

Особливості фіксації метеорологічних умов у сховищі даних кваліфікаційної експертизи сортів рослин

С. І. Мельник, Н. В. Лещук, Н. С. Орленко*, Є. М. Стариченко, К. М. Мажуга, Є. А. Шкапенко

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

Мета. Розробити багатовимірну модель сховища даних кваліфікаційної експертизи сортів рослин для фіксації метеорологічних умов у взаємозв'язку з фенологічними стадіями розвитку сортів, які проходять ВОС та ПСП експертизу. **Методи.** Для проведення досліджень з установленням основних структурних елементів багатовимірного сховища даних застосовано методи індукції, дедукції, аналізу та синтезу. У процесі проектування сховища застосовано концепцію W. H. Inmon, адаптовану для сфери сільського господарства та аграрного бізнесу. **Результати.** Проаналізовано етапи кваліфікаційної експертизи сортів рослин та розглянуто методологічні підходи до створення багатовимірної моделі сховища даних. Висвітлено особливості застосування сховищ даних для збереження результатів кваліфікаційної експертизи сортів рослин на відмінність, однорідність та стабільність (ВОС) та придатність сорту до поширення (ПСП). Особливу увагу приділено реалізації взаємозв'язку результатів кваліфікаційної експертизи сортів рослин з даними метеорологічних спостережень на різних фенологічних етапах росту й розвитку рослин, відповідно до записів в електронному польовому журналі. Спроектовано даталогічну схему сховища даних та реалізовано її у середовищі MS SQL Server. **Висновки.** Визначено джерела надходження даних у сховища даних та реалізовано багатовимірну модель сховища даних за схемою «сніжинки». Подано схему сховища даних, яка забезпечує зв'язок між метеорологічними умовами проведення польових дослідів та результатними даними кваліфікаційної експертизи й налічує чотири таблиці вимірів. Для кожної таблиці вимірів та таблиці фактів визначено атрибутивний склад даних. Сховище даних практично використовується для аналізу впливу погодних умов на показники ВОС та ПСП експертизи.

Ключові слова: метеорологічні дані; фенологічні фази; виміри; факти; атрибути; період вегетації; міжфазний період.

Вступ

Динамічні зміни метеорологічних показників, що виявляються в чергуванні вологих і посушливих періодів, теплих і холодних зим, а також аномальному впливі умов періоду вегетації, спричиняють значну мінливість показників кваліфікаційної експер-

тизи на придатність сортів до поширення. Крім того, екстремальні умови погоди (посухи, суховії, заморозки, повені тощо) ініціюють масове поширення й розвиток збудників хвороб та шкідників сільськогосподарських культур.

Світова практика показує, що під час спостереження за ростом і розвитком сортів рослин важливо фіксувати погодні умови періоду вегетації [1–4]. Тому проведення кваліфікаційної експертизи сортів рослин з визначення показників придатності їх до поширення не буде об'єктивним, якщо не враховувати вплив чинників довкілля на формування показників урожайності та якості, а також інших важливих характеристик прояву придатності сортів до поширення [5]. Прояв морфологічних ознак теж формується під впливом погодних умов періоду вегетації рослин [6], що також важливо враховувати під час дослідження.

Serhii Melnyk

<https://orcid.org/0000-0002-5514-5819>

Nadiya Leschuk

<https://orcid.org/0000-0001-625-3702>

Nataliia Orlenko

<https://orcid.org/0000-0003-0494-206>

Yevhen Starychenko

<http://orcid.org/0000-0001-8608-5268>

Kostiantyn Mazhuha

<http://orcid.org/0000-0002-1434-8687>

Yevheniya Shkapenko

<http://orcid.org/0000-0002-8600-1543>

Вплив чинників довкілля також обумовлює необхідність удосконалювати науково-методичне забезпечення експертизи шляхом визначення складу математико-статистичних інструментів, який би дав змогу врахувати ці чинники під час оцінювання результатів науково-технічної кваліфікаційної експертизи сортів рослин на відмінність, однорідність та стабільність (ВОС) та на придатність сорту до поширення (ПСП) завдяки автоматизованому отриманню точніших вихідних даних із деталізацією у розрізі фенологічних фаз розвитку рослин [7].

Відповідно врахування описаних вище чинників потребує внесення змін у структуру оперативної бази та сховища даних аналітичної інформаційної системи (АІС) УІЕСР.

Першим кроком у проектуванні багатовимірної моделі сховища даних є впорядкування сукупності фактів в інформаційні сутності. Наступний крок – це дизайн сховища даних, що передбачає об'єднання всієї структури даних. Зауважимо, що дані кваліфікаційної експертизи, як і дані погодних умов, визначені у певні фенологічні фази росту й розвитку відповідних сортів рослин містять велику кількість фактів, які взаємозалежні один від одного. Вони мають бути перетворені із сукупності інформаційних повідомлень в оперативні дані. Тому ефективність отримання даних залежить як від алгоритмів їх оброблення, так і від структури самих даних.

