methods and gain an insight into real-world applications*. Birmingham: Packt Publishing Ltd.
11. Byuyul, A., & Tsefel, P. (2002). *SPSS: iskuststvo obrabotki. Analiz sta tisticheskikh danykh i vosstanovlenie skrytykh zakonmer-*

- nostey [SPSS: Arts of processing. Analysis of statistical data and restoration of hidden patterns]. St. Petersburg: DiaSoft-YuP. [in Russian]
12. Leskovets, Yu., Radzharaman, Yu., & Ulman, D. (2016). *Analiz bol'shikh naborov dannykh* [Analysis of large data sets]. Moscow: DMK. [in Russian]
 13. Nasledov, A. D. (2013). *IMB SPSS Statistics 20 i AMOS: professional'nyy statisticheskiy analiz dannykh* [IMB SPSS Statistics 20 and AMOS: professional statistical data analysis]. St. Petersburg: Piter. [in Russian]
 14. *Derzhavnyi reestr sortiv roslyn, prydatnykh dlia poshyrennia v Ukraini na 2020 rik* [State register of plant varieties suitable for dissemination in Ukraine in 2020]. (2020). Kyiv: N.p. [in Ukrainian]
 15. Potapov, K. O., Kadyrova, L. R., & Mukhametshina, R. R. (2017). Embryological features and seed productivity of tartary buckwheat. *J. Fundam. Appl. Sci.*, 9(15), 1462–1471. doi: 10.4314/jfas.v9i1s.796
 16. Tyshchenko, V. N. (2004). Efficiency of using a new selection index in winter wheat breeding. *Fakt. Eksp. Evol. Org.* [Factors in Experimental Evolution of Organisms], 2, 266–270. [in Russian]
 17. Compton, M. E. (1994). Statistical methods suitable for the analysis of plant tissue culture data. *Plant Cell Tiss. Organ Cult.*, 37(3), 217–242. doi: 10.1007/BF00042336
 18. Marmanis, Kh., & Babenko, D. (2011). *Algoritmy intelektual'nogo Interneta. Peredovye metodiki sbora, analiza i obrabotki dannykh* [Algorithms of the intellectual Internet. Advanced techniques for data collecting, analyzing and processing]. Moscow: Simvol. [in Russian]
 19. Tyshchenko, V. M. (2005). Cluster analysis as a method of individual selection of high-yield winter wheat plants in F_2 . *Selekcii i nasamnictvo* [Plant Breeding and Seed Production], 89, 125–137. [in Ukrainian]
 20. Tishchenko, V. N., Chekalin, N. M., & Zyukov, M. E. (2004). Using cluster analysis to identify and select highly productive winter wheat genotypes in the early stages of selection. *Fakt. Eksp. Evol. Org.* [Factors in Experimental Evolution of Organisms], 2, 270–278. [in Russian]

УДК 633.12

Орленко Н. С.*, **Грынив С. Н.**, **Лекарь С. П.**, **Юшкевич М. С.** Идентификация сортов гречихи посевной *Fagopyrum esculentum* Moench по морфологическим признакам с использованием алгоритма ближайших соседей // Plant Varieties Studying and Protection. 2020. Т. 16, № 2. С. 137–143. <https://doi.org/10.21498/2518-1017.16.2.2020.209221>

Украинский институт экспертизы сортов растений, ул. Генерала Родимцева, 15, г. Киев, 03041, Украина, e-mail: n.s.orlenko@gmail.com

Цель. Оценить сорта гречихи посевной *Fagopyrum esculentum* Moench по морфологическим признакам с использованием алгоритма ближайших соседей с помощью кластеризации сортов растений и идентифицировать группы похожих сортов. **Методы.** Аналитический, математический, статистический. В качестве входной информации для статистической обработки были использованы результаты экспертизы на отличимость, однородность и стабильность (ООС) из базы данных автоматизированной информационной системы Украинского института экспертизы сортов растений. Моделирование было проведено с использованием алгоритма ближайших соседей, методом Machine Learning в среде статистического пакета IBM SPSS Statistics «Statistical Package for the Social Sciences». В качестве параметров модели были использованы следующие типы переменных: целевая переменная – признак «Растение: тип роста», фокусная переменная – «Растение: плоидность», метка наблюдений – «Название сорта». Перечень показателей (features) составляли следующие признаки: «Растение: плоидность», «Стебель: антоциановая окраска», «Соцветие: антоциановая окраска почки», «Время начала цветения», «Растение: тип роста», «Растение: по высоте», «Цветок: окраска лепестков», «Стебель: по длине», «Стебель: количество узлов», «Стебель: диаметр», «Время со-

