

РЕФЕРАТ

Пояснювальна записка: 87 с., 45 рис., 4 табл., 19 джерел.

Мета дипломної роботи: створення інформаційної технології аналізу показників енергетичної безпеки регіонів.

Об'єкт дослідження: показники моніторингу енергетичної безпеки регіонів.

Предмет дослідження: методи, моделі та інформаційні технології моніторингу та аналізу показників енергетичної безпеки регіонів.

Наукова новизна: удосконалено метод кластеризації об'єктів, що потребують врахування географічного положення.

У вступі наведено стан проблеми та обґрунтована її актуальність.

У першому розділі наведені аналіз стану енергетичної безпеки у світі, основні відомості про поняття енергетичної безпеки, виконано огляд засобів та технологій, використаних у роботі.

У другому розділі наведена загальна схема аналізу стану енергетичної безпеки, описані основні методики, що застосовувались під час аналізу, виконано моделювання взаємозв'язків між показниками.

У третьому розділі виконано моделювання угруповань серед показників, проведено оцінку якості моделювання, подано інтерпретацію результатів моделювання.

Практичне значення роботи полягає у створенні інформаційної технології аналізу показників енергетичної безпеки та її використання для аналізу показників енергетичної безпеки.

Розроблено нову інформаційну технологію аналізу показників енергетичної безпеки регіонів, що враховує географічне положення регіону.

ЕНЕРГЕТИЧНА БЕЗПЕКА, СТАЛИЙ РОЗВИТОК, МОВА ПРОГРАМУВАННЯ, R, РЕГРЕСІЙНИЙ АНАЛІЗ, КЛАСТЕРНИЙ АНАЛІЗ, КЛАСТЕРИЗАЦІЯ, ГЕОГРАФІЧНЕ ПОЛОЖЕННЯ

ABSTRACT

Explanatory note: 87 p., 45 fig., 4 tabl., 19 references.

The objective of research: creation of information technology for analysis of energy security of regions.

The object of research: indicators for monitoring energy security of regions.

The subject of research: methods, models and information technologies of monitoring and analysis of energy security indicators of regions.

Scientific novelty: the method of clustering objects that need to take into account geographical location has been improved.

In introduction presents the state of the problem and its relevance.

The first section contains analysis of worldwide energy security state, energy security common concepts and overview of used technologies.

The second section contains general scheme of analysis energy security state, the basic techniques used in the analysis are described, the relationship between the indicators is simulated.

The third section contains modeling of groupings among indicators, evaluation of modeling quality, interpretation of modeling results. The modeling results are displayed on the world map.

The practical importance of the work is to create information technology for the analysis of energy security indicators.

A new information technology for the analysis of regional energy security indicators that taking into account the geographical location of the region has been developed.

ENERGY SECURITY, SUSTAINABLE DEVELOPMENT,
PROGRAMMING LANGUAGE, R, REGRESSION ANALYSIS, CLUSTER
ANALYSIS, CLUSTERIZATION, GEOPOSITION

	5
ВСТУП.....	8
РОЗДІЛ 1 АНАЛІЗ СТАНУ ОБЛАСТІ РІШЕНЬ	9
1.1. Аналіз стану енергетичної безпеки у світі	9
1.1.1. Аналіз стану доступності електроенергії	9
1.1.2. Аналіз стану відновлювальної енергетики	10
1.1.3. Відновлювальна енергії та транспорт	11
1.1.4. Політичний аспект відновлювальної енергії	11
1.1.5. Тенденції у відновлювальній енергії	13
1.1.5.1. Біоенергетика	13
1.1.5.2. Гідроенергетика	14
1.1.5.3. Сонячна енергетика	15
1.1.5.4. Вітрова енергетика	16
1.1.5.5. Геотермальна енергетика	16
1.1.5.6. Енергія океану	17
1.1.6. Енергоефективність	18
1.1.7. Енергетична безпека	19
1.1.7.1. Юридичний аспект енергетичної безпеки	20
1.1.7.2. Політичний аспект енергетичної безпеки	21
1.1.7.3. Роль відновлювальних джерел енергії в енергетичній безпеці	21
1.2. Сталий розвиток	22
1.2.1. Основні відомості	22
1.2.2. Поняття та визначення сталого розвитку	24
1.2.3. Триєдина концепція сталого розвитку	26
1.2.3.1. Економічна складова сталого розвитку	26

	6
1.2.3.2. Соціальна складова сталого розвитку	26
1.2.3.3. Екологічна складова сталого розвитку	27
Висновки	28
РОЗДІЛ 2 ПРОЕКТНІ РІШЕННЯ	30
2.1. Засоби та технології	30
2.1.1. Мова програмування R	30
2.1.1.1. Особливості мови програмування R	31
2.1.1.2. Графічні редактори скриптів та IDE	31
2.1.2. Кластерний аналіз	32
2.1.2.1. Завдання та етапи кластерного аналізу	33
2.1.2.2. Методи кластеризації	33
2.1.2.3. Ієрархічна кластеризація	35
2.1.2.4. Оцінка якості кластеризації	36
2.1.3. Регресійний аналіз	37
2.1.3.1. Задачі регресійного аналізу	38
2.1.3.2. Метод найменших квадратів	38
2.1.3.3. Регресійні моделі	38
2.2. Застосовані методики	39
2.2.1. Багатовимірне шкалювання	39
2.2.2. Метод головних компонент	40
Висновки	40
РОЗДІЛ 3 МОДЕЛЮВАННЯ	41
3.1. Загальна схема методики	41
3.2. Схема кластеризації з урахуванням географічного положення	42
3.3. Моделювання взаємозв'язків між показниками	43

3.4. Початкові дані та показники	54
3.4.1. Показники.....	54
3.4.2. Початкові дані.....	55
3.5. Моделювання угруповань серед показників	56
3.6. Відображення показників зростання виробництва та споживання електроенергії на карті світу.....	65
Висновки.....	70
ВИСНОВКИ	71
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	72
ДОДАТОК А Відомості матеріалів дипломного проекту.....	74
ДОДАТОК Б Код програми.....	75
ДОДАТОК В ВІДГУК.....	85
ДОДАТОК Г РЕЦЕНЗІЯ.....	86

ВСТУП

Сталий розвиток – це поєднання трьох складових принципів: економічний розвиток, соціальний прогрес та відповідальність за навколишнє довкілля. Він передбачає розвиток складових таким чином, щоб покращення в одній сфері не призводило до погіршення в іншій, тим самим витримуючи баланс між задовольнянням сучасних потреб людства та захистом інтересів майбутніх поколінь.

Одним з показників сталого розвитку є енергетична безпека – стан електроенергетики який гарантує технічно та економічно безпечне задоволення поточних і подальших потреб в споживанні енергії та охорону навколишнього природного середовища. Цей показник відноситься до економічної складової сталого розвитку, проте неможна забувати про вплив енергетичної промисловості на навколишнє довкілля та про скінченність корисних копалин. Саме ці проблеми призвели до розвитку виробництва відновлювальної та екологічно чистої енергії.

Відновлювальна електроенергія є одним з найбільш перспективних напрямів розвитку людства, що дозволить майже повністю відмовитись від видобування корисних копалин, що використовуються для виробництва електроенергії. Водночас, такий вектор розвитку дозволить значно зменшити негативний вплив на навколишнє довкілля.

За останні роки розвиток відновлювальної електроенергії значно вплинув і на інші сфери, наприклад, виробники автомобілів поступово відмовляються від двигунів внутрішнього згоряння на користь електричних двигунів, а деякі з них мають намір повністю припинити виробляти екологічно нечисті автомобілі.

Для того, щоб прискорити впровадження виробництва та використання екологічно чистої електроенергії необхідно постійно створювати нові методики та технології аналізу та моніторингу показників енергетичної безпеки.

