

2. Каталог рослин дендрологічного парку "Олександрія"/ Колект. авт. – Біла Церква, 1997. – 118 с.
3. Международный кодекс ботанической литературы. – Л.: Наука, 1984. – 32 с.
4. Тахтаджян А.Л. Система Магнолиефитов. – Л.: Наука, 1987. – 439 с.

УДК 630*5(477.83).004.657

Доц. М.М. Петренко, канд. с.-г. наук;
аспір. Г.С. Домашовець¹ – Національний аграрний університет, м. Київ

ВИКОРИСТАННЯ ЗАСОБІВ OLAP ДЛЯ АНАЛІЗУ ПРОДУКТИВНОСТІ ЛІСОВОГО ФОНДУ ЛЬВІВЩИНИ

Наведено методи використання засобів OLAP для аналізу продуктивності головних лісотвірних порід Львівської області у контексті лісорослинного районування.

Ключові слова: база даних, OLAP, проектування баз даних, лісовий фонд, показники лісового фонду, аналіз даних.

*Assoc. prof. M.M. Petrenko; post-graduate G.S. Domashovets –
National agrarian university, Kyiv*

The use of OLAP tools for forest fund productivity analysis Lviv region

The methods of using of OLAP tools for the analysis of main forest forming species in Lviv region by the context of forest regions.

Keywords: data base, OLAP, data base development, forest fund data, data analysis.

Сучасне суспільство є суспільством інформаційним. Усі галузі його діяльності постійно розвивають і вдосконалюють методи нагромадження і використання даних. Корисна та цінна інформація здатна підвищити ефективність діяльності проекту чи організації за рахунок покращення якості управлінських рішень.

Для отримання цінної інформації тепер використовуються комп'ютерні технології, які поряд із збільшенням точності та об'єктивності при роботі з великими обсягами інформації дають змогу значно скоротити витрати часу і праці на її отримання. Найбільш поширеною і розвинутою технологією, яка призначена для роботи з великими масивами даних, є бази даних (БД). БД – це структурована сукупність логічно взаємопов'язаних даних, які характеризують певну предметну область [4, 6].

Поряд з іншими засобами аналізу даних під час бурхливого розвитку комп'ютерних засобів з'явилась технологія багатомірного аналізу даних OLAP (Online Analytical Processing – оперативна аналітична обробка). Це спосіб управління і представлення даних у простий і зрозумілий для кінцевого користувача спосіб [4].

Призначення систем класу OLAP – забезпечити користувача гнучким інтуїтивно зрозумілим і простим доступом до даних. Наявність такого доступу дає змогу відмовитися від використання наперед визначених звітів, робить користувачів самодостатніми, незалежними від адміністраторів БД і програмістів. В основі концепції OLAP дані представлені у вигляді багатовимірного куба, причому користувач має змогу швидко згорнути або розгорнути дані за

¹ Наук. керівник: проф. П.І. Лакида, д-р с.-г. наук – Національний аграрний університет, м. Київ

будь-яким виміром. Ця технологія не замінює, а доповнює традиційні реляційні бази даних з первинною інформацією.

Для побудови систем OLAP використовуються спеціалізовані багатовимірні БД або надбудови над звичайними реляційними БД. До останнього часу OLAP-технології асоціювалися з великими проектами зі зберігання масивів даних і складними додатками для їх аналізу. Складний і дорогий OLAP-інструментарій був доступний лише дуже великим організаціям. І все-таки, за останній час ситуація на ринку різко змінилася. Відбулося це завдяки тому, що було знайдено компромісне рішення: укомплектувати повноцінним OLAP-сервером добре зарекомендовані недорогі програмні продукти, наприклад MS SQL-сервер, починаючи з версії 7 і вище, який у всьому світі активно використовується для побудови сховищ даних. Компанія Microsoft здійснює низку важливих заходів для забезпечення найкращої підтримки сховищ даних і побудови інформаційних систем. Внаслідок вказаної зміни ситуації сучасні OLAP-системи аналізу даних стали дійсно доступні малому та середньому бізнесу [1, 5].

Метою цієї роботи є всебічний аналіз продуктивності лісового фонду Львівської області. Для реалізації цієї задачі вирішено застосовувати методи багатомірного аналізу даних на основі використання OLAP-засобів.

Варто відзначити, що Львівщина характеризується великою різноманітністю природних умов і багатством природних ресурсів, що зумовлено її географічним положенням, геологічною будовою і характером поверхні. На території області вирізняють п'ять лісорослинних областей: Українські Карпати і прилегла до них область Передкарпаття – на півдні, Опілля та Розточчя – в центральній частині та Мале Полісся – на півночі [2].

