

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

УДК 681.3:519.711:004.8

САРИЧЕВА Людмила Віссаріонівна

**МАТЕМАТИЧНЕ ЗАБЕЗПЕЧЕННЯ ТА КОМП'ЮТЕРНА
ТЕХНОЛОГІЯ ЕКОЛОГО-СОЦІАЛЬНО-ЕКОНОМІЧНОГО
МОНІТОРИНГУ РЕГІОНІВ**

**05.13.06 – Автоматизовані системи управління та прогресивні
інформаційні технології**

Автореферат
дисертації на здобуття наукового ступеня
доктора технічних наук

Київ
2004

Дисертацією є рукопис.

Робота виконана на кафедрі геоінформаційних систем Національного гірничого університету Міністерства освіти і науки України (м. Дніпропетровськ).

Науковий консультант: доктор технічних наук, професор, завідувач кафедри геоінформаційних систем Національного гірничого університету

БУСИГІН
Борис Сергійович

Офіційні опоненти:

доктор технічних наук, професор, академік НАН України, заслужений діяч науки України, радник дирекції Міжнародного науково-навчального центру інформаційних технологій і систем НАН України і Міністерства освіти і науки України

ІВАХНЕНКО
Олексій Григорович

доктор технічних наук, професор, завідувач кафедри комп'ютерних систем моніторингу Донецького національного технічного університету

АВСРІН
Геннадій Вікторович

доктор технічних наук, професор кафедри математичного забезпечення ЕОМ Дніпропетровського національного університету

ПРИСТАВКА
Олександр Пилипович

Провідна установа: Навчально-науковий комплекс "Інститут прикладного системного аналізу" НАН України і Міністерства освіти і науки України.

Захист відбудеться "12" грудня 2004 р. о 14 годині на засіданні спеціалізованої вченої ради Д 26.062.01 при Національному авіаційному університеті за адресою: 03058, м. Київ, просп. Космонавта Комарова, 1.

З дисертацією можна ознайомитись у бібліотеці Національного авіаційного університету за адресою: 03058, м. Київ, просп. Космонавта Комарова, 1.

Автореферат розісланий "___" _____ 2004 р.

Вчений секретар
спеціалізованої вченої ради Д 26.062.01



В.С. Єременко

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність проблеми. Однією з пріоритетних задач світової спільноти на ХХІ сторіччя Організація Об'єднаних Націй проголосила розв'язання проблеми сталого розвитку. Для виходу на траєкторію сталого розвитку необхідним є науково обгрунтований підхід до прогнозування розвитку регіонів (територіальних одиниць – областей, районів тощо). Тому виникла проблема кількісної оцінки рівня сталого розвитку та моніторингу екологічних, соціальних, економічних показників.

Далі під еколого-соціально-економічним (ЕСЕ) моніторингом як складової системи управління регіональним розвитком розумітимемо спеціально організоване цільове систематичне спостереження та короткотермінове прогнозування ходу ЕСЕ-процесів з метою їх аналізу, ідентифікації та вияву кола регульованих чинників для підготовки і прийняття рішень.

Математичне забезпечення ЕСЕ-моніторингу визначимо як комплекс математичних методів, алгоритмів і програм, що дають змогу автоматизувати процес розв'язання задач ЕСЕ-моніторингу на комп'ютері.

Комп'ютерна технологія ЕСЕ-моніторингу передбачає цілеспрямований процес використання математичного забезпечення в комп'ютерній системі для отримання ефективного розв'язку задач ЕСЕ-моніторингу.

Нинішній етап розвитку України ставить нові вимоги до ефективності прийняття оперативних і стратегічних рішень на всіх рівнях управління – від державного до муніципального, виконання яких неможливе без комп'ютерних систем ЕСЕ-моніторингу (ЕСЕМ). Актуальною проблемою при цьому стає видобування з «сирих» ЕСЕМ-даних раніше невідомих, нетривіальних знань та закономірностей, корисних для підтримки прийняття рішень.

При обробці соціально-економічних показників регіонів традиційно використовують загальнозживані статистичні методи, а за значеннями екологічних показників будують карти їх розподілу за територією. Такий підхід не відповідає сучасним вимогам щодо ефективності прийняття управлінських рішень. Оперативна обробка просторово-часової зашумленої ЕСЕ-інформації потребує залучення методів геоінформаційного та інтелектуального аналізу даних (ІАД), розробки відповідного математичного забезпечення, методології створення комп'ютерної системи і технології ЕСЕ-моніторингу.

Важливість проблеми підтверджено Законами України "Про Національну програму інформатизації" (№ 74/98 від 04.02.98), "Про охорону навколишнього природного середовища" (№ 1268-ХІІ від 25.06.91), "Про державне прогнозування та розроблення програм економічного і соціального розвитку України" (№ 1602-ІІІ від 23.03.2000), "Про забезпечення моніторингу соціально-економічної ситуації" (протокол №4 засідання Кабінету Міністрів України від 15.03.2000), Постановою Верховної Ради України "Про основні напрями державної політики України в галузі охорони довкілля, використання природних ресурсів та забезпечення екологічної безпеки" (№ 188/98-ВР від 05.03.98).

Отже, актуальність теми означає:

1) в науковому аспекті – гостру необхідність розв'язання проблеми розробки математичного забезпечення і комп'ютерної технології ЕСЕ-моніторингу регіонів;

2) у прикладному аспекті: а) існування нагальної потреби у розв'язанні задач комп'ютерної підтримки прийняття рішень з управління розвитком території регіонів; б) сприяння підвищенню кваліфікації кадрів – новітні знання, отримані в дисертаційній роботі, можуть увійти до навчальних програм підготовки фахівців зі спеціальностей 8.080407 Комп'ютерний еколого-економічний моніторинг і 8.080404 Інтелектуальні системи прийняття рішень освітнього напрямку 0804 Комп'ютерні науки.

Зв'язок роботи з науковими програмами, планами, темами. Тематика дисертації узгоджується із завданнями Національної програми інформатизації України і орієнтується на створення інформаційно-аналітичного забезпечення організацій, що регулюють регіональний розвиток. Мета роботи, її основні задачі відповідають планам науково-технічних робіт Міністерства освіти і науки України, Міністерства охорони навколишнього природного середовища України, Міністерства економіки та з питань європейської інтеграції України.

Дисертаційна робота виконана на кафедрі геоінформаційних систем Національного гірничого університету (НГУ) за індивідуальним планом підготовки докторанта і узагальнює отримані автором результати щодо науково-дослідних робіт за держбюджетними темами: "Моделювання розвитку екологічних процесів у регіонах України на основі актуалізованих баз даних з метою розробки та підтримки управлінських рішень" (№ держреєстрації 0104U007349), "Розробка наукових основ і принципів створення гірничої інформаційно-аналітичної системи" (№ держреєстрації 0101U001550), "Розробка наукових основ геоінформаційного забезпечення еколого-економічного моніторингу довкілля гірничопромислових регіонів" (№ держреєстрації 0103U001274), "Моделювання процесів соціально-економічного розвитку регіонів України та актуалізація динамічних рядів баз даних з метою здійснення оперативного аналізу" (№ держреєстрації 0103U006255).

Мета роботи – розробка математичного забезпечення та комп'ютерної технології ЕСЕ-моніторингу для цілей регулювання сталого розвитку регіонів.

Ідея роботи – видобування знань з різномірних ЕСЕМ-даних на основі системного підходу і сучасних інформаційних технологій для кількісної та якісної оцінки сталості й збалансованості регіонального розвитку.

Об'єкт досліджень – ЕСЕ-процеси на території регіонів, подані показниками ЕСЕ-моніторингу.

Предмет досліджень – математичне, геоінформаційне забезпечення і комп'ютерна технологія ЕСЕ-моніторингу регіонів.

Основні задачі досліджень

1. Математичне забезпечення ЕСЕ-моніторингу регіонів:

а) визначення задач, створення концепції ЕСЕ-моніторингу;

б) інтегральна ЕСЕ-оцінка регіонів для ЕСЕ-аудиту та районування території;

- в) обґрунтування і реалізація методів та технологій ІАД для інтерпретації та прогнозування ЕСЕ-показників регіонів;
- г) теоретико-графовий аналіз просторово розподілених ЕСЕМ-даних;
- д) розробка методів і алгоритмів виділення аномалій регіонального розвитку;
- е) аналіз та оптимізація мережі моніторингу;
- є) розробка математичної ЕСЕ-моделі з урахуванням інвестицій та аналіз сценаріїв розвитку регіону.

2. Геоінформаційне забезпечення ЕСЕ-моніторингу регіонів:

- а) обґрунтування й реалізація методів геоінформаційного аналізу та геоінформаційних технологій для задач ЕСЕ-моніторингу;
- б) геоіконічне моделювання ЕСЕ-показників з метою аналізу збалансованості регіонального розвитку;
- в) інтелектуальний аналіз просторових даних ЕСЕ-моніторингу в геоінформаційній системі (ГІС), генерування нових знань за допомогою ГІС-аналізу та геоіконічного моделювання;
- г) розробка методики створення електронних атласів сталого розвитку регіонів.

3. Створення комп'ютерної технології ЕСЕ-моніторингу регіонів:

- а) розробка інформаційної технології просторово-часового аналізу ЕСЕМ-даних;
- б) обґрунтування структури, логічної організації, принципів побудови, методів і засобів комп'ютерної системи ЕСЕ-моніторингу.

Методи досліджень. Для розв'язання сформульованих задач залучалися методи системного аналізу, математичного та геоіконічного моделювання, розпізнавання образів, теорії графів, геоінформаційного аналізу, математичної статистики і геостатистики.

Наукова новизна отриманих результатів.

1. Уперше запропоновано методологію геоіконічного моделювання в середовищі ГІС для ЕСЕ-моніторингу:

а) класифіковано геоіконічні моделі ЕСЕ-показників (рангові картограми; картодіаграми; аналітичні, комплексні, синтетичні карти; моделі регіональної генералізації; інтегральні регіональні моделі; динамічні моделі та анаморфози);

б) розроблено геоінформаційні технології побудови окремих типів моделей, що ґрунтуються на різних схемах класифікації, обчисленні інтегральної оцінки території, картуванні результатів ранжування, розпізнавання, кластеризації, прогнозування, виділення аномалій та відображають динаміку, незбалансованість регіонального розвитку.

2. Доведено, що геоінформаційні технології та геоіконічне моделювання дозволяють генерувати нові дані, що не враховуються офіційною статистикою (генералізація границь на основі рангових картограм, картодіаграми для аналізу просторового розподілу регресійних залишків моделей залежностей між ЕСЕ-показниками, синтез інформації на базі оверлейних ГІС-операцій та ін.), і за рахунок цього доповнити ЕСЕМ-дані.

3. Уперше з єдиних системних позицій викладено методолого-теоретичні основи створення комп'ютерної системи ЕСЕ-моніторингу:

- а) модель процесу постановки і розв'язання завдань ЕСЕ-моніторингу;
- б) мету, призначення, принципи створення, загальну характеристику, вимоги до системи в цілому та окремих підсистем, структурну схему і функції модулів основних блоків.

4. Розроблено методику інтегральної ЕСЕ-оцінки території регіонів для ЕСЕ-аудиту, завдяки якій можна кількісно оцінювати рівень сталого розвитку регіонів на підставі моніторингових даних трьох інформаційних блоків – екологічного, соціального та економічного, у кожному з яких використовують три типи показників: фактори забруднення, показники стану і керуючі впливи.

5. Запропоновано єдиний системний підхід щодо класифікації методів кластеризації та вибору ознак у розпізнаванні образів, в якому як кластеризація, так і вибір ознак є задачами дискретної оптимізації на вершинах ґратки розбивань, дозволяє використовувати існуючі алгоритми перебору вершин у виборі ознак для пошуку найкращої кластеризації.

6. Розроблене алгоритмічне забезпечення об'єктивної кластеризації (на принципах методу групового врахування аргументів та запропонованого в п.5 підходу), прогнозування (на основі структурної ідентифікації залежностей між ЕСЕ-показниками), виявлення аномалій регіонального розвитку, аналізу та оптимізації мережі моніторингу дозволяє підвищити ефективність (порівняно з відомими методами) інтегрованого аналізу і прогнозу ЕСЕ-показників, що доведено експериментальними дослідженнями на матеріалах моніторингу ЕСЕ-показників регіонів України за 1985-2000 рр.

7. Рекомендовано методику та відповідну їй комп'ютерну інформаційну технологію просторово-часового аналізу показників ЕСЕ-моніторингу, що включають етапи: ПС-аналіз і геоіконічне моделювання, інтелектуальний аналіз даних, генерування нових знань, аналіз і прийняття рішень, діалог "ОПР (особа, що приймає рішення) – ЕОМ".

Отримані результати спрямовані на інтегрований аналіз ЕСЕ-показників, підвищення ступеня видобування інформації з ЕСЕМ-даних, оперативності та достовірності регіональних прогнозів.

Вірогідність і обґрунтованість наукових положень, висновків і рекомендацій підтверджуються використанням фундаментальних положень загальної теорії систем, математичного моделювання, теорії розпізнавання образів, багатовимірного статистичного аналізу, теорії графів, прийнятими рівнями припущень при математичному описі явищ, значним обсягом експериментальних досліджень, позитивним досвідом упровадження низки наукових положень роботи при вирішенні народногосподарських завдань. При перевірці розроблених математичного забезпечення та комп'ютерної технології на реальних даних ЕСЕ-моніторингу регіонів України і зокрема Дніпропетровської області отримано добру узгодженість, тобто запропоновані моделі адекватно описують тенденції регіональних ЕСЕ-процесів.

Практична цінність і впровадження результатів дисертації.

1. Розроблене математичне і програмне забезпечення (ГІС "Охорона здоров'я") використано в управлінні охорони здоров'я Дніпропетровської міської ради для аналізу стану здоров'я та динаміки захворюваності населення Дніпропетровська.

2. Створеним "Електронним атласом Дніпропетровської області для цілей збалансованості регіонального розвитку" скористалося Дніпропетровське обласне управління статистики у звітній документації за 2001 р.

3. Математичне і програмне забезпечення задач структурної ідентифікації знайшло застосування в Дніпродіпрошахті (м. Дніпропетровськ).

4. Методичні рекомендації з обчислення інтегральної ЕСЕ-оцінки техногенно навантажених територій використані в Укрдїпруді (м. Харків).

5. Геоіконічні моделі динаміки еколого-соціально-економічних показників регіонів України застосовують у Державному науково-дослідному інституті інформатизації і моделювання економіки (м. Київ).

6. Методичними розробками з об'єктивної кластеризації та прогнозування ЕСЕ-показників користуються в департаменті регіональної політики Міністерства економіки та з питань європейської інтеграції України для аналізу збалансованості розвитку регіонів нашої держави.

7. Автор спирається на результати цих досліджень у процесі навчання студентів НГУ (спеціальності 8.080407 Комп'ютерний еколого-економічний моніторинг, 8.080404 Інтелектуальні системи прийняття рішень освітнього напрямку 0804 Комп'ютерні науки) в курсах "Комп'ютерна підтримка прийняття рішень", "Моделювання екологічних і соціально-економічних процесів".