РОЗДІЛ 1

АНАЛІЗ СТАНУ ОБЛАСТІ РІШЕНЬ

1.1. Аналіз стану енергетичної безпеки у світі

1.1.1. Аналіз стану доступності електроенергії

Енергетика має центральне значення для майже кожної з основних проблем з якими сьогодні стикається світ. Будь то робочі місця, безпека, зміна клімату, виготовлення продуктів харчування або збільшення прибутку – доступ до джерел енергії є визначаючим фактором. Стабільна енергетика необхідна для зміцнення економіки, захисту екосистем та процвітання людства.

Більше трьох мільярдів людей, більшість з яких мешкає у країнах Азії та країнах Африки на південь від Сахари, на сьогоднішній день готують їжу без використання екологічно чистих видів палива та більш ефективних технологій. Крім цього майже мільярд людей проживає без доступу до електроенергії, п'ятдесят відсотків з яких мешкає у країнах Африки на південь від Сахари. Незважаючи на те що, за останні роки спостерігається стрімке збільшення масштабів використання вітряної та сонячної енергії, доля енергії, що виробляється з цих джерел, серед загального обсягу енергоспоживання як і раніше залишається відносно незначною.

Прогресу, що досягнуто в кожному напрямі діяльності, спрямованої на забезпечення доступу до джерел енергії для всіх, та вирішення задач, пов'язаних з використанням енергії добутої з відновлювальних джерел енергії та збільшенням енергоефективності, на сьогоднішній день недостатньо. Для того, щоб досягти значних покращень необхідно збільшувати обсяг фінансування енергетичної та наукової галузей.

Серед фактів потрібно зазначити, що кожна п'ята людина в світі не має доступу до електроенергії. Близько трьох мільярдів людей залежить від традиційної біомаси, такої як деревина, рослинні залишки, що

використовуються для приготування їжі та опалення оселі. Енергетика є домінуючим фактором в області глобального потепління, на її долю припадає приблизно шістдесят відсотків від загального об'єму глобальних викидів парникового газу в атмосферу. Забруднення повітря в результаті приготування їжі та опалювання оселі з використанням горючого палива забрало майже чотири з половиною мільйони життів у 2012 році, шість з десяти загиблих – це жінки та діти. Частка використання відновлювальної енергії серед загального споживання енергії у 2015 році сягнула 17,5%.

1.1.2. Аналіз стану відновлювальної енергетики

Прогрес у відновлюваних джерелах енергії залишається зосередженим у енергетичному секторі, в той час як набагато менше зростання спостерігається у опаленні, охолодженні та транспорті. Відновлювана енергія стала більш конкурентоспроможною порівняно зі звичайною тепловою енергією. Відновлювальні джерела енергії забезпечили понад 26% світового виробництва електроенергії у 2018 році. Зростання було зумовлене стабільною політикою. Як і в попередні роки, серед відновлюваних джерел енергії спостерігалось набагато менше зростання у секторах опалення, охолодження та транспорту, при цьому прогрес стримувався відсутністю сильної політичної підтримки та повільним розвитком нових технологій. [1, с. 17]

Частка відновлювальної енергії збільшується в енергетичному секторі, а в 2018 році було встановлено електростанції із загальною потужністю 181 ГВт. Однак темпи збільшення нових потужностей знизилися після років зростання.

Глобальна потужність відновлюваної енергетики зростає приблизно до 2378 ГВт у 2018 році. Уже четвертий рік поспіль відновлювані потужності з виробництва електроенергії випереджають традиційні електростанції з викопного палива та атомної енергії за кількістю встановлень. Встановлено близько 100 ГВт сонячної фотоелектрики – 55% відновлювальної потужності, 28% з вітрової енергії 11% та гідроенергії. Загалом, частка потужностей відновлюваної енергії серед загальної кількості потужностей з виробництва електроенергії зростає більше ніж на 33%.

Відновлювальна енергія зарекомендувала себе у світовому масштабі. У 2018 році понад 90 країн встановили щонайменше 1 ГВт генеруючих потужностей, тоді як принаймні 30 країн перевищили 10 ГВт. Вітроенергетика та сонячна фотоелектрика додатково збільшили свою частку в деяких місцях, все більша кількість країн тепер має більше 20% відновлюваних джерел енергії у своєму виробництві електроенергії.[1, с.18]

1.1.3. Відновлювальна енергії та транспорт

Частка відновлюваної енергії у транспорті незначно зросла, порівняно з попереднім роком до 3,3%. Більше за все це забезпечується рідким біопаливом, однак сектор все більш відкритий до електрифікації, що відкриває можливості для подальшої інтеграції відновлюваної енергії. Незважаючи на зростання виробництва етанолу та біодизеля у 2018 році, зростання використання біопалива для транспорту залишається обмеженим невизначеністю політики та повільним прогресом у розробці відновлюваного палива для таких ринків, як авіація.

Протягом 2018 року були виявлені позитивні ознаки щодо залізничного, авіаційного та морського транспорту, з новими ідеями, партнерствами та ініціативами щодо підтримки відновлюваних джерел енергії та декарбонізації.

Частка електромобілів на дорогах світу в 2018 році збільшилась, в основному, це зумовлено зусиллями щодо зменшення забруднення повітря. Загальна кількість електричних легкових автомобілів збільшилася на 63% порівняно з 2017 роком, і все більше міст впроваджують автопарк електричного громадського транспорту. [1, с. 18]

1.1.4. Політичний аспект відновлювальної енергії

Політична підтримка відновлюваної енергетики була представлена майже у всіх країнах світу у 2018 році та спостерігається на всіх рівнях влади. Політики мають можливість розробити ефективну комбінацію законодавчих актів підтримки відновлювальної енергії з урахуванням відповідних юрисдикцій. У міру зниження витрат на відновлювані технології ці заходи продовжують розвиватися та адаптуватися.

Різноманітні переваги поновлюваних джерел енергії – такі, як: поліпшення охорони здоров'я населення через зменшення забруднення, підвищення надійності та стійкості та створення робочих місць – є рушійними політичними діями у всьому світі. Однак рамки політики щодо відновлювальної енергетики сильно відрізняються за обсягом та всеосяжністю, і більшість залишаються далеко від рівня амбіцій, необхідних для досягнення міжнародних цілей в галузі клімату.

У 2018 році законодавчі акти, що стосуються відновлювальної енергії були прийняті у 169 країнах на національному або державному рівні. Нові та переглянуті напрямки розвитку стають все більш масштабними, особливо в енергетичному секторі. Набагато менше країн мали цілі на поновлювані джерела енергії, спеціально для секторів опалення, охолодження та транспорту, а наміри для трансформації енергії в масштабах економіки залишаються рідкісними. Державні уряди часто є першими рушіями у створенні інноваційних та амбітних механізмів, включаючи амбіції на створення 100% відновлювального енергетичного сектору.[1, с. 20]

Політика щодо відновлюваної енергетики в транспортному секторі в основному зосереджена на автомобільному транспорті; однак, зростаюче використання електроенергії та сучасних біопалив, разом із посиленням зусиль щодо декарбонізації, сприяє підтримці відновлюваних джерел енергії як у залізничному, судноплавстві, так і в авіації.

У 2018 році біопаливо було центральним компонентом національних законодавчих актів в галузі поновлювальної енергії в транспортній політиці, з дотриманням мандатів, що існували в 70 країнах на національному або державному рівні до кінця 2018 року.

Енергетичний сектор отримав більшу частину політичної уваги, орієнтованої на відновлювані джерела енергії, у 2018 році. Політики продовжували звертатися до конкурентних аукціонів замість традиційної політики з фіксованою ціною, а аукціони проводилися щонайменше у 48 країнах (порівняно з 29 роком раніше), включаючи багатьох країн в Африці.

Принаймні один з аукціонів у 2018 році був нейтральним до технології (у Бразилії), тоді як принаймні шість були нейтральними щодо поновлюваних технологій. Китай призупинив фінансову підтримку сонячних проєктів на користь аукціонів, а в найближчі роки повинен відбутися перехід до аукціонів для проєктів вітроенергетики в країні.