Сьогодні є дві альтернативні теорії проектування сховища даних – це теорія W. H. Inmon [8] та R. Kimball [9]. Згідно з теорією W. H. Inmon передбачено, що дані повинні перебувати на низькому рівні ступені деталізації і в третій нормальній формі (3НФ, 3NF). За цим підходом розвиток сховища відбувається ітераційно, тобто за потреби додається одна таблиця за один раз, що забезпечує лише незначну зміну схеми даних. Тому такий підхід до проектування сховища ще називають спіральним.

Альтернативним є підхід R. Kimball, де модель архітектури сховища даних ще називають просторовим сховищем, у якому первинні дані трансформуються в інформацію, придатну для використання на етапі підготовки даних. При цьому обов'язково беруться до уваги вимоги до швидкості опрацювання інформації та якості даних. Як і в моделі W. H. Inmon, підготовка даних починається зі скоординованого добування даних із первинних джерел. Причому частина операцій відбувається централізовано, наприклад, підтримання й зберігання загальних довідкових даних, інші дії можуть бути розподіленими [10, 11].

До особливостей сховища даних порівняно з базою даних належать [12]:

- отримання інформації з різних джерел даних (зокрема й реляційних баз даних) у деталізованому та агрегованому вигляді (зберігаються результати застосування функцій агрегації – суми, середнього значення, максимуму, мінімуму тощо);
- багатовимірне подання інформації, тобто ігноруються деякі вимоги нормалізації (дотримують максимум 3-ї нормальної форми), що значно підвищує швидкість опрацювання інформації, оскільки зменшує кількість операцій з'єднання;
- наявність метаданих для опису джерел метаданих та структури самого сховища даних – у базах даних також використовують словники для опису структур даних, а у сховищах даних метадані (словники, дані про дані);
- наявність пакетного завантаження даних у сховище даних та вивантаження даних;
- наявність процедур аналізу та отримання нових даних;
- орієнтованість даних на аналітичне, а не на статистичне опрацювання.

Перспективним напрямом є створення сховища даних як багатовимірного куба та відповідних йому схем «зірок» та «сніжинок», що використовують під час проектування баз даних [13, 14]. Багатовимірне моделювання є найвдалішим підходом до проектування сховища даних для прогнозного аналізу даних [15], зокрема в аграрній сфері [16].

Застосуванню методики проектування сховища даних за концепцією W. H. Inmon у сфері сільського господарства та аграрного бізнесу присвячено низку публікації [17–20]. Особливості багатовимірної моделі аграрних даних розкрито в роботі [21].

Відповідно концептуальні підходи до створення сховища даних для використання його в кваліфікаційній експертизі сортів рослин слід напрацьовувати з використанням уже розроблених теорій та методів проектування сховищ даних.

Мета дослідження – розробити багатовимірну модель сховища даних кваліфікаційної експертизи сортів рослин для фіксації метеорологічних умов у взаємозв'язку з фенологічними стадіями розвитку сортів, які проходять ВОС та ПСП експертизу.

Методи дослідження

Під час проведення дослідження предметної сфери застосовували методи індукції, дедукції, аналізу та синтезу. Метод індукції

було застосовано для переходу від теоретичних знань щодо впливу метеорологічних умов на розвиток рослин на різних фазах вегетативного і генеративного періоду розвитку, а також на прояв морфологічних ознак та показників придатності до поширення й до складу інформаційних об'єктів сховища даних. Метод дедукції застосовано під час переходу від знання загальних закономірностей теорії проектування баз даних (БД) та сховищ даних (СД) до окремого його прояву під час реалізації багатовимірної моделі сховища даних. Метод аналізу було застосовано під час виокремлення зі складного об'єкта БД його частин (тобто підсхем БД), які відповідають даним кваліфікаційної експертизи (ВОС та ПСП). Метод синтезу дав змогу відтворити як ціле модель сховища даних, що складається з даних результатів кваліфікаційної експертизи сортів рослин ВОС та ПСП, даних польових журналів і метеорологічних умов. Проектування багатовимірної сховища даних проведено відповідно до класичної теорії W. H. Inmon. Зокрема, застосовували спіральне зведення даних за предметною орієнтацією, історичними періодами. Сховище даних організовано відповідно до етапів експертизи та технологічних процедур, які пов'язані з проведенням дослідів. Предметна організація даних у сховищі сприяє як значному спрощенню аналізу, так і пришвидшенню

виконання аналітичних запитів. Перед завантаженням даних метеорологічних умов, які отримано з бази даних програмно-апаратного комплексу Meteotrek, проведено їхнє статистичне оброблення з використанням агрегатних функцій. Сховище даних АІС УІЕСР побудоване на основі клієнт-серверної архітектури й даталогічна схема реалізована в середовищі MS SQL Server v.19.

Результати дослідження

За результатами вивчення методологічних основ створення сховища даних УІЕСР установлено, що воно повинне використовуватись для інформаційного забезпечення детального аналізу метеорологічних чинників, що впливають на результати кваліфікаційної експертизи сортів рослин, а не для задоволення потреб в автоматизації щоденних процедур оброблення інформації, які вже забезпечено функціонуванням оперативної бази даних УІЕСР.