Особистий внесок автора. Дисертація відображає результати досліджень, здійснені автором на кафедрах вищої математики (1993-1995) та геоінформаційних систем (1996-2004), під час навчання в докторантурі НГУ. Основні результати отримані безпосередньо дисертантом (2 монографії та більшість статей за темою дисертації – без співавторів). Структура комп'ютерної системи ЕСЕ-моніторингу розроблялась разом з Б.С.Бусигінім і С.Л.Нікулінім [30, 31]. Алгоритмічне і програмне забезпечення методів структурної ідентифікації і вибору ознак здійснено разом з О.П.Саричевим [16].

Апробація результатів досліджень. Основні положення дисертаційної роботи доповідалися на II Міжнародній конференції Укробраз'94 (Київ, 1994), II Міжнародному симпозіумі "Застосування математичних методів і комп'ютерних технологій при вирішенні задач геохімії й охорони навколишнього середовища" (Львів, 1994), I науково-практичній конференції "Sustainable Development: Environmental Pollution and Ecological Safety" (Дніпропетровськ, 1995), VI International Symposium on Application of Mathematical Methods and Computers in Mining, Geology and Metallurgy (Prague, 1997), Міжнародній конференції з проблем екології і ресурсозберігаючих технологій (Луцьк, 1998), I Міжнародній конференції "Проблеми і перспективи використання геоінформаційних технологій у гірництві" (Дніпропетровськ, 1999), Українському форумі з геоінформаційних технологій

ГІС-ФОРУМ–99 (Київ, 1999), Міжнародній конференції "Моделювання та оптимізація складних систем" (Київ, 2001), II Науково-практичній конференції "Проблеми і перспективи використання геоінформаційних технологій у гірництві" (Дніпропетровськ, 2000), V Всеукраїнській міжнародній конференції UkrOBRAZ-2000. Signal/Image Processing and Pattern Recognition (Київ, 2000), II Міжнародному науково-практичному семінарі "Практика і перспективи розвитку інституційного партнерства" (Донецьк, 2001), II Міжнародній науково-практичній конференції "Нестеренківські читання" (Дніпропетровськ, 2001), VI Міжнародній науково-технічній конференції "Контроль і управління в складних системах" (Вінниця, 2001), Міжнародній міждисциплінарній науково-практичній конференції "Сучасні проблеми науки та освіти" (Керч, 2001), VI Міжнародній конференції "Pattern Recognition and Information Processing" (Мінськ, 2001), III Міжнародній конференції "Проблеми геоінформатики при комплексному освоєнні надр" (Дніпропетровськ, 2001), Міжнародній конференції "Modeling of Developing Systems: Risk Assessment of Ecological, Technogenic and Sociogenic Accidents. 15 years after the Chernobyl Catastrophe MDS-XXV" (Liptovsky Mikulash, 2001), IGARSS 2001 – International Geoscience and Remote Sensing Symposium, V Міжнародній науково-технічній конференції "Геоінформаційні технології в управлінні територіальним розвитком" (Партеніт, 2002), I Міжнародній конференції з індуктивного моделювання – ІСІМ'2002 (Львів, 2002), IV Міжнародній конференції "Проблеми і перспективи використання геоінформаційних технологій у гірництві" (Дніпропетровськ, 2002), IV, VI та VII Міжнародних науково-технічних конференціях "Геоінформаційні технології в управлінні територіальним розвитком" (Ялта, 2001, 2003, 2004).

Публікації. Наукові положення та результати досліджень за темою дисертації опубліковані в 47 друкованих працях: монографії – 2, у фахових виданнях за переліком ВАК України – 23 статті (в наукових журналах – 9, у збірниках наукових праць – 14), у збірниках робіт конференцій – 9, в тезах конференцій – 9, в інших виданнях – 4. Основні публікації розміщені для ознайомлення на сайтах <http://www.sarycheval.dp.ua>; <http://gis.nmu.org.ua>.

Структура та обсяг роботи. Дисертація складається зі вступу, шести розділів, висновків, переліку умовних скорочень і позначень, списку літератури з 291 джерела посилань, актів упровадження і додатків. Загальний обсяг роботи – 367 сторінок, у тому числі 90 рисунків, 20 таблиць, 10 додатків.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі проаналізовано стан проблеми, обґрунтовано актуальність, практичну та теоретичну цінність тематики. Виділено коло основних задач, визначено мету і загальну методику досліджень, описано структуру дисертації.

Структурну діаграму логічних взаємозв'язків основних розділів дисертації зображено на рис. 1 (усередині автореферату на розгортці).

У **розділі 1** на основі аналітичного огляду стану проблеми і досліджень за темою дисертації обґрунтовано нагальну потребу розробки такої тематики:

1) ЕСЕ-моніторинг регіонів потребує спеціалізованого математичного забезпечення, яке б урахувало просторово-часовий характер і різномірність показників екологічного та соціально-економічного моніторингу;

2) підвищення оперативності та ефективності ЕСЕ-моніторингу можливе лише на сучасних інформаційних технологіях, що завдяки просторово координованому характеру ЕСЕМ-даних містять ГІС-технології.

В Україні спільного аналізу екологічних і соціально-економічних моніторингових даних у ГІС не здійснюється, аналітичний апарат ГІС для оцінки збалансованості регіонального розвитку тільки починає застосовуватись.

Викладено концепцію, проблеми та перспективи, місце в структурі регіонального управління, види і напрями, вимоги до математичного забезпечення ЕСЕ-моніторингу регіонів. ЕСЕ-моніторинг містить у собі: ведення єдиної бази, візуалізацію та аналіз даних; побудову ЕСЕ-моделей; прогноз ЕСЕ-показників; формування управлінських рішень на основі моделювання і прогнозу. Він визначається цільовим характером спостережень з орієнтацією на існуючі тенденції та локальні зміни, можливості виникнення несприятливих ситуацій; інтегрованим врахуванням ЕСЕ-інформації; включенням у сферу спостережень і аналізу не тільки числових даних, а й «статистики зв'язків», «статистики факторів» (коли зафіксовано одне явище, а дослідженню піддаються як зв'язки, так і коло явищ, що впливає на дане); широким використанням комп'ютерних методів подання та візуалізації ЕСЕ-інформації.

За сучасних умов глобалізації та переходу на інтегровану ЕСЕ-оцінку території виникає нагальна потреба розробки системи ЕСЕ-аудиту. В розділі обґрунтовано взаємодоповнювальні задачі ЕСЕ-моніторингу та ЕСЕ-аудиту, розроблено методологію ЕСЕ-аудиту. Для ЕСЕ-моніторингу та ЕСЕ-аудиту рекомендовано методику інтегральної ЕСЕ-оцінки території, яка базується на спільному аналізі трьох інформаційних блоків:

1) екологічного (хімічне, фізичне, радіаційне, біологічне забруднення, витрати на зменшення і контроль за забрудненнями тощо);

2) економічного (фінансові ресурси і механізми, зміна структури споживання, ефективність споживання ресурсів і т.д.);

3) соціального (демографічні дані, освіта, охорона здоров'я та ін.).

У цих блоках виділяються три типи показників: фактори забруднення (а), показники стану (б) та управляючого впливу (с), що являють собою змінні, які беруть участь у побудові ЕСЕ-моделей типу "вплив-стан-відгук" і пов'язані функціонально залежністю.

Визначаються узагальнені інтегральні показники a_i , b_i , c_i , $i=1, 2, 3$ (J – єдине позначення для них) для кожного інформаційного блоку і всіх n регіонів території, тобто заповнюються матриці виду

$$\Phi_j = [A_j | B_j | C_j], \quad A_j = (a_1, a_2, a_3)_j^T, \quad B_j = (b_1, b_2, b_3)_j^T, \quad C_j = (c_1, c_2, c_3)_j^T, \quad j=1, 2, \dots, n.$$

Для цього m групам базових показників (у кожному з блоків – декілька груп показників), а також базовим показникам i -ї групи (позначимо їх x_{ij} ; число показників i -ї групи – m_i) залежно від їх значущості надаються ваги P_i та p_{ij} (P_i – вага i -ї групи, $0 \leq P_i \leq 1$, $\sum_{i=1}^m P_i = 1$; p_{ij} – вага j -го показника i -ї групи, $0 \leq p_{ij} \leq 1$, $\sum_{j=1}^{m_i} p_{ij} = 1$, $i=1,2,\dots,m; j=1,2,\dots,m_i$).

Інтегральний показник обчислюється за формулою:

$$J = \sum_{i=1}^m \sum_{j=1}^{n_i} p_{ij}^* \frac{|x_{ij} - x_{ij}^{(t) \min(\max)}|}{x_{ij \max} - x_{ij \min}}, \quad (1)$$

де $x_{ij}^{(t) \min(\max)}$ – екстремальне значення показника, задане залежно від напрямку оптимізації, тобто від того, є показник стимулятором чи дестимулятором: при $t=1$ приймається мінімальне значення – $x_{ij}^{(1) \min}$, при $t=2$ – максимальне $x_{ij}^{(2) \max}$ (при $t=1$ значення $x_{ij}^{(t)}$ максимізується, при $t=2$ – мінімізується), p_{ij}^* – вага j -го показника i -ї групи з урахуванням ваги i -ї групи.

У просторі R^3 ("екологія", "економіка", "соціум") j -й регіон ($j=1,2,\dots,n$) подається векторами A_j, B_j, C_j , що характеризують фактори забруднення, показники стану та управляючі впливи. Показники $\|A_j\|, \|B_j\|, \|C_j\|$ (де $\|\cdot\|$ – норма) є атрибутами шару "регіони" в ГІС. Геоіконічне моделювання в ГІС (розд. 4) здійснюється за значеннями показників $\|A_j\|, \|B_j\|, \|C_j\|$.

У просторі R^3 ("фактори забруднення", "показники стану" та "управляючі впливи") j -й регіон ($j=1,2,\dots,n$) подається векторами $D_j=(a_1, b_1, c_1)_j, F_j=(a_2, b_2, c_2)_j, G_j=(a_3, b_3, c_3)_j$, що характеризують його екологію, економіку та соціум.

У цих же просторах подаються вектори $Q_j=(\|A_j\|, \|B_j\|, \|C_j\|), S_j=(\|D_j\|, \|F_j\|, \|G_j\|)$ та $\|F_j\|$, що слугують більш узагальнювальними характеристиками j -го регіону. За значеннями показників $\|D_j\|, \|F_j\|, \|G_j\|, \|F_j\|$ також відбувається геоіконічне моделювання в ГІС. Інтегральні ЕСЕ-оцінки територій $A^*, B^*, C^*, D^*, F^*, G^*, \Phi^*$ обчислюються як середні значення відповідних атрибутів для всіх регіонів. За допомогою кластерного аналізу (розд. 2) за 9-ма ознаками (елементи матриці Φ_j) виділяються групи регіонів зі схожими ЕСЕ-показниками, територія розбивається на однорідні за рівнем регіонального розвитку області. Результати кластерного аналізу – це атрибут "номер класу" для шару "регіони" (візуалізуються в ГІС).

Визначальні риси запропонованої методики – врахування «трубки сталості», у межах якої перебувають ЕСЕ-показники при збалансованому розвитку. Розроблена процедура обчислення інтегральної оцінки входить до блока генерації, вибору та прийняття рішень комп'ютерної системи ЕСЕ-моніторингу регіонів (розд. 5).

Розділ 2 присвячено дослідженню і реалізації методів та алгоритмів ІАД, які становлять основу математичного забезпечення ЕСЕ-моніторингу регіонів. Класифіковано задачі ІАД ЕСЕ-моніторингу: ранжування, розпізнавання, кластеризація, прогнозування, виділення аномалій, аналіз та оптимізація мережі.

Розглянуто аналіз ЕСЕМ-даних, який базується на методах математичної теорії розпізнавання образів. Доцільність їх застосування обумовлена таким:

- 1) одні й ті ж методи можна використати для розв'язання широкого класу різних за змістом задач (наприклад, методами кластеризації агрегуються економічні показники, множина регіонів розділяється на групи умовно однорідних підмножин та ін.);
- 2) методи розпізнавання образів дають змогу робити комплексний аналіз різнорідних даних трьох інформаційних блоків: екологічного, економічного, соціального.

Нехай $X = \{X_1, X_2, \dots, X_n\}$ – скінченна множина об'єктів (деяких регіонів території). Кожен об'єкт характеризується набором ознак (показників ЕСЕ-моніторингу): $X_i = (x_{i1}, x_{i2}, \dots, x_{im})$, $i = 1, 2, \dots, n$. Вихідні дані являють собою таблиці типу “об'єкт-ознака”: $X = \{x_{ij}\}$, $i = 1, 2, \dots, n$; $j = 1, 2, \dots, m$. Наявність певної спільності властивостей у різних об'єктів дозволяє групувати об'єкти в деякі підмножини K_1, K_2, \dots, K_k множини X – класи. Для групування використовуються міри близькості між об'єктами $d(X_i, X_j)$, класами об'єктів $d(K_i, K_j)$, об'єктом і класом $d(X_j, K_i)$, ознаками об'єктів $d(X^i, X^j)$. У розділі наведено міри близькості, які застосовуються далі в алгоритмах ранжування, розпізнавання, кластерного аналізу, виділення аномальних об'єктів, аналізу мережі моніторингу.

Характеру задач розпізнавання при класифікації регіонів за сукупністю показників ЕСЕ-моніторингу більше відповідають непараметричні методи. Це зумовлено тим, що кількість регіонів невелика (наприклад, у масштабах України $n=27$, а в масштабах Дніпропетровської області – $n=35$), тому важко оцінити невідомі параметри щільності розподілу $f(X | K_i)$ класу K_i у m -вимірному ознаковому просторі за вибірками малого обсягу для використання параметричних методів. Якщо все ж таки вибрати параметричні методи, наприклад, дискримінантний аналіз, то необхідно розв'язувати задачу в широкому розумінні, тобто при побудові дискримінантної функції знаходити оптимальну (з сукупності досліджуваних) множину ознак.

З єдиних системних позицій класифіковано методи кластеризації та методи вибору інформативних ознак, де і кластеризація, і вибір ознак – задачі дискретної оптимізації на вершинах ґратки розбивань.

Нехай маємо k класів об'єктів: K_1, K_2, \dots, K_k . Кожен об'єкт описано m ознаками. Клас K_l , $l = 1, 2, \dots, k$, містить n_l об'єктів: $n_1 + n_2 + \dots + n_k = n$. Під вибором інформативних ознак розуміють відображення $F: \{X^m\} \xrightarrow{F} \{X^{m_1}\}$, $m_1 < m$, при якому досягається екстремум деякого функціоналу якості $J_X(F)$. Кожному такому відображенню можна поставити у відповідність вектор $V = (v_1, v_2, \dots, v_m)$, де $v_i = 1$, якщо i -а ознака входить до складу ознак, а в протилежному разі – $v_i = 0$. Тоді функціонал $J_X(F)$ можна розглядати як функцію $g_{m1}(V)$, задану на вершинах одиничного гіперкуба $[0, 1]^m$ (рис. 2 а). З усіх вершин слід знайти таку, де досягається екстремум $g_{m1}(V)$.