зревания», «Семя: по длине», «Семя: форма», «Семя: окраска кожуры», «Семена: масса 1000 шт.». В состав модели были взяты 25 сортов гречихи посевной отечественного и зарубежного происхождения, которые включены в Государственный реестр сортов растений, пригодных для распространения в Украине на 2020 год. **Результаты.** В результате компьютерного моделирования была сформирована модель подобных сортов гречихи посевной по семнадцати морфологическим признакам, которые отображены с применением частотного анализа проявления морфологических признаков. Сформированная модель содержала 17 тренировочных объектов (сортов) и девять контрольных объектов (сортов). Было идентифицировано 22 группы подобных сортов гречихи посевной. **Выводы.** Сорта гречихи посевной *Fagopyrum esculentum* Moench оценили по морфологическим признакам с использованием алгоритма ближайших соседей с помощью кластеризации подобных сортов растений. Были идентифицированы такие, наиболее подобные, группы сортов: первая – 'Крупнозэлэна', 'Дыкуль', 'Девятка' и 'Ювилейная 100'; вторая – 'Рута', 'Мальва', 'Надийна' и 'Володар'; третья – 'Ксения', 'Симка', 'Селяночка' и 'Мальва'.

Ключевые слова: сорт; признак; модель; кластер; анализ; статистический метод.

UDC 633.12

Orlenko, N. S.*, **Hryniv, S. M.**, **Likar, S. P.**, & **Yushkevich, M. S.** (2020). Identification of buckwheat varieties *Fagopyrum esculentum* Moench by morphological characters by applying the nearest neighbors algorithm. *Plant Varieties Studying and Protection*, 16(2), 137–143. <https://doi.org/10.21498/2518-1017.16.2.2020.209221>

*Ukrainian Institute for Plant Variety Examination, 15 Henerala Rodymtseva St., Kyiv, 0341, Ukraine, *e-mail: n.s.orlenko@gmail.com*

Purpose. Evaluate buckwheat *Fagopyrum esculentum* Moench varieties by morphological characteristics by using the nearest neighbors algorithm and identify groups of similar varieties by plant varieties clustering procedure. **Methods.** Analytical, mathematical, statistical. As input information for statistical processing, the results of the examination for distinctness, uniformity and stability (DUS) from the database of the automated information system of the Ukrainian Institute for the Examination of Plant Varieties were used. The simulation was performed using Nearest Neighbors algorithm, which belong to Machine Learning method. Data processing was carried out by using trial version of the statistical package IBM SPSS Statistics “Statistical Package for the Social Sciences”. The following types of variables were used as model parameters: target (optional) variable is “Plant: growth type”, focal case identifier is “Plant: ploidy”, case name is Varieties Name, feature variables are “Cotyledon: anthocyanin coloration”, “Stem: anthocyanin coloration”, “Inflorescence: anthocyanin coloration of bud”, “Time of beginning flowering”, “Plant: height”, “Leaf blade: shape of base”, “Leaf blade:

intensity of green color”, “Flower: size”, “Flower: color of petals”, “Flower: length of pedicel”, “Plant: total number of flower clusters”, “Stem: length”, “Stem: number of nodes”, “Stem: diameter”, “Time of maturity”, “Seed: length”, “Seed: shape”, “Seed: skin color”, “Seed: 1000 seed weight”. This model contains 25 buckwheat varieties, included in the State Register of plant varieties suitable for distribution in Ukraine in 2020. These varieties are of foreign and domestic origin. **Results.** The model of similar buckwheat varieties was a result of computer modeling. This model was based on seventeen morphological features, selected using frequency analysis of buckwheat morphological features. The generated model contained 17 training objects (varieties) and nine control objects (varieties). 22 groups of similar varieties of common buckwheat were identified. **Conclusion.** The most similar groups of varieties are the following: first – ‘Krupnozelená’, ‘Dykul’, ‘Deviatka’ that ‘Yuvileina 100’; second – ‘Ruta’, ‘Malva’, ‘Nadiina’ and ‘Voldar’; third – ‘Kseniia’, ‘Simka’, ‘Selianochka’ and ‘Malva’.

Keywords: variety; feature distinctness; model; buckwheat; statistical methods.

Надійшла / Received 12.05.2020

Погоджено до друку / Accepted 14.06.2020