Предметна орієнтація сховища даних зумовлена особливостями кваліфікаційної експертизи сортів рослин, а саме визначенням критеріїв відмінності, однорідності та стабільності (ВОС-тест) та визначенням господарсько-цінних показників придатності сортів до поширення на території України (ПСП), а також забезпеченням зв'язку між стадіями фенологічного розвитку кожного дослідного сорту з метеорологічними умовами (рис. 1).

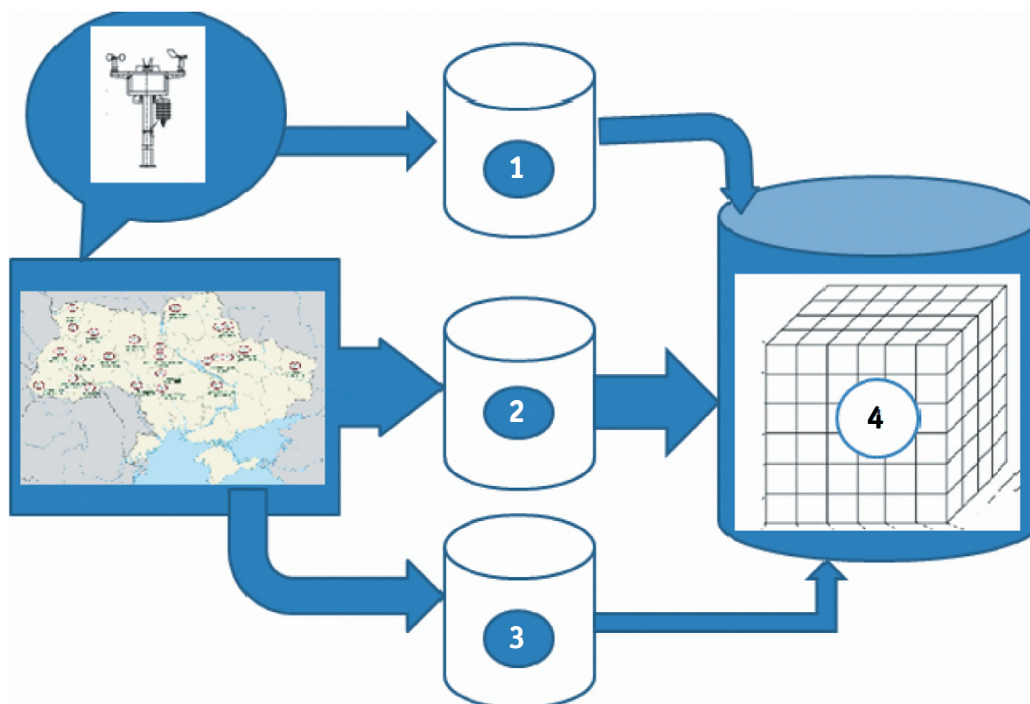


Рис. 1. Схема організації надходження даних у сховище даних кваліфікаційної експертизи сортів рослин (1 – програмно-апаратний комплекс Meteotrek; 2 – електронні польові журнали; 3 – оперативна БД АІС УІЕСР; 4 – сховище даних)

Дані до сховища даних надходять із трьох джерел: оперативної бази даних (БД) АІС УІЕСР, програмно-апаратного комплексу Meteotrek та електронних польових журналів, які ведуть у пунктах дослідження фахівці філій Українського інституту експертизи сортів рослин.

Результатні дані кваліфікаційної експертизи сортів рослин зберігаються в оперативній БД АІС УІЕСР, які потім завантажують у сховище даних з використанням стандартних механізмів реплікації SQL Server. Перед завантаженням до сховища даних відбувається інтеграція та агрегація даних з оперативної бази даних. Процес інтеграції та агрегації даних означає, що вихідні дані, які отримують з оперативної БД УІЕСР або імпортують з електронних таблиць, перевіряють, очищують, приводячи до єдиного виду, та виконують арифметичні й логічні операції (вираховують середні, мінімальні, максимальні, сумарні та інші статистичні показники) і лише потім завантажують у сховище показники. Інтеграція відбувається за сортами рослин, напрямками їх використання, ботанічними таксонами, природно-кліматичними зонами, роками проведення експертизи тощо. Ці дані є інваріантними у часі та організовані з мінімальною збитковістю інформації.

Також джерелом даних є електронні польові журнали, які ведуть фахівці філій УІЕСР, які безпосередньо працюють на дослідних полях.

Важливим джерелом інформації є програмно-апаратний комплекс Meteotrek – ав-

тономна метеостанція, яка ідентифікує та передає такі показники: температура повітря, атмосферний тиск, швидкість вітру, напрямок вітру, точка роси, опади відносні, відносна вологість повітря, температура ґрунту, вологість ґрунту. Програмне забезпечення дає змогу накопичувати та узагальнювати результати метеоспостережень, як у розрізі пунктів досліджень, так і ґрунтово-кліматичної зони проведення досліджень, розраховувати важливі показники (сума активних та ефективних температур, гідротермічний коефіцієнт тощо). Додатково метеостанції на пунктах, що проводять дослідження з кваліфікаційної експертизи сортів кукурудзи звичайної, укомплектовані датчиком сонячної радіації (піранометром).