Тоді постановка задачі вибору інформативної підсистеми ознак з початкової має вигляд

$$g_{m1}(V) \rightarrow \min_{V \in D}; \quad D = \{V \in R^m; v_j \in \{0, 1\}\}, \quad (2)$$

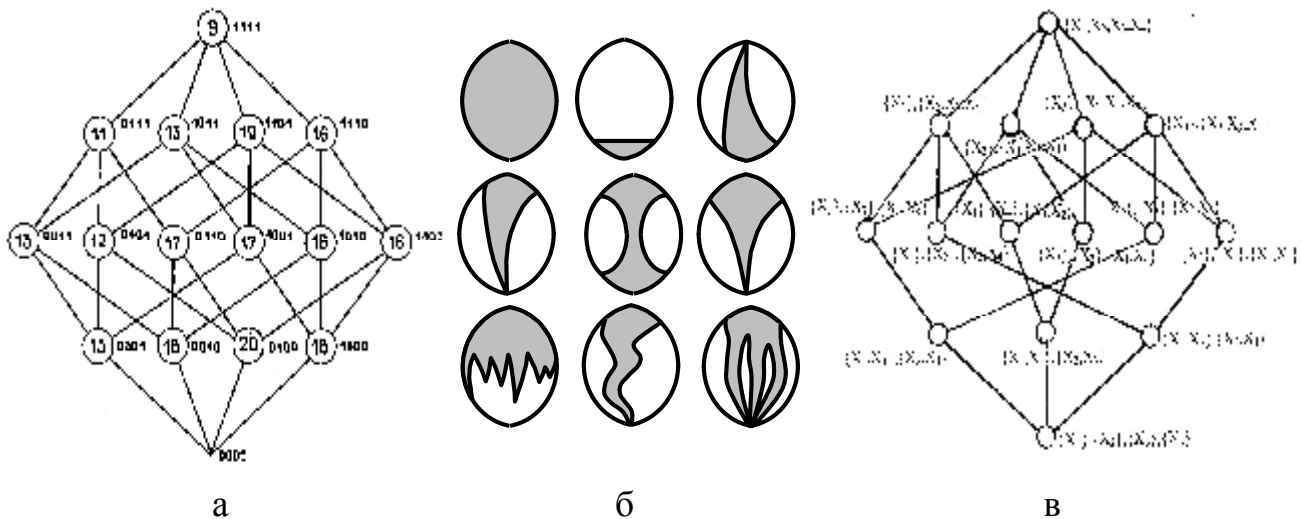


Рис. 2. Єдиний підхід до класифікації методів кластеризації та методів вибору ознак у розпізнаванні образів: а, в – приклади ґраток розбивань вибору ознак ($m=4$) та пошуку найкращої кластеризації ($n=4$) відповідно; б – типові шляхи пошуку в гіперкубі

якщо розмірність m_1 шуканої підсистеми не визначена, або

$$g_{m_1}(V) \rightarrow \min_{V \subset G}; \quad G = \left\{ V \subset D; \sum_{j=1}^{m_1} v_j = m_1 \right\}, \quad (3)$$

якщо розмірність шуканої підсистеми задана. Елементами множини D є вершини ґратки розбивань одиничного m -вимірного гіперкуба. Функція $g_{m_1}(V)$ найчастіше багатоекстремальна і її не можна дослідити аналітично. Тому для знаходження екстремуму використовують різні процедури перебору вершин з поточною оцінкою їх цінності. Критерій якості визначається за апіорними даними та з попереднього аналізу сукупності ознак. Для знаходження оптимальної підмножини ознак загалом відомий лише один метод – повний перебір усіх підмножин початкової множини ознак. На практиці застосовують субоптимальні методи пошуку, що скорочують перебір варіантів.

Кластеризація $K(X) = \{K_1(X), K_2(X), \dots, K_k(X)\}$, $1 \leq k \leq n$, скінченної множини об'єктів $X = \{X_1, X_2, \dots, X_n\}$ – сімейство непорожніх, попарно неперетинальних підмножин (кластерів) $K_l(X)$, $l=1, 2, \dots, k$, множини X , об'єднання яких збігається з X : $K_1 \cup K_2 \cup \dots \cup K_k = X$, $K_i \cap K_j = \emptyset$ при $i \neq j$, $K_l \neq \emptyset$, $l=1, 2, \dots, k$.

Число кластерів k може бути заздалегідь невідомим. Більшість алгоритмів кластеризації складається з двох етапів: 1) задається початкове розбиття множини об'єктів на класи і визначається деякий критерій якості; 2) об'єкти переносяться з класу в клас, поки значення критерію не перестане покращуватись.

У задачі вибору ознак і в задачі пошуку найкращої кластеризації використовується та ж таблиця "об'єкт-ознака". Коли під пошуком (вибором) найкращої кластеризації розуміти звужуюче відображення Φ множини вихідних об'єктів на мно-

жину кластерів K : $\{X_n\} \xrightarrow{\Phi} \{K_k\}$, $k \leq n$, при якому досягається екстремум деякого функціонала якості $I_X(\Phi)$, то постановку задачі пошуку кластеризації можна подати у вигляді, аналогічному (2), якщо число кластерів не задано, або у вигляді, аналогічному (3), якщо число k кластерів задано (рис. 2 в). Тому існує відповідність між методами вибору інформативних ознак і методами пошуку найкращої кластеризації (наприклад, методів послідовного з'єднання відповідає побудова дивізімного дерева, а методу послідовного відкидання – агломеративного).

Типові шляхи пошуку (рис. 2 б) у гіперкубі: 1) послідовний перебір варіантів (повний перебір, ранжування, послідовне приєднання, послідовне відкидання, різні комбінації послідовного приєднання і послідовного відкидання, двосторонній пошук та ін.); 2) випадковий пошук та його модифікації (метод Монте-Карло, випадковий пошук з адаптацією, випадковий пошук з поверненням, модифікований випадковий пошук з адаптацією тощо); 3) комбіновані методи (метод гілок і границь, мінімакський метод, метод синтезу та ін.); 4) генетичні методи. Такий підхід, в якому і кластеризація, і вибір ознак є задачами дискретної оптимізації на вершинах ґраток розбивань, дозволяє побачити нові напрямки розвитку цих методів. Наприклад, при побудові ієрархічних дерев кластеризації можна назустріч один одному будувати два дерева: одне за допомогою поступового збільшення числа кластерів, а друге – зменшення; там, де вони зустрінуться – буде найкраща кластеризація. Така побудова в задачі вибору ознак відповідає двосторонньому пошукові. Отримані результати враховуються далі при генерації кластерів.

Викладено нову методику об'єктивної кластеризації, розроблену відповідно до принципів методу групового врахування аргументів О.Г.Івахненка: багатоетапність пошуку кращої кластеризації, оцінка якості кластеризації за допомогою внутрішнього і зовнішнього критеріїв, використання наборів методів утворення кластерів і мір близькості між двома об'єктами.

Для обґрунтування вірогідності результатів кластерного розв'язку в запропонованій методиці використано два зовнішні критерії, перший з яких ґрунтується на розбитті вибірки на дві рівнопотужні підвибірки в просторі об'єктів, а другий – у просторі ознак. При першому способі початкова множина X , що містить n об'єктів, розбивається на підмножини A і B ($A \cap B = \emptyset$, $A \cup B = X$) розмірності $(n/2) \times m$:

$$X^T = [X_A^T \mid X_B^T],$$

при другому – на підмножини A і B ($A \cap B = \emptyset$, $A \cup B = X$) розмірності $n \times (m/2)$:

$$X = [X^A \mid X^B].$$

Кластеризація вважається об'єктивною, якщо у кластеризаціях, проведених окремо для кожної з двох рівнопотужних підмножин, – найбільший збіг.

Клас синтезованих кластеризацій має вигляд

$$\mathbf{K}(X) = \{ \mathbf{K}_1(X), \mathbf{K}_2(X), \dots, \mathbf{K}_k(X) \}, 1 \leq k \leq n, n_1 + n_2 + \dots + n_k = n,$$

де k – число кластерів, n_i – число об'єктів у кластері $\mathbf{K}_i(X)$.

Частинною кластеризацією назовемо будь-яку кластеризацію $\mathbf{Z}(X)$, отриману на деякій ітерації алгоритму як наближення об'єктивної кластеризації $\mathbf{K}(X)$. Час-

тинна кластеризація описується вектором розмірності n номерів кластерів: j -а координата вектора – номер кластера, до якого потрапив j -й об'єкт. Структурою частинної кластеризації назвемо набір параметрів n_j та k_j , що визначають \mathbf{Z}^r . Загальний вигляд матриці частинних кластеризацій \mathbf{Z}^r визначимо таким чином: $\mathbf{Z}^r = [\mathbf{z}_1^r \mathbf{M} \mathbf{z}_2^r \mathbf{M} \dots \mathbf{M} \mathbf{z}_{g+f}^r]$, де \mathbf{z}_j^r , $j=1,2,\dots,g+f$, – вектори розмірності n , частинні кластеризації; g – число кращих частинних кластеризацій, переданих від ітерації до ітерації; f – число «сусідів» кращих частинних кластеризацій ітерації $(r-1)$, $f \leq g \times 2c$, де $2c$ – число «сусідів» для однієї кращої кластеризації (якщо вона виявляє 5 кластерів, то її «сусіди» при $c=2$ – кластеризації з кількістю кластерів 3, 4, 6, 7).

Для виявлення «сусідів» рекомендовано також обчислювати міру близькості між кластеризаціями K і Q :

$$d^*(K, Q) = \frac{d(K, Q)}{\frac{1}{2} \left(\sum_{i=1}^{k_1} |K_i|^2 + \sum_{i=1}^{k_2} |Q_i|^2 \right)}; \quad (4)$$

$$d(K, Q) = \frac{1}{2} \left(\sum_{i=1}^{k_1} |K_i|^2 + \sum_{i=1}^{k_2} |Q_i|^2 \right) - \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} |K_i \mathbf{I} Q_j|^2, \quad (5)$$

де k_1, k_2 – число кластерів у кластеризаціях K і Q відповідно, $|K_i|, |Q_j|$, $i=1,2,\dots,k_1$, $j=1,2,\dots,k_2$, – число об'єктів у кластерах K_i, Q_j .

Позначимо: $\mathbf{G}^r = [\mathbf{z}_1^r \mathbf{M} \mathbf{z}_2^r \mathbf{M} \dots \mathbf{M} \mathbf{z}_g^r]$, $\mathbf{C}^r = [\mathbf{z}_{g+1}^r \mathbf{M} \mathbf{z}_{g+2}^r \mathbf{M} \dots \mathbf{M} \mathbf{z}_{g+f}^r]$.

Алгоритм складається з таких кроків.

1. Заповнюється початкова матриця частинних кластеризацій \mathbf{Z}^0 : $\mathbf{Z}^0 = [\mathbf{G}^0 \mathbf{M} \mathbf{C}^0]$. Для цього послідовно здійснюються: нормування, формування таблиці міжточкових відстаней, формування списку пар точок, ранжованих за зростанням міжточкової відстані, поділ вихідної множини на підмножини A і B першим способом (q -й об'єкт – у підмножину A , а найближчий за мірою близькості до нього s -й об'єкт – у підмножину B), генерація кластерів (одним з ієрархічних алгоритмів або алгоритмом k середніх, де число кластерів $k=2,3,\dots, n/2$) для вихідної множини X , підмножин A і B (кожної окремо). Із усіх згенерованих частинних кластеризацій початкової множини X відбираються g кластеризацій, кращих за мінімумом розбіжності структури кластеризацій на підмножинах A і B :

$$J = \frac{1}{N(A, B)} \sum_{i=1}^{N(A, B)} e(i), \quad (6)$$

$$e(i) = \begin{cases} 0, & \text{якщо } \mathbf{z}_A^r(i) \leftrightarrow \mathbf{z}_B^r(i), \\ 1 - \text{у протилежному випадку,} \end{cases}$$

де $N(A, B)$ – число об'єктів, відстань між якими менша за поріг. Поріг при першому способі розбиття на A і B визначається на підставі обчислення внутрі- та міжкластерних мір близькості в кластеризації множини X .

Відібрані кращі частинні кластеризації, ранжовані за збільшенням величини J , використовуються при формуванні матриці $\mathbf{G}^0 = [\mathbf{z}_1^0 \mathbf{M}_2^0 \mathbf{M} \cdot \mathbf{M}_g^0]$. Число кластерів у кращих кластеризаціях – $k^0(1), k^0(2), \dots, k^0(g)$. \mathbf{C}^0 формується з “сусідів” кращих частинних кластеризацій.

2. Визначається оператор R перетворення $\mathbf{z}^{r-1} \xrightarrow{R} \mathbf{z}^r$. Для заданих способу нормування, міри близькості між об'єктами, виду розбиття початкової множини на підмножини A і B , методу утворення кластерів генеруються частинні кластеризації. Відібрані кращі (за мінімумом розбіжності структури на підмножинах A і B) частинні кластеризації, ранжовані за збільшенням величини J , порівнюються з кластеризаціями попередньої ітерації $\mathbf{G}^{r-1} = [\mathbf{z}_1^{r-1} \mathbf{M}_2^{r-1} \mathbf{M} \cdot \mathbf{M}_g^{r-1}]$ і використовуються при формуванні наступної матриці $\mathbf{G}^r = [\mathbf{z}_1^r \mathbf{M}_2^r \mathbf{M} \cdot \mathbf{M}_g^r]$. Матриця \mathbf{C}^r формується з “сусідів” кращих частинних кластеризацій матриці \mathbf{G}^r .

3. Правило зупинки для ітераційної схеми: обчислення закінчуються на ітерації r^* , якщо виконано умову $J(\mathbf{z}_F^{r^*-1}) - J(\mathbf{z}_F^{r^*}) < \delta_r$, де $J(\mathbf{z}_F^{r^*})$ – значення критерію для кращої часткової кластеризації ітерації r ; δ_r – задане число.

Характерні риси методики: ітераційний пошук числа реальних кластерів, оцінка якості кластеризації за допомогою внутрішнього і зовнішнього критеріїв, використання наборів мір близькості та алгоритмів кластеризації.

Відображення результатів аналізу ЕСЕМ-даних, заснованого на методах теорії розпізнавання образів, здійснюється в ГІС, для чого запропоновано геоіконічні моделі (розд. 4).

Розглянуто питання графового аналізу просторово координованих ЕСЕМ-даних: типи графових моделей структур територіальних відносин і зв'язків, характеристики структури графа. Класифіковано параметри структури графа: 1) центральність вершин (індекси досяжності вершини, відхилення від центральної вершини, ієрархії, індекси Бавела, Бошама, Кьоніга та ін); 2) центральність графа в цілому (інтеграція, уніполярність, централізація); 3) зв'язність (α , β , ϕ індекси); 4) форма (індекс компактності, π індекс). Ці дослідження застосувань графів використано в алгоритмах аналізу та оптимізації ЕСЕМ-мережі.

ЕСЕ-прогнозування – один з вирішальних наукових факторів формування стратегії і тактики регіонального розвитку. У роботі проведено класифікацію методів прогнозування та запропоновано новий алгоритм структурної ідентифікації для прогнозування ЕСЕ-показників, розроблений відповідно до принципів методу групового врахування аргументів.

Нехай закон функціонування об'єкта має вигляд

$$y = y + \xi = \sum_{j=1}^m \Theta_j^o \cdot x_j + \xi, \quad (7)$$

де y – вихід спостережуваного об'єкта; y° – незашумлений неспостережуваний вихід; x – випадкова помилка вимірювань, що не спостерігається; x_j – j -й вхід об'єкта із множини входів χ , які беруть участь у формуванні виходу об'єкта;

m – число належних множині χ входів; $\Theta^{\circ} = (\Theta_1^{\circ}, \Theta_2^{\circ}, \dots, \Theta_m^{\circ})^T$ – вектор не рівних нулю невідомих коефіцієнтів; T – знак транспонування.