Лісовий фонд Львівщини перебуває у підпорядкуванні багатьох міністерств і відомств, що значно ускладнює можливість узагальнення і оптимізації даних про ліси області. Основна частина лісового фонду перебуває у користуванні Державного комітету лісового господарства України (68,6 %), Міністерства аграрної політики України (21,6 %), Міністерства оборони України (7,7 %). Решту лісів (2,1 %) розподілені серед інших 14 лісокористувачів. За даними головного управління статистики, у Львівській області станом на 1 січня 2006 р. зареєстровано 85 підприємств лісового господарства та пов'язаних з ним послуг. Таке різноманіття підпорядкування лісового фонду спричиняє значні труднощі для збору та аналізу статистичної інформації [2, 3].

Для виконання цієї роботи було використано дані державного лісового кадастру 1954, 1966, 1978, 1989, 2003 та 2006 рр., надані Виробничим об'єднанням "Укрдержліспроект", Державною лісовпорядною експедицією "Львівдержліспроект", державними підприємствами Львівського обласного управління лісового господарства, Сколівським, Магерівським та Старицьким військовими лісгоспами, спеціалізованим лісогосподарським підприємством ДП "Галсільліс". Процес збору даних був ускладнений фрагментарністю даних, основна частина яких зберігалася на паперових носіях у різних джерелах інформації. При цьому варто зазначити, що значна частина даних про стан лісового фонду з часом була втрачена, оскільки не була зосереджена в єдиних джерелах.

Вказані дані було вирішено записати і зберігати в реляційних таблицях СУБД Ms Access. Ця програма має вбудовані засоби OLAP аналізу. Процес розроблення OLAP моделі охоплює такі етапи:

- проектування OLAP моделі даних;
- заповнення даних;
- використання моделі.

Важливими питаннями при проектуванні моделі OLAP є рівень агрегації даних та визначення якісних ознак. Рівень агрегації визначає межу, нижче якої деталізувати дані неможливо. У цій роботі рівень агрегації даних було вибрано як рівень найменшої групи, яку формують ознаки "деревна порода" – "група віку" – "бонітет". Кожен показник, який вибирався з Державного лісового кадастру в OLAP-модель, має бути визначений за якісною роллю, яку він буде виконувати в моделі. Показник може бути ознакою кількісною або груповою, причому остання формує розмірність куба даних.

Як кількісні ознаки у моделі було обрано площу і запас. Для агрегації цих показників використовували операцію додавання. Груповими ознаками в моделі були призначені всі інші. Групова ознака слугує для формування груп, які характеризуються ознакою кількісною. Наприклад, показник "назва підприємства" може використовуватись як групувальна ознака для показника "площа". У результаті ми отримаємо перелік підприємств, для яких буде знайдена сума кількісної ознаки (операція агрегації) – площі. Математичні операції над груповими показниками недопустимі, їх основне призначення деталізація або агрегація кількісних характеристик.

Групувальні ознаки в OLAP-моделі аналізу продуктивності лісів знаходяться у певній ієрархії. Ієрархія визначає логічний зв'язок між ознаками. Наприклад, показник "підприємство" є підпорядкованим показнику "область". Ми можемо провести деталізацію площ підприємств у межах області. Взаємозв'язок показників можна побачити на рис. 1.

Процедура наповнення передбачає перенесення даних з матеріалів Державного лісового кадастру в реляційні таблиці. Цей масив у подальшому буде джерелом інформації для засобів OLAP-аналізу. На даному етапі важливою є акуратність, щоб не допускати випадкових помилок у процесі перенесення. У кінці наповнення таблиць бажано застосовувати метод логічної перевірки помилок. Наприклад, суму запасів за підприємствами розділити на суму площ – має вийти середній запас на 1 га у підприємстві.

Особливості використання OLAP-моделі визначаються потребами і задачами користувача. Для отримання таблиці користувач вибирає показник (або їх групу) для аналізу і кількісну ознаку, яку він буде аналізувати. Користувач оперує лише лісогосподарськими поняттями. Немає потреби вивчення методів і суті роботи OLAP-моделі чи іншої теорії інформатики.