Дані з Meteotrek перед завантаженням у сховища даних фільтруються, інтегруються та агрегуються (наприклад, визначають середньодобову, мінімальну та максимальну добову температуру повітря тощо).

Багатовимірною схемою сховища даних пристосована до потреб обох видів кваліфікаційних експертиз (на ВОС та ПСП) та побудована за схемою «сніжинка». Ця схема базується на математичних принципах, які підтримують нормалізовані моделі даних (рис. 2).

Прив'язка до часу даних у сховищі зумовлює зв'язок між фенологічними фазами розвитку кожного сорту рослин, який проходить кваліфікаційну експертизу, та метеорологічними умовами у відповідні для кожного сорту фенологічні фази розвитку. Такі

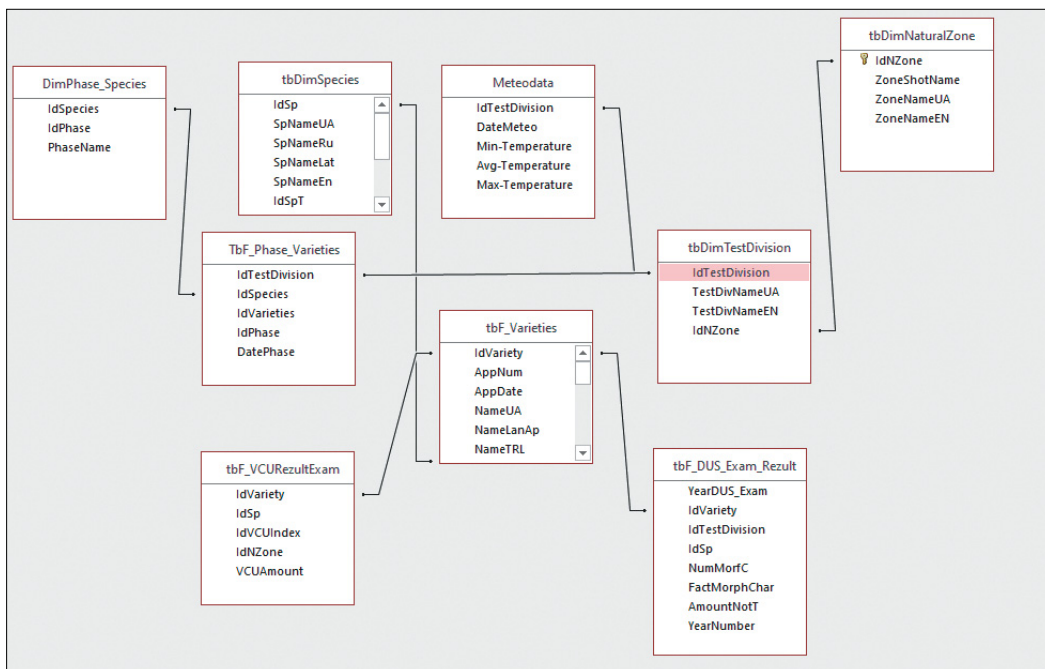


Рис. 2. Даталогічна схема сховища даних кваліфікаційної експертизи сортів рослин

дані отримано з трьох різноформатних джерел інформації й накопичено в сховищі у виді «історичних шарів», що дає змогу аналізувати тенденції в сортовивченні.

Отже, сховище виконує завдання щодо збирання інформації з баз даних, що відображають окремі бізнес-процеси УІЕСР та надходять з автоматизованих робочих місць, оперативних інформаційних систем і інших джерел інформації, зокрема й із глобальних комп'ютерних мереж (як приклад – дані щодо природно-кліматичних умов). Збір даних із різних джерел пов'язаний з тим, що інформація в них формується в різних форматах, має різноманітну структуру. Це вимагає напрацювання інформаційних технологічних процедур екстракції, очищення, фільтрації, агрегації та інтеграції даних, що будуть предметом вивчення в наступних дослідженнях.

Просторова схема сховища даних містить ту ж інформацію, що й нормалізована, але інформація структурована по-іншому, щоб полегшити її використання й виконання запитів. Ця схема включає як атомарні дані, так і узагальнювальну інформацію (агрегати у зв'язаних таблицях або багатомірних кубах) відповідно до вимог продуктивності або просторового розподілу даних. Запити в процесі виконання звертаються до все нижчого рівня деталізації без додаткового перепрограмування з боку користувачів або розробників застосунку.

До складу об'єктів сховища даних уходять таблиці вимірів і таблиці фактів. Зауважимо, що кожен вимір є сукупністю логічно пов'язаних атрибутів і розглядається як вісь для моделювання даних. У межах кожного виміру ці сутності формують рівні, на яких можна задавати різні запити для систем підтримки прийняття рішень (табл. 1).