Множина входів χ невідома. Лише відомо, що $\chi \subset C$, де C – деяка множина точно вимірюваних m входів об'єкта.

Нехай у результаті спостереження за об'єктом отримані: 1) X – матриця N спостережень m входів множини C , $\text{rang} X = m$; 2) y – вектор відповідних спостережень вихідної величини y . Відповідно до закону функціонування об'єкта (7) виконується рівність

$$y = y^{\circ} + \xi = X^{\circ} \Theta^{\circ} + \xi, \quad (8)$$

де y° – вектор значень незашумленого неспостережуваного виходу об'єкта;

X° – матриця ($N \times m$) спостережень входів об'єкта, що належать множині χ ;

X – вектор випадкових помилок вимірювань, що не спостерігаються.

Нехай відносно X виконано припущення

$$E\{x\} = \mathbf{0}_N, \quad E\{xx^T\} = \sigma^2 \mathbf{I}_N, \quad (9)$$

де E – знак математичного сподівання; $\mathbf{0}_N$ – нульовий вектор; σ^2 – невідома скінченна величина; \mathbf{I}_N – одинична матриця. Знайти: 1) множину χ ; 2) оцінку вектора коефіцієнтів Θ° ; 3) оцінку дисперсії помилок вимірювань σ^2 .

Клас синтезованих моделей має вигляд

$$y = \sum_{q=1}^s \Theta_q \cdot \prod_{j=1}^m x_j^{\alpha_{qj}}, \quad (10)$$

де y – вихідна змінна; s – число членів у моделі; Θ_q , $q=1,2,\dots,s$, – коефіцієнти; x_j , $j=1,2,\dots,m$, – вхідні змінні; m – число вхідних змінних; α_{qj} – показник степеня, в якому змінна x_j входить до q -го члена.

Нехай матриця Z^r деяких наближень вектора y має вигляд

$$Z^r = [z_1^r \ M z_2^r \ M \dots \ M z_{F+2+m+2s}^r], \quad (11)$$

де z_j^r , $j=1,2,\dots, F+2+m+2s$ – вектори, частинні описи; F – число кращих частинних описів, переданих від ітерації до ітерації; s – число членів у структурі кращого частинного опису ітерації ($r-1$).

Алгоритм складається з таких кроків.

1. Вказується початкова матриця

$$\mathbf{Z}^0 = [\mathbf{O}_{N \times F} \quad \mathbf{M}_N \quad \mathbf{M}_N \quad \mathbf{M}_N] = [\mathbf{O}_{N \times F} \quad \mathbf{M}^0] \quad (12)$$

частинних описів \mathbf{Z}^0 , де $\mathbf{O}_{N \times F}$ – нульова матриця; \mathbf{o}_N – нульовий вектор; \mathbf{I}_N – вектор з одиниць; $\mathbf{X} = [\mathbf{x}_1 \quad \mathbf{M}_2 \quad \mathbf{M}_3 \quad \mathbf{M}_m]$ – матриця спостережень вхідних змінних.

2. Визначається оператор R перетворення $\mathbf{Z}^{r-1} \xrightarrow{R} \mathbf{Z}^r$. З усіх згенерованих частинних описів відбираються F кращих за мінімумом квадратичної норми залишків на перевірній підвибірці спостережень (B). Проранжовані кращі частинні описи використовуються при формуванні перших F векторів матриці \mathbf{Z}^r , останні $2s$ векторів формуються з урахуванням структури кращого серед F відібраних частинних описів.

Синтез моделей починається з етапу за номером $p=1$ (або з будь-якого заданого номера p_0). Кожен етап являє собою ітераційну схему. Початкова матриця частинних описів етапу за номером p задається кінцевою матрицею частинних описів попереднього етапу. Обчислення закінчуються на етапі p^* , якщо виконано умову

$$J(\mathbf{Z}_{F,p^*-1}^r) - J(\mathbf{Z}_{F,p^*}^r) < \delta_p, \quad (13)$$

де $J(\mathbf{Z}_{F,p}^r)$ – значення критерію для кращого частинного опису r -ї ітерації p -го етапу; δ_p – задане число.

Характерні риси алгоритму: 1) багатоетапність пошуку моделі; 2) пошук моделі у класах як лінійних, так і нелінійних за вхідними змінними моделей; 3) прийоми вилучення окремих членів кращого частинного опису і на основі цього розширення базисного набору аргументів; 4) оптимальна за обчислювальними витратами для ітераційних алгоритмів методу групового врахування аргументів схема розрахунку критерію ковзного екзамину; 5) можливість оцінювати коефіцієнти в моделях як за методом найменших квадратів, так і за методом найменших модулів. У процесі побудови моделі відбувається вибір підмножин інформативних показників.

Систематизовано відомі підходи та запропоновано нові алгоритми виділення аномалій регіонального розвитку, що базуються на основі: 1) адаптивності та різноманітності; 2) міри близькості; 3) кластерного аналізу; 4) поляризації; 5) аналізу регресійних залишків.

Аномаліям часових тенденцій розвитку відповідає постійне зростання або спад ЕСЕ-показників регіону-аномалії, а багатовимірним аномаліям у межах вибраного тематичного сюжету (m аналізованих ЕСЕ-показників для кожного з n регіонів) відповідають полярні хвости багатовимірних розподілів. Для врахування часових тенденцій зміни ЕСЕ-показників і виділення аномалій пропонується алгоритм, що використовує поняття поляризації, адаптивності та різноманітності.

Нехай $x_i(t)$ – значення ЕСЕ-показника i -го ($i=1,2,\dots,n$) регіону в момент часу t . Позначимо дві градації показника:

$$0 \text{ – якщо } x_i(t) < \mu(t); \quad 1 \text{ – якщо } x_i(t) \geq \mu(t),$$

де $\mu(t) = \frac{1}{n} \sum_{i=1}^n x_i(t)$. Залежно від того, чи буде значення показника вище або нижче

середнього значення в послідовні моменти часу $t, t+1, t+2$, всі регіони можна розбити на 8 класів (деякі з них можуть виявитися порожніми):

ЕСЕ-показник	Клас							
	1	2	3	4	5	6	7	8
$x_i(t)$	0	0	1	1	0	0	1	1
$x_i(t+1)$	0	1	0	1	0	1	0	1
$x_i(t+2)$	0	0	0	0	1	1	1	1

До аномальних зараховуються 1-й та 8-й класи. Для трьох градацій:

0 – якщо $x_i(t) < \mu(t) - \sigma(t)$; 1 – якщо $\mu(t) - \sigma(t) \leq x_i(t) < \mu(t) + \sigma(t)$; 2 – якщо $x_i(t) \geq \mu(t) + \sigma(t)$,

де $\sigma(t) = \left(\frac{1}{n} \sum_{i=1}^n (x_i(t) - \mu(t))^2 \right)^{1/2}$, і двох моментів часу $t, t+1$, усі регіони розбиваються

на дев'ять класів. Цей процес можна продовжувати, задаючи число послідовних моментів часу та кількість градацій ЕСЕ-показника. Приклад можливих класів розбиття у разі чотирьох різних показників $x_i(t), y_i(t), z_i(t), g_i(t)$ (або одного показника для моментів часу $t, t+1, t+2, t+3$) графічно відображає рис. 2 а.

Подібним чином аналізуються три величини: значення $x_i(t)$, швидкість $v_i(t)$ та прискорення $w_i(t)$ ЕСЕ-показника $x_i(t)$ в моменти часу $t=1, 2, \dots, q$. Аналогічно розглядаються декілька показників тематичного сюжету для одного моменту часу.

У розроблених алгоритмах виділення аномалій (за тематичним сюжетом) обчислюються запропоновані міри близькості між об'єктом і класом: потенціальна функція, кутова міра близькості, відстань Махаланобіса, проекція на півпростір, відстань до центра ваги класу, функція геометричної міри близькості. Припускається, що існує закономірність функціонування всієї множини регіонів (можливо, за винятком аномальних регіонів). Ця множина являє собою один клас. Розглядаються міри близькості між об'єктом (аналізованим на аномальність) і класом. Вибрана міра близькості обчислюється для кожного з досліджуваних об'єктів. Здійснюється ранжування об'єктів за її значеннями. Об'єкти, розташовані на перших (останніх) місцях у ранжованій послідовності, приймаються за аномалії.

Виділення аномалій на основі аналізу регресійних залишків базується на припущенні, що побудована для множини регіонів регресійна модель відображає їх загальну тенденцію. Регіони, де регресійні залишки найбільші, від цієї тенденції найвіддаленіші. За допомогою розробленого алгоритму структурної ідентифікації розв'язується задача відшукування регресійної залежності одного з ЕСЕ-показників моніторингу від сукупності решти.

Нехай знайдена модель має вигляд

$$\hat{y} = f(\mathbf{X}, \hat{\Theta}_f), \quad (14)$$

де $\hat{\Theta}_f$ – вектор параметрів. Знаходиться вектор регресійних залишків $(\mathbf{y} - \hat{y})$. Регіони, що відповідають найбільшим компонентам цього вектора, приймаються за аномалії. Даний алгоритм можна застосовувати для виділення аномалій як у межах вибраного тематичного сюжету, так і для аномалій, що визначаються за часовими тенденціями розвитку.

Для виділення аномалій рекомендовано застосовувати розглянуті процедури кластеризації з подальшим аналізом відстаней між об'єктами всередині кластера і центрами кластерів з урахуванням числа об'єктів у кожному з них. Якщо центр кластера далеко знаходиться від інших центрів, і кластер містить малий процент об'єктів, то можна припустити, що він – аномалія. Знайдені за різними алгоритмами регіони-аномалії відображаються в ГІС для проведення просторового аналізу, уточнення, узгодження та інтерпретації ОПР отриманих результатів.

З усієї різноманітності підходів до побудови, аналізу та оптимізації мережі моніторингу виділено сім основних, використання яких доцільне в ЕСЕ-моніторингу регіонів: 1) геометричний, 2) статистичний, 3) теоретико-графовий, 4) інтерполяційний, 5) кластерний, 6) геостатистичний, 7) фрактальний. Кожен із цих підходів дозволяє описати окремі характеристики мережі. Мережу екологічного моніторингу рекомендовано досліджувати, використовуючи послідовно геометричний, статистичний, теоретико-графовий, кластерний, геостатистичний, інтерполяційний, фрактальний підходи. Застосування геометричного підходу для дослідження мережі соціально-економічного моніторингу регіонів України та Дніпропетровської області свідчить про їх подібність до відповідної європейської мережі.

Запропоновані в розділі 2 обчислювальні процедури становлять основу блока ІАД комп'ютерної системи ЕСЕ-моніторингу регіонів (розділ 5).

У **розділі 3** викладено методологію вибору стратегії розвитку регіону на базі ЕСЕ-моделей. На основі класифікації і порівняльного аналізу існуючих еколого-економічних моделей регіону обґрунтовано вимоги до ЕСЕ-моделі, яка б відображала функціональні зв'язки між основними ЕСЕ-показниками та давала можливість: 1) прогнозувати зміни екологічних показників від впливу виробничої діяльності; 2) порівнювати варіанти розвитку регіону за економічними показниками з урахуванням еколого-соціальних показників; 3) шукати сценарії розвитку регіону, оптимальні в сенсі раціонального природокористування.

Розроблено спрощену структурну циркуляційну схему ЕСЕ-системи регіону (рис. 3). Для математичної ЕСЕ-моделі першого наближення, що відповідає наведеній схемі, можна застосувати багатовимірні динамічні балансні системи, досліджені Л.М. Бойчуком.

Для ЕСЕ-моніторингу доцільно використовувати ЕСЕ-моделі, рівняння яких перетворені з урахуванням структури інвестицій, що створює переваги для аналізу міжгалузевих пропорцій та різноманітних планових розрахунків. Для цього запропонована балансова ЕСЕ-модель регіону (базується на моделі Гурмана В.І.):

$$\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{B}\dot{\mathbf{g}} + \mathbf{C}\mathbf{y} + \mathbf{D}\dot{\mathbf{e}} + \mathbf{p}; \quad (15)$$

$$\dot{\mathbf{z}} = \mathbf{Q}(\mathbf{z} - \mathbf{z}_0) - (\mathbf{R}\mathbf{x} + \mathbf{H}\dot{\mathbf{g}}) + \mathbf{G}\mathbf{y} - \mathbf{S}\dot{\mathbf{e}} - \mathbf{v}\mathbf{N} - \mathbf{L}\mathbf{p}; \quad (16)$$

$$\mathbf{0} \leq \mathbf{x} \leq \mathbf{g}, \quad \mathbf{0} \leq \mathbf{y} \leq \mathbf{e}, \quad \mathbf{z}_{\min} \leq \mathbf{z} \leq \mathbf{z}_{\max}, \quad (17)$$

де $\mathbf{x}(t)$, $\mathbf{g}(t)$, $\mathbf{y}(t)$, $\mathbf{e}(t)$ – вектори обсягів виробництва продукції, потужностей галузей виробництва, інтенсивності відновлення ресурсів, потужностей відновлювальних галузей у момент часу t відповідно; $\dot{\mathbf{e}}(t)$, $\dot{\mathbf{g}}(t)$ – швидкості змін потужностей $\mathbf{e}(t)$ та $\mathbf{g}(t)$; $\mathbf{p}(t)$ – вектор інтенсивності невиробничого споживання; $\mathbf{z}(t)$ – вектор показників