Технології відновлюваної енергії отримали як пряму, так і непряму підтримку через політику, спрямовану на зменшення кліматичних змін та адаптацію до них. Системи торгівлі податками на вуглець і викиди є одними з механізмів політики, які можуть стимулювати інтерес до поновлюваних джерел енергії для досягнення кліматичних цілей. До кінця 2018 року було реалізовано щонайменше 54 ініціативи щодо ціноутворення вуглецю (порівняно з 46 у 2017 році), включаючи 27 систем торгівлі викидами та 27 податків на викиди вуглецю.[1, с. 21]

1.1.5. Тенденції у відновлювальній енергії

1.1.5.1. Біоенергетика

Сучасна біоенергетика є найбільшим учасником глобального постачання відновлюваної енергії. У 2017 році сучасна біоенергетика забезпечила приблизно 5% світового споживання енергії, майже половину всього внеску відновлюваної енергії. Біоенергетика зробила свій найбільший внесок у сектор опалення та охолодження (5%), далі – транспортний сектор (3%) та електропостачання (2,1%). Використання біоенергетики зростає найшвидше в електроенергетиці та транспортному секторі, тоді як воно відстає в галузі опалення.

У 2018 році Європейський Союз (ЄС) зберігав лідируючі позиції щодо використання сучасної біоенергетики для опалення, при цьому прогрес був зумовлений найбільшим чином чином Директивою відновлюваної енергетики ЄС.

Китай є світовим лідером у виробництві біоелектрики, за ним слідує США, Бразилія, Індія та Німеччина. У виробництві біопалива лідерами є США та Бразилія, які разом виробили 69% усіх біопалив у 2018 році.

Ринки біоенергетики та промислова діяльність сильно керовані політикою, але на торговельну схему, особливо на транспортне біопаливо, значно вплинули зміни імпорتنих тарифів та інші подібні заходи.

Промисловий розвиток продовжує зосереджуватися на розробці передових біопалив, які можуть запропонувати поліпшені показники стійкості, а також використовувати їх у нових сферах, таких як авіація.[1, с. 22]

1.1.5.2. Гідроенергетика

Для гідроенергетики характерна стабільність на ринку, зростаюча конкуренція в галузі та зростаючий попит на зберігання енергії.

Світовий ринок гідроенергетики у 2018 році виглядав дуже схожим на попередній рік за ростом потужностей та концентрацією діяльності.

Приблизно 20 ГВт було додано для досягнення загальної встановленої потужності близько 1132 ГВт.

В усьому світі генерація гідроенергетики, яка змінюється щороку зі змінами погодних умов та інших місцевих умов, оцінювалась 4210 ГВт–год.

Як і в попередніх раках, Китай лідирував у введенні в експлуатацію нових гідроенергетичних потужностей, що становить понад 35% нових установок у 2018 році. Бразилія посідає друге місце, за нею йдуть Пакистан та Туреччина.

Помітною особливістю гідроенергетики у 2018 році стали набряки ряду застарілих об'єктів, які потребують ремонту та модернізації. Більше половини всіх гідроелектростанцій у всьому світі або вже пройшли, або незабаром потребуватимуть капітального ремонту та модернізації.

Ще однією тенденцією стало зростаюче визнання цінності гідроенергетичних об'єктів для ефективної інтеграції змінних джерел відновлюваної енергії, таких як сонячна фотоелектрична енергія та вітроенергетика, та потенційної синергії між гідроенергетикою та іншими технологіями відновлюваної енергії, залежно від місцевих умов.[1, с. 22]

1.1.5.3. Сонячна енергетика

1.1.5.3.1. Сонячна фотоелектрична енергія

Високий попит на сонячну фотоелектрику спостерігається на ринках, що розвиваються, та в Європі, на відміну від Китаю, що було викликано змінами політики в середині року.

Самостійне споживання є важливим рушієм ринку нових розподілених систем у деяких регіонах, а корпоративні закупівлі сонячної фотоелектрики значно збільшилися, особливо в США та Європі. У всьому світі гірничо-шахтна, обробна та інші галузі промисловості споруджували сонячні електростанції (та інші відновлювані) для управління їх роботою.

Промисловість сонячної фотоелектрики зазнала значного струсу у 2018 році. Рішення Китаю обмежувати внутрішній попит призвело до світових потрясінь, оскільки китайські модулі заповнили світовий ринок, а торговельні суперечки торкнулися галузі в деяких країнах. Рекордно низькі ціни на аукціонах, зумовлені сильною конкуренцією та нижчими цінами на панелі, принесли подальшу консолідацію. Тим не менш, також спостерігаються інвестиції в нові, більш ефективні виробничі потужності та додаткові досягнення в галузі сонячної фотографії.[1, с.23]

1.1.5.3.2. Сонячна енергія для опалення та охолодження

Приблизно 33,3 ГВт нових сонячних теплових потужностей було введено в експлуатацію в 2018 році, збільшивши загальну глобальну потужність на 2% до приблизно 480 ГВт.

На частку Китаю припадає близько 74% глобальних потужностей, за ним ідуть Туреччина, Бразилія та США. Незважаючи на те, що валовий приріст Китаю зменшується п'ятий рік поспіль, більшість найбільших ринків сонячного опалення та охолодження за межами Китаю вперше почали зростати тільки у 2015 році. Це було пов'язано з кількома факторами, включаючи політику щодо чистого повітря та підвищення конкурентоспроможності сонячних теплових систем. Крім того, зростання попиту на Близькому Сході та в Східній та Центральній Африці дозволило

декільком виробникам сонячних колекторів південної Європи збільшити обсяги виробництва.

Щонайменше 37 нових масштабних систем було введено в експлуатацію в усьому світі для теплопостачання районних мереж або великих будівель.[1, с.23]

1.1.5.4. Вітрова енергетика

Станом на 2018 рік світовий ринок вітру був досить стабільним: було додано близько 51 ГВт потужностей, що збільшило накопичувальну потужність до 591 ГВт. Після рекордного року вітроенергетики в Європі та Індії в 2017 році обидва ринки скоротилися в 2018 році, але помітне зростання відбулося в кількох інших регіонах та країнах. Азія була найбільшим регіональним ринком, що представляє майже 52% доданої потужності.

В офшорному сегменті сім країн Європи та дві в Азії підключили 4,5 ГВт, збільшивши сукупну глобальну потужність на 24% до 23,1 ГВт. Успіх офшорної вітроенергетики в Європі викликав інтерес майже в кожному іншому регіоні.

Незважаючи на те, що падіння цін сприяє переміщенню вітроенергетики на нові ринки та прискоренню продажів, глобальний перехід до більш конкурентоспроможних механізмів, таких як аукціони та тендери, призвів до інтенсивної цінової конкуренції, яка здавлює весь ланцюжок вартості та викликає конкуренцію серед виробників вітрових турбін.

Постійні досягнення в промислових технологія (включаючи більш великі турбіни), які збільшують виробництво енергії на турбіну, підвищують ефективність та виробництво електростанції та знижують низькі витрати електроенергії на енергію вітру.[1, с.23]

1.1.5.5. Геотермальна енергетика

Геотермальний енерговипуск у 2018 році оцінювався у 630 ПДж, приблизно половина цього у вигляді електроенергії (89,3 ТВт-год), а інша половина – як тепло.

У 2018 році глобальна геотермальна індустрія продовжувала виражати мірний оптимізм щодо геотермального розвитку, загартований постійними занепокоєннями щодо різних проблем, пов'язаних з галуззю, а також сприйняттям недостатньої або колишньої підтримки уряду. Міжнародні агенції та банки розвитку вивчали можливості подолання деяких із цих викликів та фінансування нових розробок.[1, с.22]

1.1.5.6. Енергія океану

Деякі технології в галузі видобутку енергії з океану є близькими до комерціалізації, але послідовна політика підтримки та гарантії доходу залишаються критичними факторами успішності.

Енергетика океану являє собою найменшу частину ринку відновлюваної енергії, більшість проектів орієнтовані на відносно невеликі демонстраційні та пілотні проекти менше 1 МВт. Чисті доповнення в 2018 році становили приблизно 2 МВт, при прогнозній експлуатаційній потужності на кінець року – 532 МВт.