Таблиця 1

Характеристика об'єктів сховища даних

№ з/п	Назва таблиці	Ідентифікатор таблиці	Тип таблиці	Джерело надходження даних
1	Довідник природно-кліматичних зон	tbDimNaturalZone	виміри	Оперативна БД УІЕСР
2	Довідник пунктів дослідження УІЕСР	tb_DimTestDivision	виміри	Оперативна БД УІЕСР
3	Довідник фенологічних фаз розвитку рослин	DimPhase_Species	виміри	Оперативна БД УІЕСР
4	Довідник ботанічних таксонів	tbDimSpecies	виміри	Оперативна БД УІЕСР
5	Метеорологічні дані	Meteodata	факти	БД програмно-апаратного комплексу Meteotrek
6	Фенологічні фази розвитку сортів рослин		факти	Електронні польові журнали філій УІЕСР
7	Опис сортів рослин, що проходили кваліфікаційну експертизу	F_Varieties	факти	Оперативна БД УІЕСР
8	Фактичні значення проявів морфологічних ознак за результатами аналітичного оброблення даних ВОС експертизи	tbF_DUSResultExam	факти	Результат процедур аналізу даних
9	Фактичні значення польових та лабораторних випробувань ПСП експертизи	tbF_VCUResultExam	факти	Результат процедур аналізу даних

До складу таблиць вимірів входять: Довідник фенологічних фаз розвитку рослин, Довідник ботанічних таксонів, Довідник природно-кліматичних зон, Довідник пунктів дослідження УІЕСР.

Таблиця tbDimNaturalZone містить такі атрибути: код природно-кліматичної зони (IdNZone), коротка назва природно-кліматичної зони (ZoneShotName), назва природно-кліматичної зони українською (ZoneNameUA), назва природно-кліматичної зони англійською (ZoneNameEN), а от таблиця tb_DimTestDivision містить атрибути: пункт дослідження (IdTestDivision), назва пункту дослідження українською (TestDivNameUA), назва пункту досліджен-

ня англійською (TestDivNameEn), код природно-кліматичної зони (IdNZone), таблиця DimPhase_Species містить: код ботанічного таксону (IdSpecies), код фази фенологічного розвитку рослин (IdPhase), назва фази фенологічного розвитку рослин (PhaseName).

Таблиця tbDimSpecies містить такі атрибути: код типу ботанічних таксонів (IdSp), назва типу ботанічного таксону українською (SpNameUA), назва типу ботанічного таксону латинською (SpNameLat), назва типу ботанічного таксону англійською (SpTNameEn), синонім назви ботанічного таксону (SpNameSininim), ознака на проведення випробувань на ВОС (ExpDUS), ознака на проведення випробувань на ПСП (ExpVCU код

типу ботанічних таксонів (IdSpT), кількість років (періодичність) оновлення зразка у сховищі, одиниця виміру для кількості оновлення зразка.

Запити DSS можуть отримувати доступ до фактів із багатьох різних вимірів та рівнів. Рівні в кожному вимірі полегшують пошук фактів з різних вимірів. Склад даних повинен мати узагальнений збір атрибутів даних, що робить пошук інформації ефективнішим.

До таблиць фактів належать: Метеорологічні дані; Фенологічні фази розвитку сортів рослин; Опис сортів рослин, що проходили кваліфікаційну експертизу; Фактичні значення проявів морфологічних ознак за результатами аналітичної обробки даних ВОС експертизи; «Фактичні значення польових та лабораторних випробувань ПСП експертизи».

Таблиця Meteorodata, що формується на підставі зовнішніх джерел, а саме програмно-апаратного комплексу Meteotrek, має такі атрибути: код пункту дослідження (IdTestDivision), дата (DateMeteo), середньодобова температура повітря (Avg-Temperature), максимальна добова температура повітря (Max-Temperature), мінімальна добова температура повітря (Min-Temperature), середньодобова температура ґрунту (Avg-Soil_Temper), максимальна добова температура ґрунту (Max-Soil_Temper), мінімальна добова температура ґрунту (Min-Soil_Temper), а також показники: атмосферний тиск, швидкість вітру, напрямок вітру, точка роси, опади відносні, відносна вологість повітря, вологість.

Таблиця F_Varieties містить атрибути: код сорту (IdVariety), номер заявки (AppNum), дата заявки (AppDate), назва сорту українською (NameUA), назва сорту на мові заявника (NameLanAp), назва сорту транслітерацією (NameTRL), код ботанічного таксону (IdSp), кодова формула сорту (CodFormula), відмінність (Distinctness), однорідність (Uniformity), стабільність (Stability), наявність новизни (Novelty), код причини визнання сорту таким, що не має новини (IdReason), текст загального висновку ВОС (GeneralText), додатковий опис сорту (AdditionalInf), номер експертного висновку (NumEV), дата експертного висновку (DateEV), номер методики, за якою проводились випробування (NumMetod), код тексту для експертного висновку (IdResom), номер блоку (IdBlock), рік коли сорт був включений до Державного реєстру сортів рослин, придатних для поширення в Україні (StateRegYear).

Таблиця tbF_DUSRezultExam містить атрибути: код сорту (IdVariety), код типу

морфологічної характеристики (NumMorfC), код ботанічного таксону (IdSp), узагальнене значення прояву морфологічної ознаки (GenMorphChar), примітка щодо прояву морфологічної ознаки (Comments). Таблиця tbF_VCURRezultExam призначена для збереження результатів статистичного оброблення результатів польових і лабораторних випробувань та містить такі атрибути: код сорту (IdVariety), код культури (IdSp), номер показника (IdVCUIndex), код природно-кліматичної зони (IdNZone), значення показника а (VCUAmount).