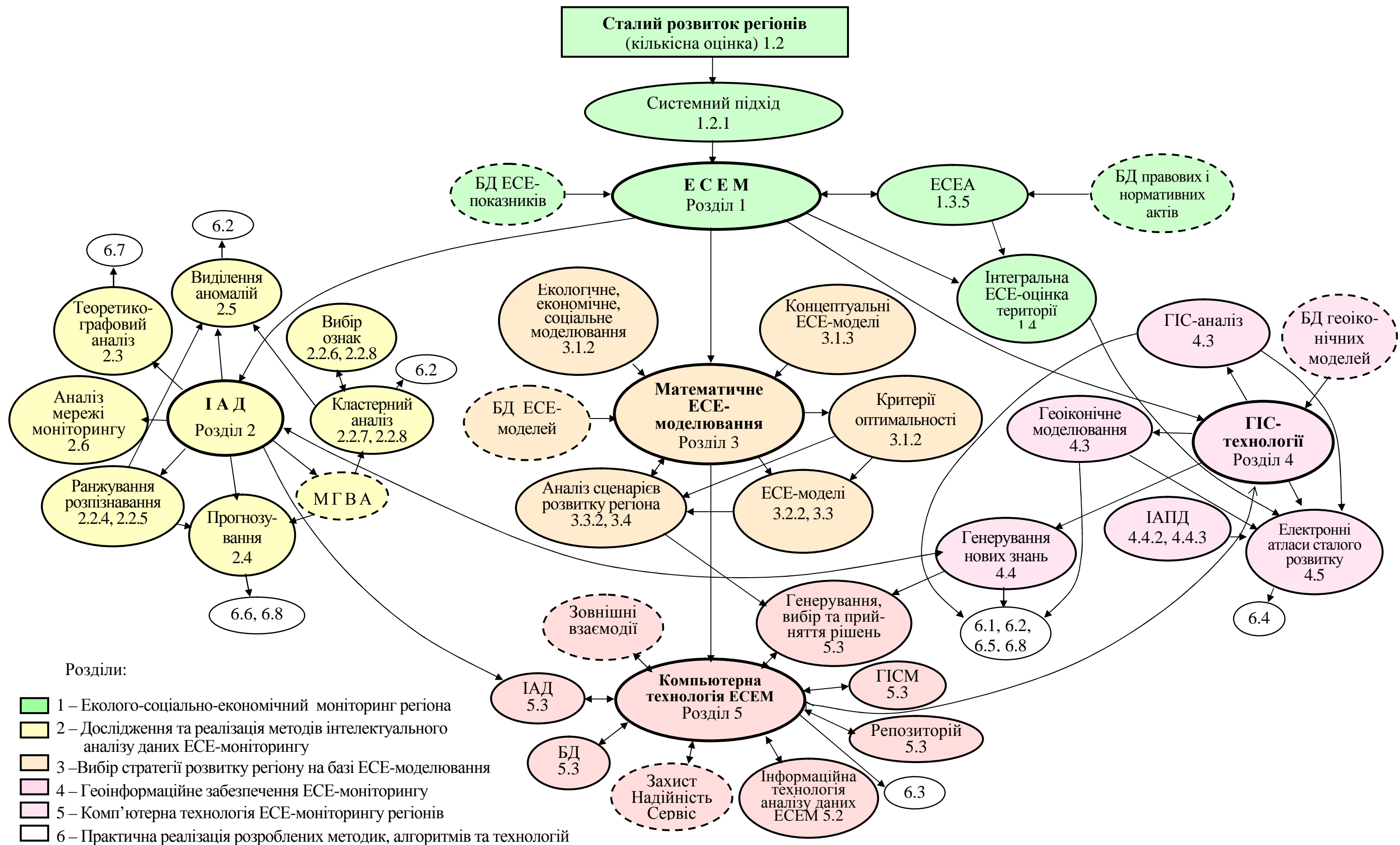


Рис.1. Структурна діаграма логічних взаємозв'язків основних розділів дисертації

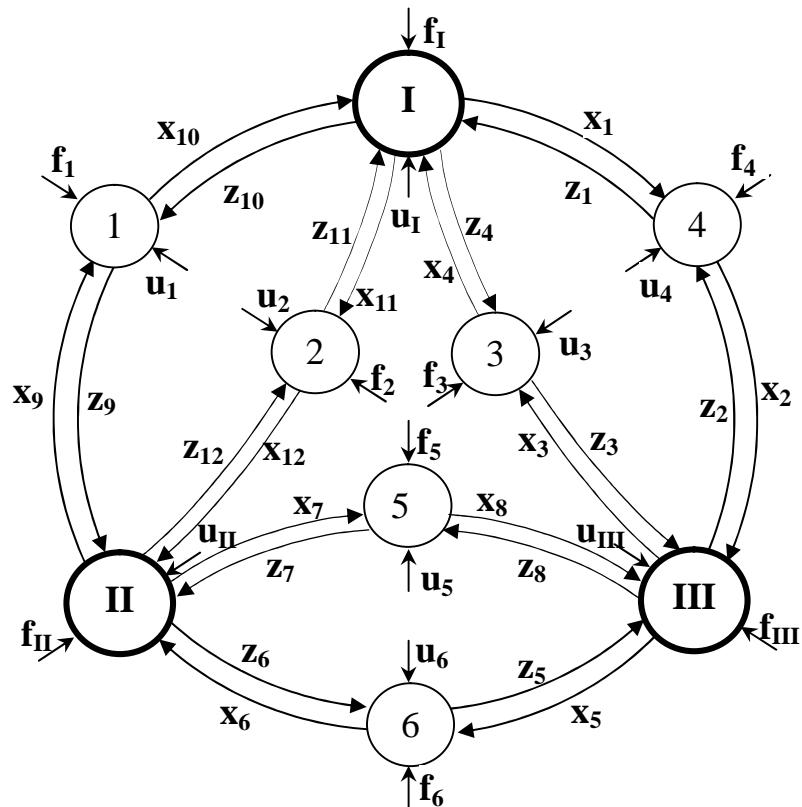


Рис. 3. Структурна циркуляційна схема ЕСЕ-системи регіону:

I – економіка, **II** – екологія, **III** – соціум; 1 – видобуток ресурсів і сільське господарство, 2 – забруднення, 3 – трудові ресурси, 4 – кінцевий продукт, 5 – якість довкілля, 6 – відходи життєдіяльності;

x_1 – обсяги виробництва товарів народного споживання, x_2 – обсяги придбаних товарів, x_3 – пропоновані трудові ресурси, x_4 – обсяги придбаних трудових ресурсів, x_5 – обсяги побутових відходів, x_6 – обсяги щорічних викидів забруднювачів, супутніх соціальним процесам, x_7 – показники якості навколишнього середовища, x_8 – показники якості довкілля, що впливають на життєдіяльність та здоров'я людей, x_9 – показники відтворюваних природних ресурсів, x_{10} – обсяги щорічного видобутку відтворюваних і невідтворюваних ресурсів, x_{11} – галузеві обсяги виробництва, x_{12} – обсяги щорічних викидів забруднювачів, супутніх виробничим процесам;

z_1 – кошти за продані товари, z_2 – кошти за куплені товари, z_3 – кошти за продані трудові ресурси, z_4 – витрати на придбання ресурсів, z_5 – витрати на перетворення відходів в екологічно чисті речовини, z_6 – витрати на захист від відходів-забруднювачів, z_7 – витрати на підвищення якості довкілля та відновлення природних ресурсів, z_8 – витрати на природоохоронні заходи, z_9 – витрати на відновлення природних ресурсів, z_{10} – витрати на набування невідтворюваних і відтворюваних ресурсів, z_{11} – показники економічних збитків від забруднення навколишнього середовища, z_{12} – витрати на захист від забруднювачів;

f_1, f_2, \dots, f_{12} – зовнішні впливи (централізовані капіталовкладення, атмосферні впливи і т.д.);

u_1, u_2, \dots, u_{12} – управляючі впливи.

природного середовища; \mathbf{z}_0 – незбурений стан \mathbf{z} ; \mathbf{v} – вектор, компоненти v_i якого показують зміну показника z_i під впливом одиниці населення; $\mathbf{N}(t)$ – населення в момент t ; $\mathbf{A}=(a_{ij})_1^n$, $\mathbf{B}=(b_{ij})_1^n$, $\mathbf{C}=(c_{ij})_{i,j}^{n,m}$ ($a_{ij} \geq 0$, $b_{ij} \geq 0$, $c_{ij} \geq 0$) матриці коефіцієнтів прямих витрат продукції i -ї галузі на виробництво одиниці продукції j -ї галузі, фондотвірних витрат продукції i -ї галузі на одиничний приріст потужності j -ї галузі, витрат продукції i -ї галузі на одиничну інтенсивність відновлення j -го ресурсу відповідно; $\mathbf{D}=(d_{ij})_{i,j}^{n,m}$, $d_{ij} \geq 0$, d_{ij} – коефіцієнт витрат продукції i -ї галузі на одиничний приріст потужності відновлення j -го ресурсу; $\mathbf{Q}=(q_{ij})_1^m$, q_{ij} – коефіцієнт взаємного впливу показників природного середовища z_i і z_j ; $\mathbf{R}=(r_{ij})_{i,j}^{m,n}$, $\mathbf{H}=(h_{ij})_{i,j}^{m,n}$ – матриці коефіцієнтів питомих ресурсних витрат при випуску продукції і розвитку потужностей галузей виробництва; $\mathbf{G}=(g_{ij})_1^m$, $g_{ij}=0$ при $i \neq j$, $g_{ii}=1$, якщо відновлення i -го ресурсу сприяє збільшенню показника z_i , у противному разі $g_{ii}=-1$; $\mathbf{S}=(s_{ij})_1^m$, s_{ij} – коефіцієнт, що показує зміну показника z_i при одиничному збільшенні потужності відновлення j -го ресурсу; $\mathbf{L}=(l_{ij})_{i,j}^{n,m}$, коефіцієнт l_{ij} показує зміну i -го ресурсу, обумовлену одиничним невиробничим споживанням продукції j -ї галузі.

Основні змінні в моделі – $\mathbf{x}(t)$, $\mathbf{z}(t)$. Різні сценарії розвитку регіону визначаються заданням функцій $\mathbf{y}(t)$, $\dot{\mathbf{e}}(t)$, $\dot{\mathbf{g}}(t)$. Для ОПР зручніше оперувати фінансовими ресурсами, тому $\mathbf{y}(t)$, $\dot{\mathbf{e}}(t)$, $\dot{\mathbf{g}}(t)$ виражено в роботі функціями від $\Phi(t)$, де $\Phi(t)$ – сумарні інвестиції (у вартісному вигляді) в розвиток регіону. Припускаючи, що витрати на зростання швидкостей потужностей $\dot{g}_i(t)$ дорівнюють величині інвестицій $\Phi_{ii}(t)$ на підвищення потужностей по галузях виробництва, витрати на збільшення інтенсивності $y_i(t)$ відновлення i -го ресурсу – величині інвестицій $\Phi_{Pi}(t)$ на відновлення власно ресурсів, затрати на $\dot{e}_i(t)$ – величині інвестицій $\Phi_{Ci}(t)$, $i=1,2,\dots,m$, системи рівнянь для визначення функцій $\mathbf{y}(t)$, $\dot{\mathbf{e}}(t)$, $\dot{\mathbf{g}}(t)$ подано у вигляді

$$\mathbf{y}(t)=\mathbf{F}_y(t)\Phi(t), \quad \dot{\mathbf{g}}(t)=\mathbf{F}_g(t)\Phi(t), \quad \dot{\mathbf{e}}(t)=\mathbf{F}_e(t)\Phi(t) \quad (18)$$

де $\mathbf{F}_y(t)$, $\mathbf{F}_g(t)$, $\mathbf{F}_e(t)$ – вектори коефіцієнтів, а з (15) – (17) отримано систему рівнянь:

$$\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}+(\mathbf{B}\mathbf{F}_g+\mathbf{C}\mathbf{F}_y+\mathbf{D}\mathbf{F}_e)\Phi+\mathbf{p}, \quad (19)$$

$$\dot{\mathbf{z}}=\mathbf{Q}(\mathbf{z}-\mathbf{z}_0)-\mathbf{R}\mathbf{x}+(\mathbf{G}\mathbf{F}_y-\mathbf{H}\mathbf{F}_g-\mathbf{S}\mathbf{F}_e)\Phi-\mathbf{v}\mathbf{N}-\mathbf{L}\mathbf{p}, \quad (20)$$

$$\mathbf{0} \leq \mathbf{x} \leq \mathbf{g}, \quad \mathbf{0} \leq \mathbf{y} \leq \mathbf{e}, \quad \mathbf{z}_{\min} \leq \mathbf{z} \leq \mathbf{z}_{\max}. \quad (21)$$

Змінюючи частки інвестицій (на зростання потужностей виробництва, переробку відходів, відтворення власне ресурсів і т.д.), а, отже, і розподіл асигнованих фінансувань $\Phi(t)$ на підвищення потужностей галузей виробництва та відтворення ресурсів, можна вибирати кращі щодо раціонального природокористування сценарії розвитку регіону. Можна робити аналіз для передбачуваних гіпотетичних сценаріїв: а) відновлення ресурсів відбувається на повну потужність пропорційно основним фондам; б) відсутнє зростання потужностей економіки та вкладань у відновлювані потужності; в) задано постійний темп зростання потужностей економіки і т.д.

Процес прийняття рішень при виборі стратегії розвитку регіону на підставі ЕСЕ-моделювання слід розглядати як ітераційну процедуру, що включає три етапи: 1) пошук функціональних закономірностей між ЕСЕ-показниками; 2) побудова ЕСЕ-моделі та дерева припустимих сценаріїв розвитку регіону; 3) пошук на дереві припустимих сценаріїв прийняттого з погляду ОПР сценарію в режимі діалогу. Процедура вибору стратегії розвитку регіону на базі ЕСЕ-моделювання дає можливість ОПР робити пошук стратегій, націлених на траєкторію сталого розвитку регіону.

У **розділі 4** досліджено геоінформаційне забезпечення ЕСЕМ. Регіони в ГІС – це просторово координовані об’єкти – полігони, а ЕСЕ-показники – їх атрибути-індикатори, які зберігаються в атрибутивних таблицях, пов’язаних з відповідними об’єктами. Схематичне подання вихідних ЕСЕМ-даних в ГІС наведено на рис. 4.

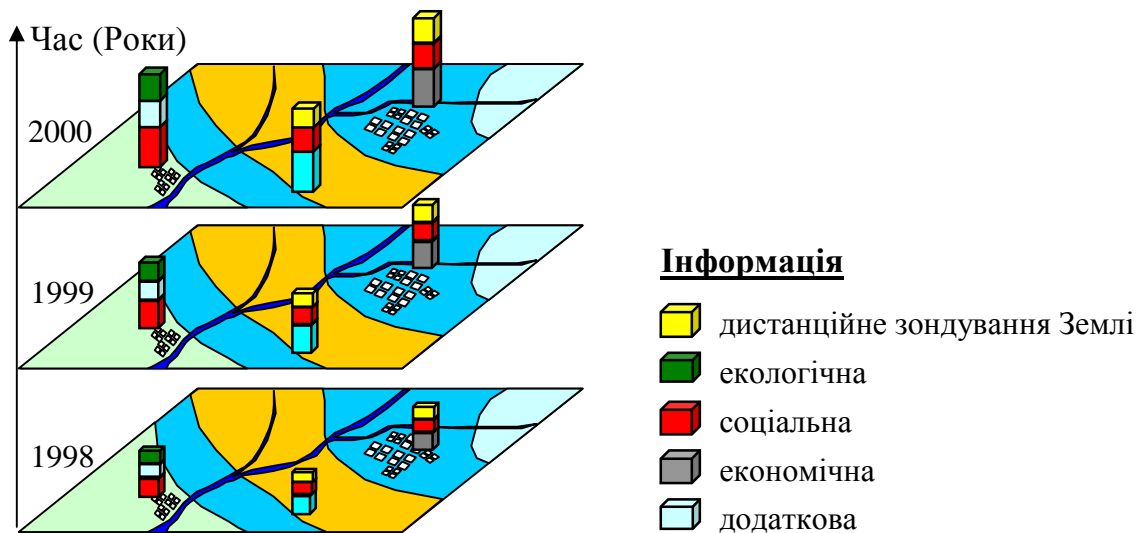


Рис. 4. Схематичне зображення вихідних ЕСЕМ-даних у ГІС

Типова для ЕСЕМ-задач структура даних включає географічний район і декілька його розбивань на однорідні інформаційні шари: 1) полігони (адміністративні одиниці, зони екологічного районування і т.д.); 2) лінійні об’єкти (ріки, залізниці тощо); 3) точкові об’єкти (міста, вузли мережі спостережень і т.д.); 4) цифрові моделі растрових полів (щільність населення і т.д.); 5) зображення (аерокосмічні знімки та ін.). Шари можна комбінувати, накладаючи один на один і створюючи нові.

У ГІС атрибути регіонів можна поділити на три групи: 1) атрибути картографічного подання (набір констант і алгоритмів картографічного відображення даних, таких як вибір палітри та піктограм, кольору і методу фарбування, обчислення розміру піктограм і т.д.); 2) індикатори (значення ЕСЕ-показників); 3) документи (гіпертекст, зображення, звук).