Діяльність з розробки спостерігається у всьому світі, але зосереджена в основному в Європі, особливо у берегів Шотландії, де в 2018 році було розгорнуто декілька масивів припливних турбін. Ресурсний потенціал океанічної енергії величезний, але він залишається значною мірою невикористаним, незважаючи на десятиліття зусиль з розвитку.

2018 рік став одним із надзвичайних контрастів для енергетики океану. З одного боку, виробники припливних турбінних масивів, зокрема, зазначили технологічний успіх та прогрес у напрямку комерціалізації. З іншого боку, негативний прогноз спонукав одного видатного розробника припливних технологій відмовитися від усіх планів щодо подальшого виготовлення та впровадження. Фінансова та інша підтримка з боку урядів, особливо в Європі та Північній Америці, продовжувала посилювати приватні інвестиції в енергетичні технології океану, особливо припливні течії та пристрої енергетичної хвилі.[1, с.22]

1.1.6. Енергоефективність

Міжнародні зусилля щодо планування траєкторій досягнення цілей сталого розвитку, як правило, визнають взаємодоповнюваність розгортання відновлюваної енергії та заходів з енергоефективності. Наприклад, у 2011 році ініціатива Організації Об'єднаних Націй «Стала енергія для всіх», визнаючи спільну роль відновлюваних джерел енергії та ефективність у забезпеченні універсального доступу до сталої енергії, ставлять мету подвоїти як частку відновлюваної енергії у світовому споживанні кінцевої енергії (до 36% до 2030 року), так і темп підвищення енергоефективності.

У 2018 році Міжурядова група з питань зміни клімату представила декілька шляхів пом'якшення зміни клімату, що відповідають відносно високій ймовірності обмеження довгострокового підвищення середньої глобальної температури на 1,5 градусів Цельсія вище доіндустріального рівня. Кожен із шляхів частково характеризується відносним скороченням світового попиту на енергію. Зменшення попиту на енергію вимагає прогресу як енергоефективності (специфічної технології), так і енергозбереження.

Докладаються зусилля для розмежування наслідків трьох основних детермінант загальної кінцевої потреби в енергії: структурні зміни в економіці, зміни рівня активності в кожному економічному секторі та зміни ефективності використання енергії в кожному секторі. Аналіз показав що без покращення енергоефективності глобальний попит на кінцеву енергію в 2017 році був би на 12% вище, ніж у 2000 р. Це означає середньорічне переміщення потреби в енергії нижче 0,7% за цей період. Між тим, між 2005 та 2017 роками частка загального споживання кінцевої енергії, що задовольняється сучасними відновлюваними джерелами енергії, зросла із середньорічним темпом 2,9%. [1, с.169]

По всьому світі політична підтримка енергоефективності значно зросла між 2010 та 2017 роками. Однак найбільший прогрес був у розповсюдженні національних планів дій та цілей енергоефективності, тоді як кількість конкретних національних мандатів зростала повільніше. Уряди міст прийняли

будівельні енергетичні кодекси, мінімальні стандарти ефективності енергоефективності та інші тверді зобов'язання, а міста продовжували відігравати визначну роль у розробці та реалізації політики щодо енергоефективності.

Хоча як відновлювані джерела енергії, так і ефективність є важливим елементом більш стійких енергетичних систем, розробники політики ініціатив можуть боротися з тим, де найбільш ефективно розподіляти ресурси: на стороні пропозиції (поновлювані джерела енергії) або на попиті (енергоефективність). У США розповсюдження відновлюваних джерел енергії з нульовою змінною вартістю, таких як вітроенергетика та сонячна фотоелектрична енергія – а також низька вартість природного газу – перетворюється на менші витрати, які уникнуть для програм енергоефективності в електроенергетиці, потенційно зменшуючи економічну ефективність заходів з підвищення ефективності, що рухаються вперед. У поєднанні з мінливістю швидко зростаючих сонячних фотоелектричних і вітроенергетичних результатів це більший інтерес серед політиків у використанні майбутніх заходів з енергоефективності із досягненням реагування на попит, розподілену генерацію, сховища та електромобілі, щоб забезпечити різні системні послуги, які відображають потреби в електромережі, що залежать від місця та часу.[1. с. 170]

1.1.7. Енергетична безпека

Енергетична безпека – стан електроенергетики, який гарантує технічно та економічно безпечне задоволення поточних і перспективних потреб споживачів в енергії та охорону навколишнього природного середовища.

Під енергетичною безпекою слід розуміти здатність держави в особі її органів управління забезпечити кінцевих споживачів енергією в необхідному обсязі та належної якості у звичайних умовах, а також під час дії дестабілізуючих факторів (надзвичайних ситуацій) внутрішнього чи зовнішнього характеру у межах гарантованого покриття мінімального обсягу

найважливіших потреб країни, окремих її районів, міст, селищ чи об'єктів у паливно–енергетичних ресурсах.

Поняття енергетичної безпеки розділяють на три рівні:

- Політична – енергетична безпека, яка відноситься до енергетичної безпеки держави, її суб'єкту чи регіону;
- Економічна – включає в себе тарифи та запаси енергоресурсів;
- Техногенна – має на увазі техногенний характер ризиків для людини, майна та навколишнього середовища, пов'язаний з експлуатацією будь–яких енергоустановок. Цей тип енергобезпеки включає в себе комплексну оцінку техногенної небезпеки об'єкта енергетики: електробезпека і пожежна безпека, електромагнітна і механічна безпека, екологічна та промислова безпека, вибухову безпеку, радіаційна, ядерна та хімічна безпеки.[13]

1.1.7.1. Юридичний аспект енергетичної безпеки

Торгівля енергоносіями юридично регулюється різними міжнародними угодами. Частина угод спеціально розробляються для підвищення енергетичної безпеки. Часто вони ставлять собі за мету відокремлення виробничої діяльності від транспортної та розподільчої і націлені на відкриття та лібералізацію енергетичного ринку.

Перший важливий документ, присвячений енергетичній безпеці, був створений країнами–виробниками, які за допомогою нього намагалися закріпити національний суверенітет над природними ресурсами. На початку 1960–х ООН офіційно визнала державний суверенітет і суверенні права на енергоресурси. Ця резолюція ООН в певній мірі враховує питання енергетичної безпеки, однак вона більше благоволить країнам–споживачам. У зв'язку з цим була створена організація країн – експортерів нафти, яка спробувала повернути країнам–виробникам контроль над енергоресурсами.

В діючих на сьогодні угодах є деякі суперечності. Так звані «загальні» торгові угоди іноді вступають в протиріччя з юридичними аспектами

енергетичної безпеки. Більшість торгових угод регулюють ринкові відносини, але вони не можуть вирішити проблему енергетичної безпеки.

1.1.7.2. Політичний аспект енергетичної безпеки

Величезне значення для забезпечення енергобезпеки має сталий розвиток і стабільність держав–виробників в сфері енергетики. Така стійкість може опинитися під загрозою в зв'язку з необхідністю вести розвідку і видобуток енергоресурсів у все більш складних умовах, що можливо лише при використанні надсучасних технологій, нанесенні шкоди навколишньому середовищу, а також створює труднощі з транспортуванням.

Енергетична безпека залежить від промислової політики, яку проводять в країнах–постачальниках енергії. У сучасному світі порушення в забезпеченні енергетичної безпеки, як внаслідок природних причин, так і внаслідок зовнішнього, штучного впливу, є однією з головних причин дестабілізації соціальної, а в майбутньому і політичної ситуації в країні.

1.1.7.3. Роль відновлювальних джерел енергії в енергетичній безпеці

Впровадження розробки відновлювальних джерел енергії, як правило, збільшує різноманітність джерел електроенергії та, завдяки місцевій генерації, сприяє гнучкості системи та її стійкості до центральних ударів. Для тих країн, де зростаюча залежність від імпортного газу є важливим питанням енергетичної безпеки, технології розробки відновлювальних джерел енергії можуть забезпечити альтернативні джерела електроенергії, а також витіснити попит на електроенергію за рахунок прямого виробництва тепла. Поновлюване біопаливо для транспорту є ключовим джерелом диверсифікації нафтопродуктів.