Висновки

Реалізовано багатовимірну модель даних, що поєднує результатні дані кваліфікаційної експертизи та погодні умови, за якими проводиться дослідження. Джерелами для завантаження даних є: оперативна БД АІС УІЕСР, програмно-апаратний комплекс Meteotrek та електронні польові журнали з пунктів дослідження філій Українського інституту експертизи сортів рослин.

Сховище даних побудовано за схемою «сніжинка» й містить таблиці вимірів: Довідник фенологічних фаз розвитку рослин, Довідник ботанічних таксонів, Довідник природно-кліматичних зон, Довідник пунктів дослідження УІЕСР та таблиці фактів: Метеорологічні дані; Опис сортів рослин, що проходили кваліфікаційну експертизу; Фактичні значення проявів морфологічних ознак за результатами аналітичної обробки даних ВОС експертизи; Фактичні значення польових та лабораторних випробувань ПСП експертизи. Зв'язок між даними кваліфікаційної експертизи та даними метеорологічних умов забезпечує таблиця Фенологічні фази розвитку сортів рослин.

Використана література

1. Anderson P. K., Cunningham A. A., Patel N. G. et al. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends Ecol. Evol.* 2004. Vol. 19, Iss. 10. P. 535–544. doi: 10.1016/j.tree.2004.07.021
2. Audsley E., Pearn K. R., Simota C. et al. What can scenario modelling tell us about future European scale agricultural use, and what not? *Environ. Sci. Pol.* 2006. Vol. 9, Iss. 2. P. 148–168. doi: 10.1016/j.envsci.2005.11.008
3. Baker R. H. A., Sansford C. E., Jarvis C. H. et al. The role of climatic mapping in predicting the potential distribution of non-indigenous pests under current and future climates. *Agric. Ecosyst. Environ.* 2000. Vol. 82. P. 57–71. doi: 10.1016/S0167-8809(00)00216-4
4. Bale J. S., Masters G. J., Hodkinson I. D. et al. Herbivory in global climate change research: direct effects of rising temperature on insect herbivores. *Global Change Biol.* 2002. Vol. 8, Iss. 1. P. 1–16. doi: 10.1046/j.1365-2486.2002.00451.x
5. Chloupek O., Hrstkova P., Schweigert P. Yield and its stability, crop diversity, adaptability and response to climate change, weather and fertilisation over 75 years in the Czech Republic in comparison to some European countries. *Field Crops Res.*