У розділі розглянуто особливості застосування ГІС для задач ЕСЕМ, методи і засоби ГІС-аналізу просторово розподіленої ЕСЕ-інформації. Розроблено інформаційну технологію геоіконічного моделювання ЕСЕМ-даних.

Геоіконічні ЕСЕ-моделі – новий клас просторових геоінформаційних ЕСЕ-моделей, що ґрунтується на іконічному (візуальному, образному) способі відображення ЕСЕ-інформації, упорядкованої в ГІС за допомогою географічної карти. Карта в геоіконічних моделях – носій не тільки наочної інформації, а й ефективний засіб аналізу. Процес створення карт і географічний аналіз не є чимось принципово новим. Однак ГІС-технології надають новий ефективний спосіб роботи з просто-риво координованою інформацією, що включає ГІС-аналіз об'єктів.

Основна ідея геоіконічного моделювання показників ЕСЕ-моніторингу регіонів полягає в інтегрованому аналізі даних екологічного та соціально-економічного моніторингу на основі ГІС-технологій, у системному підході до інтерпретації ЕСЕ-процесів, що відбуваються у регіонах.

При розв'язанні задач збалансованості ЕСЕ-розвитку регіонів автор пропонує використовувати геоіконічні моделі:

- 1) *рангові картограми* (застосовують схеми класифікації за одним атрибутом);
- 2) *картодіаграми* (показують розподіл якого-небудь явища за допомогою різних типів діаграм, локалізованих по території);
- 3) *аналітичні карти* (для неузагальнених показників);
- 4) *комплексні карти* (одночасно враховують декілька різних взаємопов'язаних явищ – кожне в своїй системі показників);
- 5) *синтетичні карти* (на основі інтеграції множини показників або аналітичних і комплексних карт показують результати розпізнавання, кластеризації та ін., отримані методами ІАД);
- 6) *геоіконічні моделі регіональної генералізації* (будують після порівняльного аналізу різних картограм і пошуку просторових аналогів);
- 7) *інтегральні регіональні моделі* (суперпозиція карт ЕСЕ-показників заданого тематичного сюжету; інтегруванню в ГІС відповідає «логічне додавання»);
- 8) *динамічні геоіконічні моделі* (з урахуванням часових змін показників моніторингу) та *анаморфози* або топологічні картограми (графічні зображення, похідні від традиційної карти, масштаб яких трансформується і варіюється залежно від величини регіонального ЕСЕ-показника).

Автор не лише класифікує геоіконічні моделі, а й робить їх дієвими, вводить для практичного застосування на основі дослідження ГІС-технологій для задач ЕСЕ-моніторингу регіонів, аналізу традиційних екологічних і соціально-економічних карт, а також нових власних розробок. Стандартні ГІС поки що не містять багатьох необхідних для геоіконічного моделювання ЕСЕ-показників процедур – кластеризації, розпізнавання, прогнозування, обчислення інтегральної оцінки і т.д.

Розроблено технології побудови окремих геоіконічних моделей, що базуються на різних схемах класифікації, районуванні за допомогою інтегральної оцінки (див. розд. 1), відображенні результатів ІАД (див. розд. 2) – ранжування, розпізнавання, кластеризації, прогнозування, виділення аномалій.

При моделюванні часової динаміки ЕСЕ-показників рекомендовано одночасне врахування поточного значення самого ЕСЕ-показника та його першої і другої похідних за часом, що дозволяє будувати інтегральну модель, яка враховує не тільки величину показника, а й тенденцію його зміни.

У розділі розроблено схеми класифікації в ГІС за двома та трьома атрибутами (природне розбиття, рівні інтервали, квантиль, стандартне відхилення), аналогічні існуючим схемам за одним атрибутом. Викладено технологію створення електронних атласів сталого розвитку.

Обґрунтовано процедури отримання нових знань на основі ГІС-аналізу, ІАД та геоіконічного моделювання: 1) регіональна генералізація границь на основі рангових картограм, оверлею полігонів; 2) виявлення закономірностей територіального розподілу багатовимірних аномалій; 3) комплексний аналіз ЕСЕ-розвитку регіонів на базі синтезу та агрегування в ГІС результатів статистичної звітності; 4) виявлення наявності взаємозв'язків між різнорідними економічними, екологічними і соціальними показниками за допомогою порівняльного аналізу рангових картограм; 5) співставлення регіональних багатовимірних регресійних моделей залежностей між ЕСЕ-показниками та результатів геоіконічного моделювання залишків регресійних моделей для знаходження пояснення залишкам і уточнення регресійної моделі; 6) виявлення тенденцій динаміки зміни регіональних ЕСЕ-показників.

Інформаційно насичені інтерактивні геоіконічні моделі надають зручний інструментарій для ОПР при аналізі регіональних проблем.

Розділ 5 присвячено викладу методики просторово-часового аналізу ЕСЕМ-даних і відповідної комп'ютерної інформаційної технології.

Концептуальна схема зазначеної методики, наведена на рис. 5, охоплює п'ять основних блоків: 1) геоінформаційний аналіз та геоіконічне моделювання; 2) інтелектуальний аналіз даних; 3) генерування нових знань; 4) аналіз і прийняття рішень (вибір стратегії розвитку регіону на базі ЕСЕ-моделювання); 5) діалог ОПР–ЕОМ.

Остаточний вибір технологічного ланцюжка методів залежить від конкретних даних ЕСЕ-моніторингу, цільових настанов і досвіду дослідника, але при цьому він може скористатись інформаційною технологією, що відповідає даній методиці, взявши за основу готові блоки ланок технологічного ланцюжка.

На першому етапі дослідження виконують ГІС-аналіз і геоіконічне моделювання вихідних ЕСЕ-показників. Алгоритм застосування методів при цьому такий.

1. Будуються стандартні географічні (містять інформацію про границі регіонів), аналітичні та комплексні карти ЕСЕ-показників.

2. Аналізуються рангові картограми ЕСЕ-показників. Для моніторингу часових змін ЕСЕ-показника будується набір рангових картограм. Досліджується, які схеми класифікації при цьому краще використовувати (квантиль, рівні інтервали і т.д.). Ураховується, що діапазони зміни градацій показника повинні бути співставними. Здійснюються порівняльний аналіз картограм і пошук просторових аналогів, виділяються регіональні зони.

3. Моніторинг тимчасових змін одного ЕСЕ-показника відображається за допомогою картодіаграм.

4. Створюються геоіконічні моделі, одержувані шляхом порівняльного аналізу різних картограм і пошуку просторових аналогів. При цьому використовуються результати попередніх кроків (виділені регіональні зони).



Рис. 5. Концептуальна схема просторово-часового аналізу даних ЕСЕ-моніторингу

5. Будуються синтетичні карти (показують швидкість, прискорення та поточне значення показника).

6. Інформація, що отримана на попередніх кроках і потребує спільного розгляду, подається в інтегральних геоіконічних моделях. Виділяються аномальні за сукупністю ЕСЕ-показників регіони та ін.

7. На основі часових рядів значень ЕСЕ-показників будуються динамічні карти, а за потреби – анаморфози.

Другий етап передбачає ІАД. Основні кроки другого етапу.

1. Кореляційний аналіз ЕСЕ-показників.
2. Моделювання залежностей між досліджуваними ЕСЕ-показниками з використанням розробленого в розділі 2 алгоритму; виявлення просторової автокореляції та автокореляції в регресійних залишках.
3. Кластеризація регіонів за сукупністю ЕСЕ-показників; кластеризація регіонів за залишками декількох задовільних регресійних моделей; перегляд результатів у ГІС, створення синтетичної карти.
4. Ранжування регіонів за мірою близькості до заданого регіону або сукупності регіонів; задання еталонних класів і розпізнавання з учителем; подання і перегляд отриманих результатів у ГІС.

5. Виявлення просторово-часових аномалій.
6. Прогнозування значень ЕСЕ-показників.

На третьому етапі виконується синтез знань, отриманих на перших двох етапах. Основні кроки третього етапу.

1. Генералізація меж регіональних зон на основі аналізу карт кластеризації та рангових картограм.

2. Генерування гіпотез про зв'язок явищ.

3. Прогноз значень ЕСЕ-показників регіонів (на основі результатів шостого кроку другого етапу і перших двох кроків третього етапу).

4. Інтегральний регіональний аналіз збалансованості розвитку та визначення "вузьких" місць.

5. Верифікація даних.

У ході четвертого етапу – аналіз та прийняття управлінських рішень. Процес прийняття ОПР рішень треба розглядати як ітераційну процедуру в діалозі ОПР–ЕОМ, коли ОПР аналізує результати попередніх етапів, задає певні установки для уточнення рішення, оцінює ситуацію в регіонах.

На рис. 6 подано окремі приклади аналізу тематичного сюжету "Народжуваність і смертність населення, екологічна обстановка регіонів України, 1997-2000".

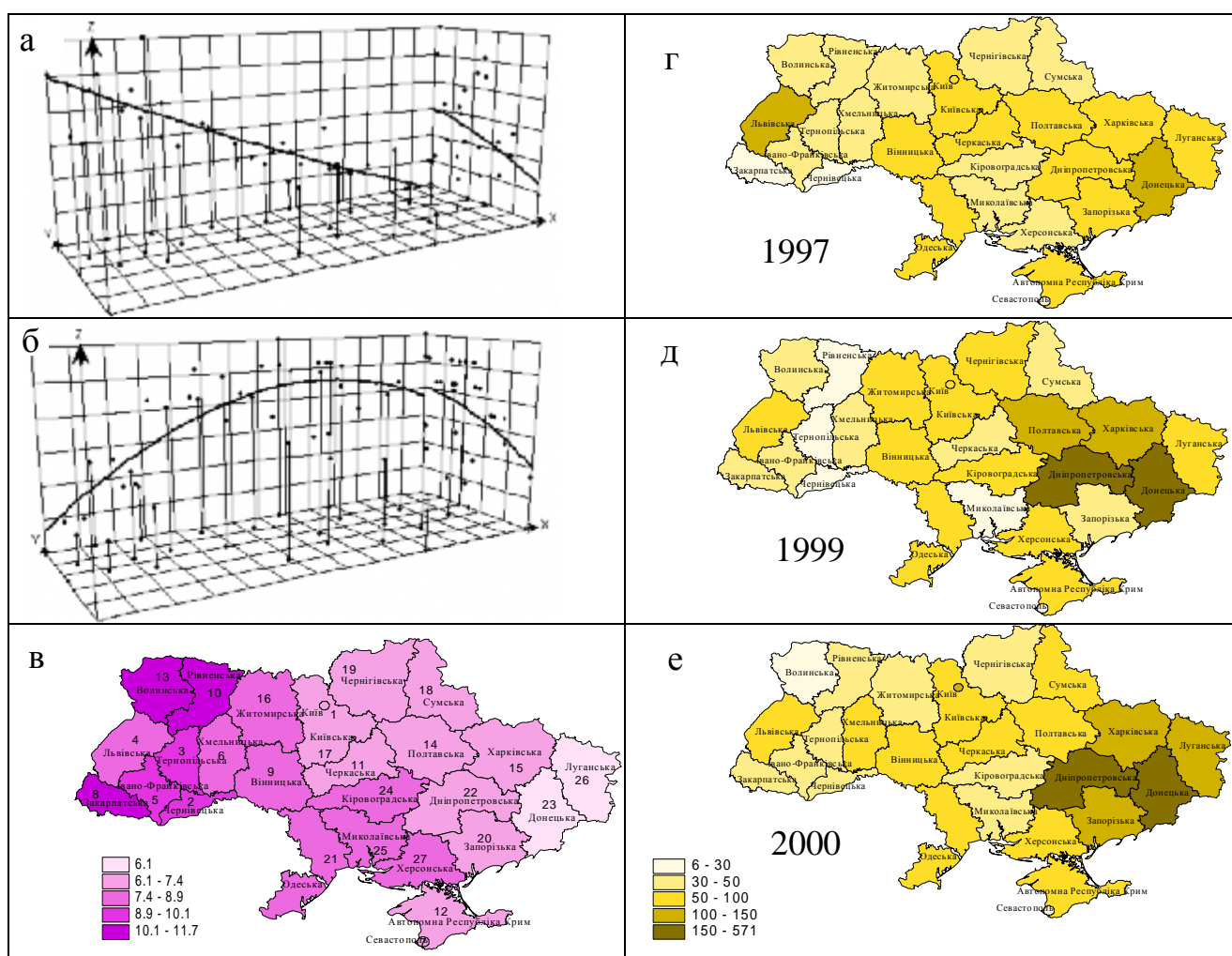


Рис. 6. Фрагменти просторового (а, б, в) та часового (г, д, е) аналізу ЕСЕМ-даних

Простежуються тренди, присутні у розподілі показників «Народжуваність (рис. 6 а) і смертність (рис. 6 б) населення регіонів України в 1997 р.». Кожна вертикальна лінія на графіках – це значення показника в місці розташування регіону (початок координат – на заході). На перпендикулярних (до XOY) площинах видно діаграми розсіяння та лінії регресії: народжуваність населення зменшується, а смертність зростає в напрямку із заходу на схід. Рангова картограма показника "Кількість народжених (загалом, на 1000 населення)" з індексом тривалості життя за 2000 р. (рис. 6 в) виявляє таку ж тенденцію показника народжуваності. Динаміку зміни показника "Викиди шкідливих речовин в атмосферу, тис.т" відображено на рис. 6 г–е.

У розділі вказано призначення, мету і принципи створення комп'ютерної системи ЕСЕ-моніторингу регіонів, вимоги до системи в цілому і до окремих її підсистем, розроблено загальну структурно-функціональну схему (рис. 7). Проаналізовано характеристики системи та функції її підсистем.

У **розділі 6** викладено результати практичної реалізації розроблених методів, алгоритмів та технологій. Розглянуто задачі ГІС-аналізу ЕСЕ-показників моніторингу регіонів України, районування, виділення просторово-часових аномалій регіонального розвитку за сукупністю ЕСЕ-показників, розробки фрагментів ГІС "Охорона здоров'я" м. Дніпропетровська, створення електронного атласу сталого розвитку Дніпропетровської області, кількісної оцінки показника грошових витрат на експлуатацію виробки, геоіконічного моделювання і прогнозування за реальними даними ЕСЕ-моніторингу України.

У задачі короткострокового прогнозування вхідними даними є значення 8-ми ЕСЕ-показників (на душу населення, у фактичних цінах) для регіонів України за 6 років (1996-2001): y – валова додана вартість; x_1 – виробництво промислової продукції; x_2 – інвестиції в основний капітал (за рахунок коштів держбюджету); x_3 – інвестиції в основний капітал у порівняльних цінах 1996 р.; x_4 – прямі іноземні інвестиції; x_5 – середньомісячна заробітна плата працівників; x_6 – місцеві бюджети; x_7 – виробництво товарів народного споживання.

На першому етапі виконано розвідувальний аналіз (коефіцієнтів кореляції, головних компонент і кластерний). Значення попарних коефіцієнтів кореляції окремого показника за 6 років достатньо високі (0,9), тому в матриці коефіцієнтів виділяється деяка квадратно-прямокутна структура, а це дає підставу вважати, що різьких поворотів траєкторії розвитку регіонів у 1996-2001 рр. не спостерігалось. Головні компоненти деяких регіонів утворюють окремі, віддалені від загальної сукупності точок, «викиди». Цей же висновок підтверджують результати кластеризації, що проведено для різної кількості показників. Такі регіони (Дніпропетровський, Донецький, Запорізький, Полтавський, Київ, Севастополь) потребують окремого розгляду.