Оскільки ресурси, що мають настільки важливе значення для виживання у світі, почнуть скорочуватись, країни почнуть усвідомлювати, що потреба у відновлюваних джерелах палива та енергії буде настільки ж важливою, як ніколи. З виробництвом нових видів енергії, включаючи сонячну, геотермальну, гідроелектричну, біопаливо та вітрову енергію. При кількості сонячної енергії, яка потрапляє на світ за одну годину, є достатньо енергії для

живлення світу протягом одного року. З додаванням сонячних панелей у всьому світі знижується тиск на необхідність видобутку нафти.

Геотермальний потенціал може призвести до інших джерел палива, якщо компанії братимуть тепло з внутрішнього ядра землі для нагріву джерел води, ми могли б по суті використовувати пар, що створюється з нагрітої води до енергетичних машин, цей варіант є одним із найчистіших ефективні варіанти.[14]

1.2. Сталий розвиток

1.2.1. Основні відомості

Концепція сталого розвитку стала логічним переходом від екологізації наукових знань і соціально–економічного розвитку, що бурхливо почався у 1970–і роки. Питанням обмеженості природних ресурсів, а також забруднення природного середовища, що є основою життя, економічної і будь–якої діяльності людини, в 1970–і роки був присвячений ряд наукових робіт. Реакцією на це було створення міжнародних неурядових наукових організацій з вивчення глобальних процесів на Землі, таких як Міжнародна Федерація Інститутів Перспективних Досліджень (МФІП), Римський клуб (з його знаменитою доповіддю «Межі зростання»), Міжнародний інститут системного аналізу.

Проведення в 1972 році в Стокгольмі конференції ООН по навколишньому середовищі і створення програми ООН з навколишнього середовища ознаменувало включення міжнародного співтовариства на державному рівні у вирішення екологічних проблем, які стали стримувати соціально–економічний розвиток. Стала розвиватися екологічна політика і дипломатія, право навколишнього середовища, з'явилася нова інституційна складова – міністерства і відомства з навколишнього середовища.

У 1980–х роках почали вести мову про екорозвиток, розвиток без руйнування, необхідність сталого розвитку екосистем. Всесвітня стратегія охорони природи, прийнята в 1980, вперше містила згадку про сталий

розвиток. Друга редакція ВСОП отримала назву «Турбота про планету Земля – Стратегія стійкого життя» і була опублікована в жовтні 1991. У ній підкреслюється, що розвиток має базуватися на збереженні живої природи, захисту структури, функцій і різноманітності природних систем Землі, від яких залежать біологічні види. Для цього необхідно: зберігати системи підтримки життя (життєзабезпечення), зберігати біорізноманіття та забезпечити стале використання поновлюваних ресурсів. З'явилися дослідження з екологічної безпеки як частини національної і глобальної безпеки.

У 1980–і роки Програма ООН з навколишнього середовища (ЮНЕП) закликала до необхідності переходу до «розвитку без руйнування». У 1980 році вперше набула широкого розголосу концепція сталого розвитку у Всесвітній стратегії збереження природи, розробленої з ініціативи ЮНЕП, Міжнародного союзу охорони природи і Всесвітнього фонду дикої природи. У 1987 році в доповіді «Наше спільне майбутнє» Міжнародна комісія з навколишнього середовища і розвитку приділила основну увагу необхідності «сталого розвитку», при якому «задоволення потреб теперішнього часу не зменшує здатність майбутніх поколінь задовольняти свої власні потреби». Саме таке тлумачення поняття «сталий розвиток» зараз широко використовується в якості базового в багатьох країнах світу. [3]

Теорія і практика показали, що екологічна складова є невід'ємною частиною людського розвитку. В основі діяльності Міжнародної комісії з навколишнього середовища і розвитку і її доповіді «Наше спільне майбутнє» була покладена нова триєдина концепція сталого (еколого–соціально–економічного) розвитку. Всесвітній саміт ООН зі сталого розвитку (міжурядовий, неурядовий і науковий форум) у 2002 році підтвердив прихильність всього світового співтовариства ідеям сталого розвитку для довгострокового задоволення основних людських потреб при збереженні систем життєзабезпечення планети Земля.

1.2.2. Поняття та визначення сталого розвитку

Сталий розвиток – це «така координація економічного, екологічного і людського розвитку, яка відповідає потребам нинішнього покоління, не ставлячи під загрозу здатність майбутніх поколінь задовольняти свої власні потреби» (відповідно до термінології Всесвітньої комісії з навколишнього середовища і розвитку). Теорія сталого розвитку стала пріоритетною галуззю науки, швидко розвивається, а також базовою ідеологією всіх концептуальних і стратегічних національних і міжнародних документів за останні десятиліття.[1]

Цілі сталого розвитку є універсальним заклик до дії для подолання бідності, захисту планети і забезпечення того, щоб всі люди насолоджувалися миром і процвітанням.

Фактично мова може йти не про негайне припинення економічного зростання взагалі, а про припинення, на першому етапі, зростання нераціонального використання ресурсів навколишнього середовища. Останнє важко здійснити в світі зростаючої конкуренції та зростання таких нинішніх показників успішної економічної діяльності як продуктивність і прибуток. У той же час перехід до «інформаційного суспільства» – економіці нематеріальних потоків фінансів, інформації, зображень, повідомлень, інтелектуальної власності – призводить до так званої «дематеріалізації» господарської діяльності: Вже зараз обсяги фінансових операцій перевищують обсяги торгівлі матеріальними товарами в багато разів. Нову економіку рухають не тільки дефіцит матеріальних (і природних) ресурсів, але у все більшій мірі наявність ресурсів інформації і знань. Питома енергоємність господарської діяльності продовжує знижуватися, хоча загальне енергоспоживання поки росте. [2]

Значна більшість міжнародних організацій системи ООН додало до своєї діяльності вагому екологічну складову, орієнтовану на перехід до сталого розвитку. Експерти Світового банку визначили сталий розвиток як процес управління сукупністю (портфелем) активів, спрямований на збереження і

розширення можливостей, наявних у людей. Активи в даному визначенні включають не тільки традиційно підраховуваний фізичний капітал, але також природний і людський капітал. Щоб бути стійким, розвиток має забезпечити зростання – або принаймні не зменшення – в часі всіх цих активів. Для раціонального управління економікою країни застосовується та ж логіка, що використовується для раціонального управління особистою власністю.

Відповідно до наведеного визначення сталого розвитку головним показником стійкості, розробленим Світовим банком, є «справжні темпи (норми) заощадження» або «справжні норми інвестицій» в країні. Прийняті зараз підходи до вимірювання накопичення багатства не враховують виснаження і деградацію природних ресурсів, таких як ліси і нафтові родовища, з одного боку, а, з іншого – інвестиції в людей – один з найцінніших активів будь-якої країни. При переході на обчислення дійсних темпів заощаджень (інвестицій) цей недолік виправляється коригуванням розрахованих традиційними методами темпів заощаджень: в сторону зменшення – шляхом оцінки виснаження природних ресурсів і шкоди від забруднення навколишнього середовища (втрата природного капіталу), і в бік збільшення – шляхом обліку зростання людського капіталу (перш за все через інвестиції в освіту і базове медичне обслуговування).

В результаті шестирічного міжнародного діалогу з метою вироблення загальнолюдських цілей і загальних цінностей з'явився документ «Хартія Землі». Він був підготовлений за ініціативою громадянського суспільства і офіційно прийнятий на зборах Комісії Хартії Землі в штабі ЮНЕСКО в Парижі, в березні 2000 року. Місією «Хартії Землі» є пропаганда переходу до сталого способу життя і формування глобального співтовариства, заснованого на загальних етичних цінностях, що включають в себе повагу і турботу про все співтовариство живого, принципи екологічної цілісності, загальні права людини, повага до культурної різноманітності, економічну справедливість, демократію і культуру світу. [5]

1.2.3. Триєдина концепція сталого розвитку

Концепція сталого розвитку з'явилася в результаті об'єднання трьох основних точок зору: економічної, соціальної та екологічної.