2004. Vol. 85, Iss. 2–3. P. 167–190. doi: 10.1016/S0378-4290(03)00162-X
6. Leschuk N. V., Orlenko N. S., Khareba O. V., Dydiv O. J. The use of grouping morphological characteristics of *Lettuce varieties* L. var. *capitata* for the difference test in Ukraine. *Int. J. Bot. Stud.* 2020. Vol. 5, Iss. 6. P. 516–522.
 7. Лещук Н. В., Мажуга К. М., Орленко Н. С., Стариченко Є. М., Шкапенко Є. А. Порівняльний аналіз статистичних програмних продуктів для кваліфікаційної експертизи сортів рослин на придатність до поширення. *Plant Var. Stud. Prot.* 2017. Т. 13, № 4. С. 429–435. doi: 10.21498/2518-1017.13.4.2017.117757
 8. Inmon W. H. *Building the Data Warehouse*. 3rd ed. New York, NY: John Wiley & Sons, 2002. 428 p.
 9. Kimball R., Ross M. *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*. New York, NY: John Wiley & Sons, 2002. 416 p.
 10. Шаховська Н. Б., Пасичник В. В. *Сховища та простори даних*. Львів: Львівська політехніка, 2009. 244 с.
 11. Cooper B. L., Watson H. J., Wixom B. H., Goodhue D. L. Data Warehousing supports corporate strategy at First American Corporation. *MIS Quarterly*. 2000. Vol. 24, Iss. 4. P. 547–567. doi: 10.2307/3250947
 12. Туманов В. Е. Проектирование хранилищ данных для приложений систем деловой осведомленности (Business Intelligence Systems). 2-е изд. Москва: Нац. Открытый Ун-т «ИНТУИТ», 2016. 958 с.
 13. Abdullah A., Brobst S., Pervaiz I. et al. Agri Data Mining/Warehousing: Innovative tools for analysis of integrated agricultural and meteorological data: to appear in proceedings of International Workshop on Frontiers of Information Technology, COMSATS (CIIT), Islamabad, Pakistan, 2003. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.9159&rep=rep1&type=pdf>
 14. Jindal R., Taneja S. Comparative study of data warehouse design approaches: A survey. *Int. J. Database Manag. Syst.* 2011. Vol. 4, No. 1. P. 205–210. doi: 10.5121/ijdms.2012.4104
 15. Agrawal R., Gupta A., Sarawagi S. Modelling Multidimensional Databases. *Proc. 13th Int. Conf. on Data Engineering*. 1997. P. 232–243. doi: 10.1109/ICDE.1997.581777
 16. Sharma L., Mehta N. Data Mining Techniques: A Tool for Knowledge Management System in Agriculture. *Int. J. Sci. Technol. Res.* 2012. Vol. 1, Iss. 5. P. 67–73.
 17. Gupta A. K., Mazumdar B. D. Multidimensional schema for agricultural. *Int. J. Res. Eng. Technol.* 2013. Vol. 2, Iss. 3. P. 245–253. doi: 10.15623/IJRET.2013.0203006
 18. Yost M., Nealon J. Using a dimensional data warehouse to standardize survey and census metadata. National Agricultural Statistics Service, U.S. Department of Agriculture, 1999. URL: https://nces.ed.gov/FCSM/pdf/II_B_Yost_FCSM1999.pdf
 19. Sharma S. D., Singh R., Rai A. Integrated National Agricultural Resources Information System (INARIS). New Delhi: Indian Agricultural Statistics Research Institute, 2000. URL: <https://www.geospatialworld.net/article/integrated-national-agricultural-resources-information-system-inaris/>
 20. Abdullah A., Brobst S., Pervaiz I. et al. Agri Data Mining/Warehousing: Innovative tools for analysis of integrated agricultural and meteorological data: to appear in proceedings of International Workshop on Frontiers of Information Technology, COMSATS (CIIT), Islamabad, Pakistan. 2003. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.9159&rep=rep1&type=pdf>
 21. Nilakanta S., Scheibe K., Rai A. Dimensional issues in agricultural data warehouse designs. *Comput. Electron. Agric.* 2008. Vol. 60, No. 2. P. 263–278. doi: 10.1016/j.compag.2007.09.009
 - technology drivers. *Trends Ecol. Evol.*, 19(10), 535–544. doi: 10.1016/j.tree.2004.07.021
 2. Audsley, E., Pearn, K. R., Simota, C., Cojocar, G., Koutsidou, E., Rounsevell, M. D. A., Trnka, M., & Alexandrov, V. (2006). What can scenario modelling tell us about future European scale agricultural and use, and what not? *Environ. Sci. Pol.*, 9(2), 148–168. doi: 10.1016/j.envsci.2005.11.008
 3. Baker, R. H. A., Sansford, C. E., Jarvis, C. H., Cannon, R. J., MacLeod, A., & Walters, K. F. A. (2000). The role of climatic mapping in predicting the potential distribution of non-indigenous-pests under current and future climates. *Agric. Ecosyst. Environ.*, 82(1–3), 57–71. doi: 10.1016/S0167-8809(00)00216-4
 4. Bale, J. S., Masters, G. J., Hodkinson, I. D., Awmack, C., Bezemer, M. T., Brown, V. K., ... Whittaker, J. B. (2002). Herbivory in global climate change research: direct effects of rising temperature on insect herbivores. *Global Change Biol.*, 8(1), 1–16. doi: 10.1046/j.1365-2486.2002.00451.x
 5. Chloupek, O., Hrstkova, P., & Schweigert, P. (2004). Yield and its stability, crop diversity, adaptability and response to climate change, weather and fertilisation over 75 years in the Czech Republic in comparison to some European countries. *Field Crops Res.*, 85(2–3), 167–190. doi: 10.1016/S0378-4290(03)00162-X
 6. Leschuk, N. V., Orlenko, N. S., Khareba, O. V., & Dydiv, O. J. (2020). The use of grouping morphological characteristics of *Lettuce varieties* L. var. *capitata* for the difference test in Ukraine. *Int. J. Bot. Stud.*, 5(6), 516–522.
 7. Leschuk, N. V., Mazhuha, K. M., Orlenko, N. S., Starychenko, Ye. M., & Shkapenko, Y. A. (2017). Comparative analysis of statistical software products for the qualification examination of plant varieties for suitability for distribution. *Plant Var. Stud. Prot.*, 13(4), 429–435. doi: 10.21498/2518-1017.13.4.2017.117757 [in Ukrainian]
 8. Inmon, W. H. (2002). *Building the Data Warehouse* (3rd ed.). New York, NY: John Wiley & Sons.
 9. Kimball, R., & Ross, M. (2002). *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling*. New York, NY: John Wiley & Sons.
 10. Shakhovskaya, N. B., & Pasichnyk, V. V. (2009). *Skhovyshcha ta prostory danykh* [Data warehouses and spaces]. Lviv: Lvivska politekhnika. [in Ukrainian]
 11. Cooper, B. L., Watson, H. J., Wixom, B. H., & Goodhue, D. L. (2000). Data Warehousing supports corporate strategy at First American Corporation. *MIS Quarterly*, 24(4), 547–567. doi: 10.2307/3250947
 12. Tumanov, V. E. (2016). *Proektirovanie khranilishch dannykh dlya prilozheniy sistem delovoy osvedomlennosti (Business Intelligence Systems)* [Designing Data Warehouses for Business Intelligence Systems Applications (Business Intelligence Systems)] (2nd ed.). Moscow: Natsional'nyy Otkrytyy Universitet "INTUIT" [in Russian]
 13. Abdullah, A., Brobst, S., Pervaiz, I., Umer, M., & Nisar, A. (2003). *Agri Data Mining/Warehousing: Innovative tools for analysis of integrated agricultural and meteorological data: to appear in proceedings of International Workshop on Frontiers of Information Technology*, COMSATS (CIIT), Islamabad, Pakistan. Retrieved from <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.9159&rep=rep1&type=pdf>
 14. Jindal, R., & Taneja, S. (2011). Comparative study of data warehouse design approaches: A survey. *Int. J. Database Manag. Syst.*, 4(1), 205–210. doi: 10.5121/ijdms.2012.4104
 15. Agrawal, R. R., Gupta, A., & Sarawagi, S. (1997). Modelling Multidimensional Databases. *Proc. 13th Int. Conf. on Data Engineering* (pp. 232–243). doi: 10.1109/ICDE.1997.581777
 16. Sharma, L., & Mehta, N. (2012). Data Mining Techniques: A Tool for Knowledge Management System in Agriculture. *Int. J. Sci. Technol. Res.*, 1(5), 67–73.
 17. Gupta, A. K., Mazumdar, B. D. (2013). Multidimensional schema for agricultural. *Int. J. Res. Eng. Technol.*, 2(3), 245–253. doi: 10.15623/IJRET.2013.0203006