Застосування розробленого алгоритму структурної ідентифікації обумовлено особливостями задачі: 1) відсутністю апріорної інформації про структуру моделі та дисперсію похибок спостережень; 2) співвідношенням числа ($m=7$) вихідних змінних і числа ($N=21 \times 6=121$) спостережень. Розрахунки зроблено для трьох виборок даних – А (навчальна, 1996-1998), В (перевірна, 1999-2000), С (контрольна, 2001), двох видів структури моделі (лінійна, нелінійна), трьох критеріїв оцінки якості моделей: регулярності $R(s)$, остаточної суми квадратів $J(s)$, ковзного екзамену $U(s)$.

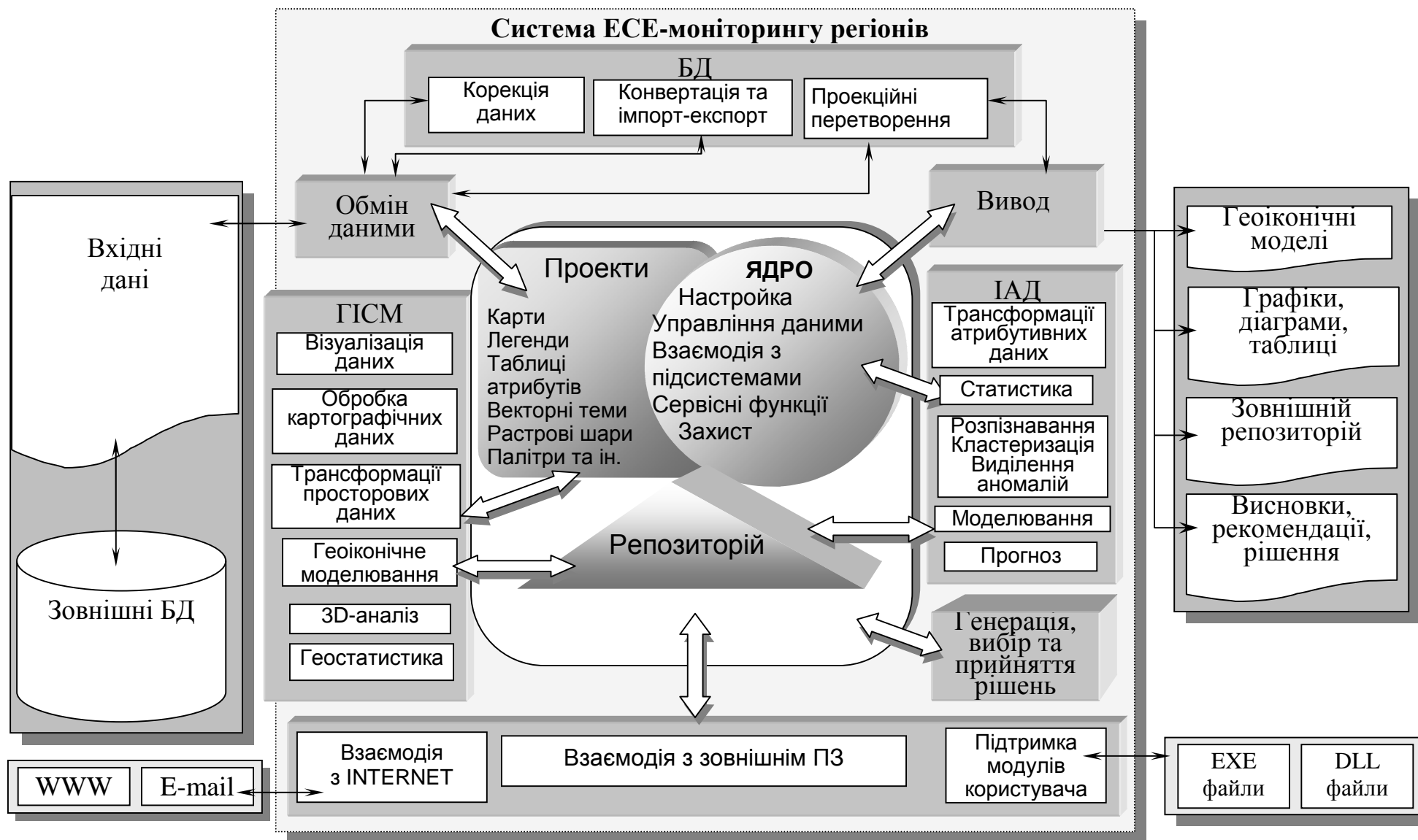


Рис.7. Структурно-функціональна схема системи ЕСЕ-моніторингу регіонів

Якість отриманих моделей оцінювалася середньоквадратичним відхиленням модельних значень \hat{y} від спостережуваних і коефіцієнтом множинної кореляції $R_{y,x}$. Для моделей, у яких $R_{y,x} > 0.7$, в таблиці подано змінні x_i , що ввійшли до відповідних рівнянь регресії. Як випливає з таблиці, якісна структура шуканої залежності має вигляд $y=f(x_1, x_5, x_7)$. Якщо значенням критеріїв на контрольній вибірці надавати найбільшої ваги, то найкращі моделі:

$$y=0,18x_1+8,04x_5+0,55x_7, \quad (22)$$

$$y=0,144x_1+0,754x_5-0,052x_5x_6+0,26x_7. \quad (23)$$

Значення критеріїв, коефіцієнти множинної кореляції та номери ознак у кращих моделях

Структура моделі	Критерій		
	J(s)	G(s)	U(s)
Лінійна	J(3)=137 $R_A=0.95, R_C=0.72$ x_1, x_5, x_7	G(3)=183 $R_A=0.9, R_B=0.92$ $R_{AB}=0.95, R_C=0.75$ x_1, x_5, x_7	U(3)=156 $R_A=0.95$ $R_C=0.79$ x_1, x_5, x_7
Нелінійна	J(4)=0.286 $R_A=0.96, R_C=0.78$ x_1, x_5, x_6, x_7	G(3)=0.364 $R_A=0.89, R_B=0.93$ $R_{AB}=0.96, R_C=0.78$ x_5, x_6, x_7	U(5)=0.266 $R_A=0.96$ $R_C=0.75$ x_1, x_4, x_5, x_6, x_7

На рис. 8 показано динаміку зростання показника «Валова додана вартість» (1996-2001) з прогностичними значеннями 2002 р. – за (22).

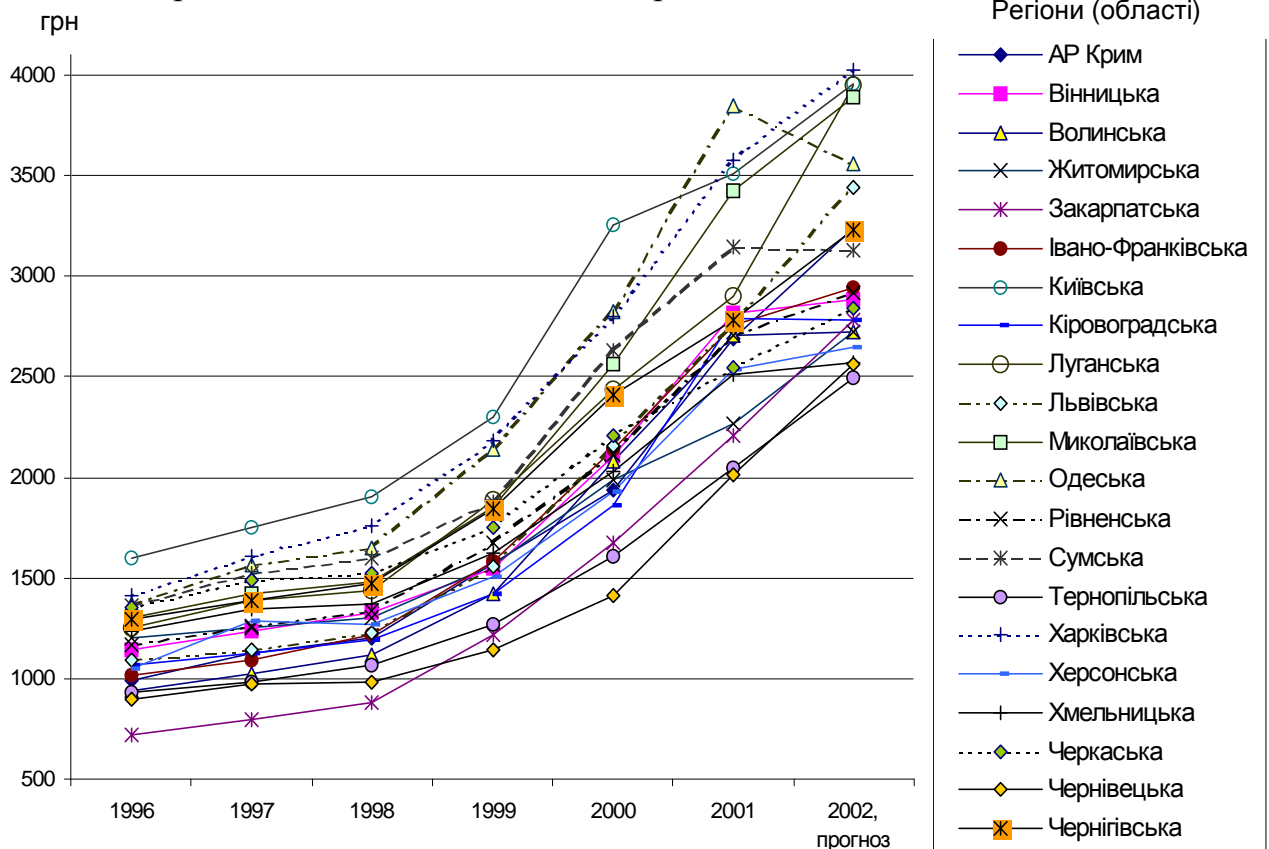


Рис. 8 . Динаміка зростання показника «Валова додана вартість» (1996-2001) з прогностичними значеннями 2002 р.

Отримані моделі дають інформацію для роздумів ОПР: «Чому до структури кращих моделей не ввійшли показники інвестицій x_2 , x_4 ?»

Результати досліджень використано у Міністерстві економіки та з питань європейської інтеграції України і в Державному науково-дослідному інституті інформатизації та моделювання економіки.

ВИСНОВКИ

Дисертаційна робота – закінчене наукове дослідження, результатом якого є вирішення науково-технічної проблеми математичного забезпечення і комп'ютерної технології ЄСЕ-моніторингу регіонів. Проведені дослідження сприяють видобуванню знань з різномірних даних ЄСЕ-моніторингу для оцінки сталості та збалансованості регіонального розвитку. Основні наукові результати, висновки та рекомендації полягають у нижчезазначеному.

1. Запропонована концепція ЄСЕ-моніторингу як складової системи управління регіональним розвитком визначає його місце в регіональному управлінні, види та напрямки, структуру математичного забезпечення, проблеми і перспективи, необхідність упровадження ЄСЕ-моніторингу для спостереження і прогнозування ходу найважливіших ЄСЕ-процесів з метою їхнього аналізу, ідентифікації та виявлення кола факторів, що регулюються при підготовці і прийнятті рішень.

2. Розроблена методика інтегральної ЄСЕ-оцінки території для ЄСЕ-аудиту дає змогу кількісно оцінювати рівень сталого розвитку регіонів на основі моніторингових даних трьох інформаційних блоків – екологічного, соціального й економічного, в кожному з яких використано три типи показників: фактори забруднення, показники стану і керуючі впливи.

3. Запропонований єдиний системний підхід до класифікації методів кластеризації та методів вибору ознак у розпізнаванні образів, в якому як кластеризація, так і вибір ознак – задачі дискретної оптимізації на вершинах ґратки розбивань, дозволяє використовувати існуючі алгоритми перебору вершин у виборі ознак для пошуку найкращої кластеризації.

4. Розроблене методичне та алгоритмічне забезпечення об'єктивної кластеризації та прогнозування сприяє підвищенню ефективності (порівняно з відомими методами) інтегрованого аналізу та прогнозу ЄСЕ-показників, що підтверджено експериментальними дослідженнями на матеріалах моніторингу ЄСЕ-показників регіонів України за 1985-2000 рр.

5. Досліджено методи та алгоритми виділення аномалій регіонального розвитку, що охоплюють різні підходи (на основі застосування кластерного аналізу, поляризації, аналізу регресійних залишків, адаптивності та різномірності; мір близькості між об'єктом і класом) і дозволяють виявляти аномалії тенденцій часового розвитку та аномалії в межах заданого тематичного сюжету.

6. Класифіковано методи оптимізації мережі моніторингу (геометричні, статистичні, інтерполяційні, кластерні, графові, геостатистичні, фрактальні), використання яких доцільне в ЄСЕ-моніторингу регіонів. Подано рекомендації щодо застосування цих методів для аналізу та оптимізації мережі екологічного і соціально-економічного моніторингу.

7. Запропоновано методологію вибору стратегії розвитку регіону на базі ЕСЕ-моделювання. Розглянуто балансову ЕСЕ-модель регіону першого наближення, зручну для моніторингових досліджень. Рівняння моделі перетворено з урахуванням структури інвестицій, що надає можливість аналізувати міжгалузеві пропорції та робити різні планові розрахунки. Викладено питання визначення ЕСЕ-показників за заданих гіпотетичних сценаріїв розвитку регіону.

8. Вперше розроблено методологію використання геоіконічного моделювання в середовищі ГІС для ЕСЕ-моніторингу:

а) класифіковано геоіконічні моделі ЕСЕ-показників (рангові картограми; картодіаграми; аналітичні карти; комплексні карти; синтетичні карти; геоіконічні моделі регіональної генералізації; інтегральні регіональні геоіконічні моделі; динамічні геоіконічні моделі);

б) запропоновано технології побудови окремих типів моделей, заснованих на різних схемах класифікації, районуванні за допомогою інтегральної оцінки території, картуванні результатів ранжування, розпізнавання, кластеризації, прогнозування, виділення аномалій регіонального розвитку.

9. Доведено, що ГІС-технології, методи ГІС-аналізу і геоіконічного моделювання дозволяють генерувати нові дані, що не враховуються офіційною статистикою (генералізація границь на основі рангових картограм, картодіаграм, аналізу просторового розподілу регресійних залишків тощо), і за рахунок цього доповнити статистичні дані.

10. Розроблено методику та відповідну комп'ютерну інформаційну технологію просторово-часового аналізу показників ЕСЕ-моніторингу, що включає процедури ГІС-аналізу і геоіконічного моделювання, інтелектуальний аналіз даних, генерування нових знань, аналіз і прийняття рішень. Перевага даної технології – у спрямуванні на оперативне видобування знань з різномірних моніторингових даних.

11. Викладено методологію побудови комп'ютерної системи ЕСЕ-моніторингу: визначено призначення, мету і принципи створення, вимоги до системи в цілому та до окремих її підсистем; розроблено структурно-функціональну схему, проаналізовані характеристики системи ЕСЕ-моніторингу та функції її підсистем.

12. Побудовано геоіконічні моделі для аналізу ЕСЕ-показників моніторингу регіонів України та Дніпропетровської області, що відбивають динаміку, незбалансованість регіонального розвитку і дозволяють ОПР інтерпретувати перебіг ЕСЕ-процесів у регіонах.