Наприклад, у процесі виконання аналізу продуктивності лісів знадобилось отримати таблицю із розподілом площ лісів різних деревних порід у різних лісорослинних областях. У формі аналізу даних OLAP зі списку показників вибираємо показник "лісорослинна область" – як групова ознака за рядками; "група порід" – як групова ознака за стовпчиками і деталізуємо її по-

казником "деревна порода". Як кількісну ознаку вибираємо зі списку показників "запас" і метод агрегації – "сума". У результаті отримаємо розподіл запасів головних лісотвірних порід (рис. 2).

Рис. 1. Форма заповнення реляційних таблиць даними Державного лісового кадастру

GrPorName	GrVik	Rik	DL Gname
Исключая: (Пусто)	(несколько элементов)	2003	Все
PorName	Бук	Береза	Вільха
Зона	Сумма "М (26)"	Сумма "М (26)"	Сумма "М (26)"
Карпати	13926,01499	261,0769392	609,7696808
Мале Полісся	2160,081444	685,5346013	2606,764363
Опілля	9019,29375	564,1894537	1139,87195
Передкарпаття	4461,080969	411,2660901	947,9942945
Розточчя	2738,773942	677,5436662	774,3281056
Общие итоги	32305,2451	2599,610751	6278,728393

Рис. 2. Розподіл запасів головних лісотвірних порід, тис. м³

Таблиця розподілу площ і запасів деревних порід у розрізі груп віку, яка виконується вручну з використанням паперових матеріалів і калькулятора, потребує близько 8 год. якщо ми зводимо ці дані для області. Отримання такої таблиці за допомогою OLAP-засобів, вбудованих у СУБД Access, потребує коло 1 хв. Однак, останній підхід вимагає витрат значних зусиль на проектування моделі, підготовку даних, їх перенесення в реляційні таблиці і перевірку. Тим не менше, такі зусилля себе виправдовують. Аналіз ефективності застосування різних інформаційних технологій наведено в табл. 1.

Таким чином, якщо є потреба в одній або декількох таблицях, використання комп'ютерних засобів є неефективним порівняно з ручним. Але, якщо необхідно отримати багато різних форм аналітичної інформації, комп'ютерна обробка продуктивніша більш як у 10 разів. Крім того, до переваг, які були помічені у процесі роботи, можна віднести:

- зменшення кількості помилок;
- простота формування таблиць;
- просте перенесення аналітичних таблиць до інших програм (текстові редактори, електронні таблиці);
- можливість використання цих даних для вирішення інших задач (наприклад, інформація була використана у процесі дослідження фітомаси лісів у Львівській області).

Табл. 1. Аналіз витрат часу за двома методами обробки даних

Вид робіт	Витрати часу, год.			
	ручний спосіб		комп'ютерна обробка	
	на весь масив	на 1 таблицю	на весь масив	на 1 таблицю
Підготовка матеріалів	–	–	1	
Розроблення OLAP-МОДЕЛІ	–	–	8	
Перенесення даних	–	–	16	
Перевірка даних	50	1	2	
Отримання аналітичної інформації	400	8	1	0,02
Представлення	50	1	5	0,1
Сума	500	10	33	0,12

Внаслідок розроблення бази даних і засобів OLAP-аналізу з продуктивності лісового фонду Львівської області був створений інформаційний ресурс на базу СУБД Ms Access із елементами OLAP-аналізу. Його можна рекомендувати для:

- використання під час досліджень лісового фонду різних лісокористувачів області;
- досліджень продуктивності лісів області;
- дослідження динаміки лісового фонду області;
- базу даних можна доповнювати даними інших областей, що скоротить витрати часу на розробку структури OLAP-моделі.

Література

1. Дженнингс, Роджер. Руководство разработчика баз данных на Visual Basic 6/ Пер. с англ. – К.-М.; СПб.: Изд. дом "Вильямс", 2001. – 976 с.
2. Домашовець Г.С., Василюшин Р.Д. Дослідження продуктивності лісів Львівщини в контексті лісорослинного районування// Наук. вісник НАУ: Зб. наук. праць. – К.: НАУ. – 2007, вип. 105. – С. 212-219.
3. Інструкція про порядок ведення державного лісового кадастру і первинного обліку лісів. – К.: Міністерство лісового господарства України, 1995. – 26 с.
4. Малыгина М.П. Базы данных: основы проектирования, использования. – СПб.: БХВ – Петербург, 2004. – 412 с.
5. Усольцев В.А. Формирования банков данных о фитомассе лесов. – Екатеринбург: УрО РАН, 1998. – 543 с.
6. Харитонов И.А., Михеева В.А. Microsoft Access 2000. – СПб.: БХВ-Петербург, 2002. – 1088 с.