References

1. Anderson, P. K., Cunningham, A. A., Patel, N. G., Morales, F. J., Epstein, P. R., & Daszak, P. (2004). Emerging infectious diseases of plants: pathogen pollution, climate change and agro-

18. Yost, M., & Nealon, J. (1999). Using a dimensional data warehouse to standardize survey and census metadata. National Agricultural Statistics Service, U.S. Department of Agriculture. Retrieved from https://nces.ed.gov/FCSM/pdf/II_B_Yost_FCSM1999.pdf
19. Sharma, S. D., Singh, R., & Rai A. (2000). *Integrated National Agricultural Resources Information System (INARIS)*. New Delhi: Indian Agricultural Statistics Research Institute. Retrieved from <https://www.geospatialworld.net/article/integrated-national-agricultural-resources-information-system-inaris/>
20. Abdullah, A., Brobst, S., Pervaiz, I., Umer, M., & Nisar, A. (2003). *Agri Data Mining/Warehousing: Innovative tools for analysis of integrated agricultural and meteorological data: to appear in proceedings of International Workshop on Frontiers of Information Technology*, COMSATS (CIIT). Islamabad, Pakistan. Retrieved from <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.9159&rep=rep1&type=pdf>
21. Nilakanta, S., Scheibe, K., & Rai, A. (2008). Dimensional issues in agricultural data warehouse designs. *Comput. Electron. Agric.*, 60(2), 263–278. doi: 10.1016/j.compag.2007.09.009

UDC 631.559.2:004.652

Melnyk, S. I., Leschuk, N. V., Orlenko, N. S.*, Starychenko, E. M., Mazhuha, K. M., & Shkapenko, Y. A. (2021). Features of recording of meteorological conditions in the data warehouse of qualification examination of plant varieties. *Plant Varieties Studying and Protection*, 17(3), 254–261. <https://doi.org/10.21498/2518-1017.17.3.2021.242980>

*Ukrainian Institute for Plant Variety Examination, 15 Henerala Rodymtseva St., Kyiv, 03041, Ukraine, *e-mail: n.s.orlenko@gmail.com*

Purpose. To develop a multidimensional model of the data storage for the qualification examination of plant varieties for fixing meteorological conditions in conjunction with the phenological stages of development of varieties that undergo DUS and SVD expertise. **Methods.** To conduct research with the establishment of the main structural elements of a multidimensional data warehouse, methods of induction, deduction, analysis and synthesis were used. In the design process of the storage facility, W. H. Inmon's concept was applied, adapted for the agricultural and agricultural business. **Results.** The stages of qualification examination of plant varieties were analyzed and methodological approaches to the creation of a multidimensional data warehouse model were considered. The features of the use of data storages for storing the results of qualification examination of plant varieties for distinctness, uniformity and stability (DUS) and suitability of a variety for dissemination in Ukraine (SVD) were highlighted. Particular attention was paid to the implementation of the interconnection between

the results of the qualification examination of plant varieties with the data of meteorological observations at various phenological stages of plant growth and development, according to the records in the electronic field journal. The logical data model of the data warehouse was designed and implemented in the MS SQL Server environment. **Conclusions.** Sources of data entry into data warehouses were determined and a multidimensional data warehouse model was implemented according to the "snowflake" scheme. The diagram of the data warehouse was presented, which provided a link between the meteorological conditions of the field experiments and the initial data of the qualification examination, and had four tables of measurements. For each dimension table and fact table, an attribute composition of the data was defined. The data warehouse was practically used to analyze the influence of weather conditions on the indicators of DUS and SVD examinations.

Keywords: meteorological data; phenological stages; dimensions; facts; attributes; growing season; interphase period.

*Надійшла / Received 08.06.2021
Погоджено до друку / Accepted 22.07.2021*