Подальші дослідження в даному напрямку передбачають розробку методів і алгоритмів інтегрованого аналізу ЕСЕМ-даних з використанням інтерактивних можливостей ГІС майбутніх поколінь.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ

1. Сарычева Л.В. Компьютерный эколого-социально-экономический мониторинг регионов. Математическое обеспечение: Монография. – Днепропетровск: НГУ, 2003. – 222 с.

2. Сарычева Л.В. Компьютерный эколого-социально-экономический мониторинг регионов. Геоинформационное обеспечение: Монография. – Днепропетровск: НГУ, 2003. – 174 с.
3. Сарычева Л.В. Разработка и применение фрагментов геоинформационной системы «Охрана здоровья» города Днепропетровска // Сб. науч. тр. НГУ, №14, Т.2. – Днепропетровск: НГУ, 2002. – С.124-136.
4. Сарычева Л.В. Информационная технология пространственно-временного анализа данных мониторинга регионов // Наук. вісник НГУ, 2004.–№2.–С.72-76.
5. Сарычева Л.В. Геоиконическое моделирование показателей устойчивого развития регионов // Актуальні проблеми автоматизації та інформаційних технологій. – Т.7. Зб. наук. пр. – Дніпропетровськ: Навч. книга, 2003. – С.116-133.
6. Сарычева Л.В. Единый подход к классификации методов кластеризации и методов выбора информативных признаков // Актуальні проблеми автоматизації та інформаційних технологій. – Т.6. Зб. наук. пр. – Дніпропетровськ: Навч. книга, 2002. – С.137-144.
7. Сарычева Л.В. Виртуальное картографирование показателей эколого-социально-экономического мониторинга в растровой ГИС // Сб. науч. тр. НГА Украины. – №10. – Днепропетровск: НГА Украины, 2000. – С.98-103.
8. Сарычева Л.В. Геоинформационный анализ и эколого-социально-экономический мониторинг // Сб. науч. тр. НГА Украины №11, Т.1. – Днепропетровск: НГА Украины, 2001. – С.129-138.
9. Сарычева Л.В. Теоретико-графовый анализ структуры пространственно привязанных данных // Сб. науч. тр. НГА Украины.– №12, Т.1. – Днепропетровск: НГА Украины, 2001. – С.117-127.
10. Sarycheva L.V. Using GMDH for Modeling Economical Indices of Mine Opening // Systems Analysis and Modeling Simulation (SAMS).– 2003, Vol.43, №10, P.1341-1350.
11. Сарычева Л.В. Эколого-социально-экономический мониторинг: выделение аномалий регионального развития // Сб. науч. тр. НГУ.– №18, Т.2. – Днепропетровск: НГУ, 2003. – С.124-136.
12. Сарычева Л.В. Прогнозирование в задачах эколого-социально-экономического-мониторинга регионов // Моделювання та інформатизація соціально-економічного розвитку України. Зб. наук. пр.– К.: ДНДПМЕ, 2003.–№3.–С.147-165.
13. Sarycheva L. Using GMDH in Ecological and Socio-economical Monitoring Problems // Systems Analysis and Modeling Simulation (SAMS). – 2003, Vol.43, №10. – P.1409-1414.
14. Сарычева Л.В. Компьютерная система эколого-социально-экономического мониторинга регионов // Сб. науч. тр. НГУ.– №19. – Днепропетровск: НГУ, 2004. – С.186-193.
15. Романюк С.А., Сарычева Л.В. О методологии эколого-социально-экономического аудита // Моделювання та інформатизація соціально-економічного розвитку України. Зб. наук. пр. – К.: ДНДПМЕ, 2003. – №3. – С.10-31.
16. Sarychev A.P., Sarycheva L.V. The optimal set features determination in discriminant analysis by the group method of data handling // Systems Analysis and Modeling Simulation (SAMS). – 1998, Vol.31, P.153-167.

17. Куриленко В.А., Сарычева Л.В. Интегрирование динамических моделей и ГИС в системе эколого-экономического мониторинга // Сб. науч. тр. НГА Украины.– №9, Т.1. – Днепропетровск: НГА Украины, 2000. – С.103-108.
18. Сарычева Л.В., Литвиненко С.Ю. Районирование территорий на торговые зоны с помощью геоинформационных технологий и теории нечетких множеств // Наук. вісник НГА України, 2001. – №4. – С.40-42.
19. Сарычева Л.В., Куриленко В.А. Интегральная эколого-социально-экономическая оценка территории и картографирование в планировании устойчивого регионального развития // Актуальні проблеми автоматизації та інформаційних технологій.– Т.3. Зб. наук. пр. – Дніпропетровськ: Навч. книга, 2000.– С.98-106.
20. Сарычева Л.В., Бусыгин Б.С., Никулин С.Л. Компьютерный эколого-экономический мониторинг промышленного региона // Сб. науч. тр. НГА Украины.– №7, Т.1. – Днепропетровск: НГА Украины, 1999.– С.58-63.
21. Сарычева Л.В., Замковая Л.Д. Анализ сценариев развития региона в системе эколого-экономического мониторинга // Наук. вісник НГА України, 2001.– №5. – С.47-48.
22. Сарычева Л.В., Качанов О.В. Схеми класифікації регіонів за показниками еколого-соціально-економічного моніторингу в геоінформаційній системі // Геоінформатика. – 2002.– №4. – С.53-63.
23. Сарычева Л.В., Криворучко О.В. Моделирование водного баланса озера Солоний лиман за даними гідрогеологічного моніторингу // Наук. вісник НГУ, 2003.– №10.– С.63-66.
24. Замковая Л.Д., Сарычева Л.В. О практической устойчивости конечноразностных систем при возмущениях различных порядков // Наук. вісник НГА України, 2000. – №4. – С.103-104.
25. Бусыгин Б.С., Сарычева Л.В., Никулин С.Л. Использование ГИС для оценки запасов месторождений // Сб. науч. тр. НГА Украины.–№7, Т.1.– 1999.– С.107-111.
26. Сарычева Л.В. Методика объективной кластеризации для зонирования территории по эколого-социально-экономическим показателям // МКІМ`2002. Міжнарод. конф. з індуктивного моделювання. Праці в 4-х томах. – Т.1., Ч.2. – Львів: Державний НДІ інформаційної інфраструктури, 2002. – С.260-265.
27. Сарычева Л.В. Создание электронного атласа Днепропетровской области для сбалансированности регионального развития // Сб. докл. IV Междунар. науч.-практ. конф. „Проблемы и перспективы использования геоинформационных технологий в горном деле”. – Днепропетровск: НГУ, 2002.– С.76-78.
28. Сарычева Л.В. Картирование эколого-социально-экономических показателей для управления природоохранной деятельностью в регионах // Сб. докл. II Междунар. науч.-практ. конф. „Проблемы и перспективы использования геоинформационных технологий в горном деле”.– Днепропетровск: НГА Украины, 2000. – С.258-260.
29. Sarycheva L., Zhuikov A. Cluster Analysis of Territories by the Totality of Ecological and Socio-Economic Indices // IEEE 2001 International Geoscience and Remote Sensing Symposium, UNSW. – P.663-664.
30. Бусыгин Б.С., Никулин С.Л., Сарычева Л.В. Векторно-растровая моделирующая геоинформационная система интегрированного анализа данных // Тези доп.

Міжнар. конф. „Моделювання та оптимізація складних систем” (МОСС-2001). – Київ: КНУ ім. Шевченка, 2001. – Т.3. – С.74.

31. Автоматизированная система и компьютерная технология анализа ситуаций в геологии и экологии / Б.С. Бусыгин, С.Л. Никулин, С.В. Яковлев, Л.В. Мирошниченко // Тези доп. II Міжнар. симпозіуму “Застосування математичних методів і комп’ютерних технологій при вирішенні задач геохімії і охорони навколишнього середовища”. – Львів, 1994. – С. 15-17.

Особистий внесок автора в роботи, опубліковані в співавторстві: [15 – 23] – формування основної ідеї та розробка теоретичних основ методів; [24, 25] – обґрунтування та приклади практичного використання досліджень; [29 – 31] – розробка теоретичних та алгоритмічних основ технології.

АНОТАЦІЯ

Саричева Л.В. Математичне забезпечення та комп’ютерна технологія еколого-соціально-економічного моніторингу регіонів. – Рукопис.

Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.13.06 Автоматизовані системи управління і прогресивні інформаційні технології. Національний авіаційний університет. Київ, 2004.

Дисертацію присвячено проблемі видобування знань з різномірних екологічних і соціально-економічних (ЕСЕ) моніторингових даних для кількісної оцінки сталості та збалансованості регіонального розвитку. Запропоновано нові методи та алгоритми інтелектуального аналізу ЕСЕ-даних, інтегральної ЕСЕ-оцінки території, виділення аномалій регіонального розвитку.

Розглянуто геоінформаційне забезпечення ЕСЕ-моніторингу регіонів. Викладено методи та засоби геоінформаційного аналізу просторово розподіленої ЕСЕ-інформації. Класифіковано геоіконічні моделі ЕСЕ-показників, рекомендовано геоінформаційні технології їх побудови.

Запропоновано методику та відповідну комп’ютерну інформаційну технологію просторово-часового аналізу ЕСЕ-показників моніторингу регіонів. Визначено призначення, мету і принципи створення, розроблено структурно-функціональну схему комп’ютерної системи ЕСЕ-моніторингу регіонів. Проаналізовано вимоги до системи в цілому та її окремих підсистем.

Ключові слова: геоінформаційні системи та технології, комп’ютерний еколого-соціально-економічний моніторинг регіонів, інтелектуальний аналіз даних, геоіконічне моделювання.

АННОТАЦИЯ

Сарычева Л.В. Математическое обеспечение и компьютерная технология эколого-социально-экономического мониторинга регионов. – Рукопись.

Диссертация на соискание ученой степени доктора технических наук по специальности 05.13.06 Автоматизированные системы управления и прогрессивные информационные технологии. Национальный авиационный университет. Киев, 2004.

Диссертация посвящена проблеме извлечения знаний из разнородных экологических и социально-экономических (ЭСЭ) мониторинговых данных для количественной оценки устойчивости и сбалансированности регионального развития. Изложены концепция, задачи и перспективы ЭСЭ-мониторинга и ЭСЭ-аудита. Предложены новые методы и алгоритмы интеллектуального анализа ЭСЭ-данных, выделения аномалий регионального развития, интегральной ЭСЭ-оценки территории на основе показателей трех информационных блоков – “экология”, “экономика”, “социум”, каждый из которых разделен на три типа показателей – факторы загрязнения, показатели состояния и управляющие воздействия. С единых системных позиций классифицированы методы кластеризации и выбора информативных признаков. Освещены методические вопросы ЭСЭ-моделирования.

Рассмотрено геоинформационное обеспечение ЭСЭ-мониторинга регионов. Исследованы вопросы использования геоинформационных систем и технологий для анализа показателей ЭСЭ-мониторинга. Изложены методы и средства геоинформационного анализа пространственно распределенной ЭСЭ-информации, технология подготовки электронных атласов устойчивого развития регионов.

Классифицированы геоиконические модели ЭСЭ-показателей: 1) ранговые картограммы; 2) картодиаграммы; 3) аналитические карты; 4) комплексные карты; 5) синтетические карты; 6) геоиконические модели региональной генерализации; 7) интегральные региональные модели; 8) динамические модели и анаморфозы.

Разработаны технологии построения отдельных типов моделей, основанные на разных схемах классификации, районирования на основе интегральной оценки территории, картировании результатов ранжирования, распознавания, кластеризации, прогнозирования, выделения аномалий регионального развития. Рассмотрены вопросы генерирования новых знаний на базе геоинформационного анализа и геоиконического моделирования ЭСЭ-показателей регионов.

Предложена методика и соответствующая компьютерная информационная технология пространственно-временного анализа ЭСЭ-показателей мониторинга регионов. Концептуальная схема методики включает пять основных блоков: 1) геоинформационный анализ и геоиконическое моделирование; 2) интеллектуальный анализ данных; 3) генерирование новых знаний; 4) анализ и принятие решений; 5) диалог "лицо, принимающее решения, – ЭВМ".

Изложена методология построения системы компьютерного ЭСЭ-мониторинга: определены назначение, цели и принципы создания, требования к системе в целом и ее отдельным подсистемам. Разработана структурно-функциональная схема системы ЭСЭ-мониторинга, проанализированы характеристики системы и функции ее подсистем.

Приведены результаты практической реализации разработанных методов, алгоритмов и технологий. Рассмотрены задачи районирования территории по совокупности ЭСЭ-показателей с использованием методики объективной кластеризации и геоинформационного анализа, задачи выделения аномальных регионов, раз-

работки фрагментов ГИС "Охрана здоровья" города Днепропетровска, прогнозирования ЭСЭ-показателей по реальным данным мониторинга регионов Украины.

Ключевые слова: геоинформационные системы и технологии, компьютерный эколого-социально-экономический мониторинг регионов, интеллектуальный анализ данных, геоиконическое моделирование.

ANNOTATION

Sarycheva L.V. Mathematical Supplying and Computer Technology of Ecology-Socio-Economic Region Monitoring. – Manuscript.

Thesis on competition of a scientific degree of the doctor of engineering science on a speciality 05.13.06 Computer-aided control systems and progressive information technologies. National Air University. Kyiv. 2004.

The dissertation is devoted to a problem of extraction of knowledge from diverse ecological and socio economic (ESE) monitoring data for a quantitative rating of the stability and equilibrium of regional development. The concept, tasks and prospects of ESE-monitoring and ESE-audit are stated. New methods and algorithms of an intellectual analysis of ESE-data based on the Group Method of Data Handling, algorithm of an integrated ESE-rating of a territory are offered on the basis of parameters of three information blocks – ecology, economy, society, each of which is divided into pollution factors, state parameters and control influences. Methods of clustering and choice of informative attributes are classified from uniform system positions. The methodical questions of ESE-modeling are covered.

The geoinformation support of ESE-monitoring of regions is considered. Issues of using the geoinformation systems and technologies for the analysis of parameters of ESE-monitoring are investigated. Methods and means of the geoinformation analysis of the spatially distributed ESE-information, a technology of preparation of the electronic atlases of steady development of regions, geoiconic modeling are presented. Geoiconic models of ESE-parameters are classified, technologies of construction of separate types of models are offered based on different classification schemes, zoning on the basis of an integrated rating of a territory, of ranging results mapping, recognition, clustering, forecasting, allocation of anomalies of regional development.

A technique and an appropriate computer information technology of the spatial – temporal analysis of ESE-parameters of regions monitoring are offered. A methodology for construction of a computer ESE-monitoring system is suggested: assignment, purposes and principles of creation, requirements to the system as a whole and to its separate subsystems are determined. A structural and functional scheme of an ESE-monitoring system is developed, characteristics of the system and functions of its subsystems are analyzed.

Key words: geoinformation systems and technology, computer ecology-socio-economical region monitoring, intellectual data analysis, geoiconic modelling.

Саричева Людмила Віссаріонівна

Математичне забезпечення та комп'ютерна технологія еколого-
соціально-економічного моніторингу регіонів

(Автореферат)

Підписано до друку 04.11.04. Формат 30×42/4.
Папір Cartain. Ризографія. Умовн. друк. арк. 1,9.
Обліково-видавн. арк. 1,9. Тираж 100 прим. Зам. №

РВК НГУ
49027, м. Дніпропетровськ, просп. К. Маркса, 19.