1.2.3.1. Економічна складова сталого розвитку

Економічний підхід до концепції сталого розвитку заснований на теорії максимального потоку сукупного доходу Хікса–Ліндаль, який може бути проведений за умови, принаймні, збереження сукупного капіталу, за допомогою якого і проводиться цей дохід. Ця концепція має на увазі оптимальне використання обмежених ресурсів і використання екологічних – природо–, енерго–, і матеріало– заощаджуючих технологій, включаючи видобуток і переробку сировини, створення екологічно прийнятної продукції, мінімізацію, переробку та знищення відходів. Однак при вирішенні питань про те, який капітал повинен зберігатися (наприклад, фізичний або природний, або людський) і в якій мірі різні види капіталу взаємозаміні, а також при вартісній оцінці цих активів, особливо екологічних ресурсів, виникають проблеми правильної інтерпретації і підрахунків. З'явилися два види стійкості: слабка – коли мова йде про незменшуваного в часі природному та проведений капіталі, і сильна – коли не повинен зменшуватися природний капітал (причому частина прибутку від продажу невідновлюваних ресурсів повинна спрямовуватися на збільшення цінності відновлюваних природного капіталу).

1.2.3.2. Соціальна складова сталого розвитку

Соціальна складова сталого розвитку орієнтована на людину і спрямована на збереження стабільності соціальних і культурних систем, в тому числі, на скорочення числа руйнівних конфліктів між людьми. Важливим аспектом цього підходу є справедливий розподіл благ. Бажано також збереження культурного капіталу і різноманіття в глобальних масштабах, а також більш повне використання практики сталого розвитку, наявної в невідоміючих культурах. Для досягнення сталого розвитку, сучасному суспільству доведеться створити більш ефективну систему ухвалення рішень, що враховує історичний досвід і заохочує плюралізм. Важливо досягнення

справедливості не тільки серед поколінь, але і в періоди між поколіннями також. В рамках концепції людського розвитку людина є не об'єктом, а суб'єктом розвитку. Спираючись на розширення варіантів вибору людини як головну цінність, концепція сталого розвитку має на увазі, що людина повинна брати участь в процесах, які формують сферу її життєдіяльності, сприяти прийняттю і реалізації рішень, контролювати їх виконання.

1.2.3.3. Екологічна складова сталого розвитку

З екологічної точки зору, сталий розвиток має забезпечувати цілісність біологічних і фізичних природних систем. Особливе значення має життєздатність екосистем, від яких залежить глобальна стабільність всієї біосфери. Більш того, поняття «природних» систем і ареалів можна розуміти широко, включаючи в них створене людиною середовище, таке як, наприклад, міста. Основна увага приділяється збереженню здібностей самовідновлення і динамічної адаптації таких систем до змін, а не збереження їх в деякому «ідеальному» статичному стані. Деградація природних ресурсів, забруднення навколишнього середовища і втрата біологічного різноманіття скорочують здатність екологічних систем до самовідновлення.

Узгодження цих різних точок зору і їх переклад на мову конкретних заходів, які є засобами досягнення сталого розвитку – завдання величезної складності, оскільки всі три елементи сталого розвитку повинні розглядатися збалансовано. Важливі також і механізми взаємодії цих трьох концепцій. Економічний і соціальний елементи, взаємодіючи один з одним, породжують такі нові завдання, як досягнення справедливості всередині одного покоління (наприклад, щодо розподілу доходів) і надання цілеспрямованої допомоги бідним верствам населення. Механізм взаємодії економічного та екологічного елементів породив нові ідеї щодо вартісної оцінки та інтерналізації (обліку в економічній звітності підприємств) зовнішніх впливів на навколишнє середовище. Нарешті, зв'язок соціального та екологічного елементів викликала інтерес до таких питань як внутрішньогенераційна і міжгенераційна

рівність, включаючи дотримання прав майбутніх поколінь, і участі населення в процесі прийняття рішень. [4]



Рис. 1.1. Триєдина концепція сталого розвитку

Висновки

В більшості країн світу спостерігається прогрес у розвитку виробництва енергії з відновлювальних джерел, проте більше трьох мільярдів людей на сьогоднішній день готує їжу без використання екологічно чистих видів палива та більш ефективних технологій. Крім цього майже мільярд людей проживає без доступу до електроенергії.

Прогрес у відновлюваних джерелах енергії залишається зосередженим у енергетичному секторі.

Відновлювана енергія стала більш конкурентоспроможною порівняно зі звичайною тепловою енергією та забезпечила понад 26% світового виробництва електроенергії у 2018 році.

Загальна кількість електричних легкових автомобілів збільшилася на 63% порівняно з 2017 роком. Міста активно впроваджують автопарк електричного громадського транспорту.

Створення інформаційної технології аналізу показників енергетичної безпеки дозволить більш точно координувати дії для покращення стану енергетичної безпеки в залежності від політики розвитку та географічного положення регіону.

РОЗДІЛ 2

ПРОЕКТНІ РІШЕННЯ

2.1. Засоби та технології

Для розробки фрагменту інформаційної системи аналізу показників(індексів) сталого розвитку регіонів був застосований метод регресійного аналізу. Усі обчислення виконані за допомогою мови програмування R та деяких сторонніх бібліотек. В результаті були згенеровані дані для побудови геоіконічних моделей засобами ESRI ArcGis.

2.1.1. Мова програмування R

R – мова програмування і програмне середовище для статистичних обчислень, аналізу та зображення даних в графічному вигляді. Розробка R відбувалась під істотним впливом двох наявних мов програмування: мови програмування S з семантикою успадкованою від Scheme[6]. R названа за першою літерою імен її засновників Роса Іхаки (Ross Ihaka) та Роберта Джентлмена (Robert Gentleman) працівників Оклендського Університету в Новій Зеландії. Незважаючи на деякі принципові відмінності, більшість програм, написаних мовою програмування S запускаються в середовищі R.

R поширюється безкоштовно за ліцензією GNU General Public License у вигляді вільнодоступного вихідного коду або відкомпільованих бінарних версій більшості операційних систем: Linux, FreeBSD, Microsoft Windows, Mac OS X, Solaris. R використовує текстовий інтерфейс, однак існують різні графічні інтерфейси користувача.

R має значні можливості для здійснення статистичних аналізів, включаючи лінійну і нелінійну регресію, класичні статистичні тести, аналіз часових рядів (серій), кластерний аналіз і багато іншого. R легко розбудовується завдяки використанню додаткових функцій і пакетів доступних на сайті Comprehensive R Archive Network (CRAN). Більша частина стандартних функцій R, написана мовою R, однак існує можливість підключати код написаний C, C++, або Фортраном. Також за допомогою

програмного коду на C або Java можна безпосередньо маніпулювати R об'єктами.

2.1.1.1. Особливості мови програмування R

R належить до інтерпретованих мов програмування і для роботи використовується командний інтерпретатор.

Мова програмування – це штучна мова, створена для передачі команд машинам, зокрема комп'ютерам. Мови програмування використовуються для створення програм, котрі контролюють поведінку машин, та запису алгоритмів.

Інтерпретована мова програмування – мова програмування, в якій сирцевий код програми не перетворюється попередньо повністю у машинний код для виконання, як у компільованих мовах, а виконується рядок за рядком з допомогою спеціальної програми–інтерпретатора.

R підтримує концепцію Об'єктно–орієнтованого програмування (ООП) включаючи generic функції, результат виконання якої залежить від аргументів (типу об'єктів), що передаються generic функції. В мові програмування R всі змінні є об'єктами, кожен об'єкт належить до певного класу.[17]

Об'єктно–орієнтоване програмування – це метод програмування, заснований на поданні програми у вигляді сукупності взаємодіючих об'єктів, кожен з яких є екземпляром певного класу, а класи є членами певної ієрархії наслідування. Спочатку пишеться клас, а на його основі при виконанні програми створюються конкретні об'єкти (екземпляри класів). На основі класів можна створювати нові, які розширюють базовий клас і таким чином створюється ієрархія класів.[7]

2.1.1.2. Графічні редактори скриптів та IDE

Для роботи з R існує кілька графічних інтерфейсів (GUI)

- Графічна оболонка RGui разом з командною оболонкою (терміналом) R Console входять до базового пакету R у версії для Windows;

- RStudio – зручне кросплатформне середовище розробки з відкритим кодом (існує можливість запуску на віддаленому linux сервері);
- RapidMiner і розширення RapidMiner R – середовище розробки для аналізу і обробки даних з використанням R;
- Java Gui for R (JGR) – кросплатформний термінал і редактор R написаний на Java;
- Deducer – графічний інтерфейс для аналізу даних з використанням системи меню (подібний до SPSS). Розроблений для використання разом з JGR та RGui;
- Rattle GUI – кросплатформний графічний інтерфейс, розроблений збору та аналізу даних;
- R Commander – кросплатформний GUI з системою меню і доступними додатковими плагінами;
- RExcel – додаток до Microsoft Excel, який дозволяє використовувати можливості R;
- Sage – середовище для математичних розрахунків з використанням інтерфейсу веб-браузера, бібліотек R;
- Red-R – інтерфейс для аналізу, що використовує R;
- Tinn-R – графічний інтерфейс;

2.1.2. Кластерний аналіз

Кластерний аналіз – це задача розбиття заданої вибірки об'єктів (ситуацій) на підмножини, які називаються кластерами, так, щоб кожен кластер складався з схожих об'єктів, а об'єкти різних кластерів істотно відрізнялися. Завдання кластеризації відноситься до статистичної обробки, а також до широкого класу завдань навчання без вчителя.

Нехай X – множина об'єктів, Y – множина номерів кластерів, $\rho(x, x')$ – функція відстані між об'єктами. Задано кінцеву вибірку об'єктів: $X^m = \{x_1, \dots, x_m\} \subset X$. Потрібно розбити вибірку на непересічні підмножини,

що називаються кластерами, так, щоб кожен кластер складався з об'єктів, близьких за метрикою ρ , а об'єкти різних кластерів істотно відрізнялися. При цьому кожному об'єкту $x_i \in X^m$ задається номер кластеру u_i . Алгоритм кластеризації – це функція $a: X \rightarrow Y$, яка будь-якому об'єкту $x \in X$ ставить у відповідність номер кластера $u \in Y$. Множина Y в деяких випадках відома заздалегідь, проте частіше ставиться завдання визначити оптимальне число кластерів, з погляду деякого критерію якості кластеризації.

2.1.2.1. Завдання та етапи кластерного аналізу

Основні завдання кластерного аналізу:

- Розробка типології або класифікації;
- Дослідження корисних концептуальних схем групування об'єктів;
- Породження гіпотез на основі дослідження даних;
- Перевірка гіпотез або дослідження для визначення, чи дійсно групи, виділені тим чи іншим способом, присутні в наявних даних.

Незалежно від сфери застосування кластерного аналізу, він передбачає наступні етапи:

- Відбір вибірки для кластеризації;
- Визначення множини характеристик, за якими будуть оцінюватися об'єкти у вибірці;
- Обчислення значень міри схожості між об'єктами;
- Застосування одного з методів кластерного аналізу для створення груп схожих об'єктів;
- Перевірка достовірності результатів кластеризації.

2.1.2.2. Методи кластеризації

Оскільки поняття «кластеру» не може бути точно визначено, то це є однією з причин чому існує так багато різних методів кластеризації[9]. Але є і спільна риса — це об'єднання схожих об'єктів у групи. Однак, різні дослідники

використовують різні моделі кластерів і для кожної з цих моделей можуть бути застосовані різні алгоритми. Поняття кластера, які отримуються у різних алгоритмах, різняться властивостями. Розуміння цих «кластерних моделей» є ключовим для розуміння відмінностей між різними алгоритмами. Типовими кластерними моделями є:

- Моделі зв'язності. Наприклад, ієрархічна кластеризація або таксономія будуються на основі відстані між вузлами;
- Центроїдні моделі. Наприклад, метод К–середніх представляє кожен кластер єдиним усередненим вектором;
- Статистичні моделі. Кластери будуються ґрунтуючись на статистичних розподілах. Таких як багатовимірний нормальний розподіл з допомогою EM–алгоритму;
- Моделі засновані на щільності. Наприклад, в DBSCAN і в OPTICS кластери визначаються як зв'язані області відповідної щільності у просторі даних;
- Групові моделі. Деякі алгоритми не забезпечують вдосконалену модель для своїх результатів, а просто описують групування об'єктів;
- Графові моделі. Поняття кліки (така підмножина вершин, в якій кожна пара вершин з'єднана ребром) у графі слугує прототипом кластеру. Пом'якшення вимоги до повної зв'язності (тобто, частина ребер може бути відсутня) призводить до поняття відомого як квазі–кліка. Вони будуються алгоритмом HCS;
- Нейронні моделі. Найбільш відомою моделлю нейронної мережі з навчанням без учителя є нейронна мережа Кохонена. Ці моделі, як правило, можна охарактеризувати як схожі на одну або подібні якійсь з наведених вище моделей, включаючи моделі у підпросторах, коли нейронні мережі реалізують метод головних компонент або аналіз незалежних компонент.

2.1.2.3. Ієрархічна кластеризація

Ієрархічна кластеризація – метод кластерного аналізу, який намагається побудувати ієрархію кластерів. Стратегії побудови ієрархічної кластеризації діляться на два типи[15]:

- Агломераційні (об'єднувальні) – це підхід, в якому спочатку кожна точка має власний кластер, а далі пари кластерів об'єднуються при підйомі вгору по ієрархії.
- Розділювальні – це підхід, в якому спочатку всі точки знаходяться у єдиному кластері, потім відбувається рекурсивне розбиття їх на окремі кластери при русі вниз по ієрархії.

Результати ієрархічної кластеризації зазвичай відображають у вигляді дендрограми – графу без циклів побудованого на основі матриці близькості між об'єктами. Дендрограма дозволяє зобразити взаємні зв'язки між об'єктами з заданого переліку[16].

Для того, щоб вирішити, які кластери слід поєднувати або де кластер слід розділити, міра несхожості між множинами спостережень. У більшості методів ієрархічної кластеризації це досягається використанням відповідної метрики (міра відстані між парами спостережень) та критерію зв'язку, який визначає несхожість множин як функції парних відстаней спостережень у множинах.

Вибір відповідної метрики (відстані між об'єктами) впливатиме на форму кластерів, оскільки деякі елементи можуть бути близькими один до одного відповідно до однієї відстані, одна далекими, відповідно до іншої.

Найбільш розповсюдженні метрики для ієрархічної кластеризації[17] між точками a та b , де a_i , та b_i – координати точки a та b , відповідно, на i -вимірній площині:

- Евклідова відстань $\sqrt{\sum_i (a_i - b_i)^2}$
- Квадрат Евклідової відстані $\sum_i (a_i - b_i)^2$
- Відстань міських кварталів (Manhattan distance) $\sum_i |a_i - b_i|$

Критерій зв'язку визначає відстань між наборами спостережень як функцію парних відстаней між спостереженнями.

Найбільш розповсюджені критерії зв'язку для ієрархічної кластеризації між множинами A та B , d – обрана метрика:

- Метод одиночного зв'язку (метод найближчого сусіда):
$$\min\{d(a, b): a \in A, b \in B\}$$
- Метод повного зв'язку (метод дальнього сусіда):
$$\max\{d(a, b): a \in A, b \in B\}$$
- Метод Варда (Ward): як відстань між кластерами обирається приріст суми квадратів відстаней об'єктів до центрів кластерів, що отримується в результаті їх об'єднання. На кожному кроці алгоритму об'єднуються такі два кластери, які призводять до мінімального збільшення цільової функції, тобто внутрішньогрупової суми квадратів. Цей метод направлений на об'єднання близько розташованих кластерів і «прагне» створювати кластери малого розміру.

2.1.2.4. Оцінка якості кластеризації

Одна з проблем навчання без вчителя полягає в тому, що методи кластеризації формуватимуть групи, навіть якщо аналізований набір даних є повністю випадкову структуру.

Після створення кластерного рішення виникає питання наскільки воно статистично значиме. У теоретичному плані проблема перевірки адекватності кластеризації не вирішена, принаймні, без використання іншого виду аналізу або апіорного знання приналежності об'єктів до відповідних класів. В літературі пропонується багато методів та критеріїв оцінки якості результатів кластеризації, серед яких можна виділити декілька підходів[10]:

- Зовнішня перевірка, яка полягає в порівнянні підсумків кластерного аналізу з заздалегідь відомим результатом (тобто мітки кластерів відомі апіорі);

- Відносна перевірка, яка оцінює структуру кластерів, змінюючи різні параметри одного і того ж алгоритму (наприклад, число груп k);
- Внутрішня перевірка, яка використовує внутрішню інформацію процесу об'єднання в кластери (якщо зовнішня інформація відсутня).

2.1.2.4.1. Індекс Данна

Індекс Данна (DI) (введений Дж. К. Данном в 1974 р.) – це показник для оцінки алгоритмів кластеризації.[18] Це частина групи індексів дійсності, включаючи індекс Девіса–Болдіна або індекс Silhouette, оскільки це внутрішня схема оцінки, де результат базується на самих кластерних даних. Як і у всіх інших подібних індексах, мета полягає у визначенні компактних наборів кластерів з невеликою дисперсією між членами кластера, де представники різних кластерів досить далеко один від одного, порівняно з кластером всередині. Для заданого розподілу кластерів більш високий індекс Данна вказує на кращу кластеризацію. Одним з недоліків використання цього є обчислювальна вартість у міру збільшення кількості кластерів та розмірності даних.

2.1.2.4.2. Індекс Девіса–Болдіна

Індекс Девіса–Болдіна (індекс силуету) – це міра того, наскільки об'єкт схожий на власний кластер порівняно з іншими кластерами. Силует коливається від -1 до $+1$, де високе значення вказує на те, що об'єкт добре відповідає власному кластеру і погано відповідає сусіднім кластерам. Якщо більшість об'єктів мають високе значення, то конфігурація кластеризації вважається об'єктивною. Якщо багато точок мають низьке або негативне значення, то в конфігурації кластеризації може бути занадто багато або занадто мало кластерів.[11]

2.1.3. Регресійний аналіз

Регресійний аналіз – статистичний метод дослідження впливу однієї або декількох незалежних змінних X_1, X_2, \dots, X_n на залежну змінну Y .

При використанні лінійної регресії взаємозв'язок між даними моделюється за допомогою лінійних функцій, а невідомі параметри моделі оцінюються за вхідними даними. Подібно до інших методів регресійного аналізу лінійна регресія повертає розподіл умовної імовірності у в залежності від X , а не розподіл спільної імовірності Y та X , що стосується області мультиваріативного аналізу.

При розрахунках параметрів моделі лінійної регресії як правило застосовується метод найменших квадратів (МНК), але також можуть бути використані інші методи. Так само метод найменших квадратів може бути використаний і для нелінійних моделей. Тому МНК та лінійна регресія хоч і є тісно пов'язаними, але не є синонімами.[19]

2.1.3.1. Задачі регресійного аналізу

1. Визначення ступеня детермінованості варіації критеріальної (залежної) змінної предикторами (незалежними змінними);
2. Прогнозування значення залежної змінної за допомогою незалежної;
3. Визначення внеску окремих незалежних змінних у варіацію залежної;
4. Регресійний аналіз не можна використовувати для визначення наявності зв'язку між змінними, оскільки наявність такого зв'язку і є передумова для застосування аналізу;

2.1.3.2. Метод найменших квадратів

Метод найменших квадратів – метод знаходження наближеного розв'язку надлишково–визначеної системи. На практиці найчастіше використовується лінійний метод найменших квадратів, що використовується у випадку системи лінійних рівнянь. Зокрема важливим застосуванням у цьому випадку є оцінка параметрів у лінійній регресії, що широко застосовується у математичній статистиці і економетриці.[8]

2.1.3.3. Регресійні моделі

До регресійної моделі відносяться наступні параметри і змінні:

- Невідомі параметри, що позначаються як β , і які задають скаляри або вектор;

- Незалежні змінні, X ;
- Залежна змінна, Y ;
- Регресійна модель співвідносить Y до функції по X і β : $Y \approx f(X, \beta)$.

2.2. Застосовані методики

2.2.1. Багатовимірне шкалювання

Оскільки візуалізацію багатовимірного простору виконати важко, а іноді зовсім неможливо, використовують багатовимірне шкалювання, яке надає кращу виразність даних, ніж інші способи.

Ідея полягає в тому, щоб спроектувати дані багатовимірного простору на площину або тривимірний простір, які добре сприймаються людиною. Знаючи відстань між точками у просторі, треба розташувати точки на площині так, щоб відстані між ними були такі ж, як і у початковому багатовимірному просторі. Точніше, щоб відстані між точками на площині мали найменше відхилення від «правильних» відстаней у багатовимірному просторі.

Це дозволяє відразу виділити кластери у даних та найбільш віддалені точки від усіх інших.

2.2.2. Метод головних компонент

Ідея методу полягає в тому, щоб відібрати найбільш мінливі змінні, і тим самим, зменшити кількість змінних вцілому. Це один із способів зменшити розмірність даних і втратити при цьому найменшу кількість інформації. Тобто, якщо дано три змінні, що корелюють між собою, за допомогою метода головних компонент ми можемо замінити їх на дві і вони будуть описувати деякий відсоток дисперсії. При цьому кожна головна компонента обирається так, щоб описувати якомога більшу частину всієї дисперсії.[19]

Висновки

Для подальших розрахунків і аналізу показників використовується мова програмування R, оскільки вона створювалась з метою полегшення статистичних розрахунків і дозволяє оперувати такими поняттями як: вектор, таблиця, матриця, функція та ін. Її стандартна бібліотека містить багато функцій для розрахунку статистичних величин, також наявні функції для побудови графіків різних типів, побудови регресійних моделей(не тільки лінійних) та багато іншого. Крім цього існує багато сторонніх пакетів, написаних спільноту. Для їх простого встановлення і використання існує «Всеосяжна Архітектурна Мережа R»(The Comprehensive R Archive Network, CRAN), що є сховищем усіх бібліотек для R.

Для кластерного аналізу за відстань між об'єктами обрано Евклідову відстань – геометричну відстань між точками у багатовимірному просторі, оскільки вона найбільш «нейтрально» описує схожість між об'єктами, на відміну від іншого кандидату – манхеттенської відстані, яка зменшує вплив далеко розташованих об'єктів (викидів).

За відстань між кластерами обрано метод Варда, оскільки він добре розподіляє щільні шарові скупчення об'єктів на окремі кластери.

РОЗДІЛ 3

МОДЕЛЮВАННЯ

3.1. Загальна схема методики

Використана методика аналізу та оцінки включає в себе відбір даних, моделювання взаємозв'язку між показниками та моделювання угруповань (кластерів) серед показників за допомогою мови програмування R, побудову карт показників енергетичної безпеки для країн світу та аналіз сукупності показників енергетичної безпеки(рис. 3.1.).



Рис. 3.1. Загальна схема методики

3.2. Схема кластеризації з урахуванням географічного положення

Кластеризація з урахуванням географічного положення об'єкту включає в себе декілька етапів: розрахунок матриці відстаней між об'єктами та матриці географічних відстаней між об'єктами; нормалізація цих відстаней; коригування відстаней між об'єктами, використовуючи різні варіації поєднань відстаней; кластеризацію кожної варіації за допомогою одного з відомих методів; оцінку якості кластеризації.

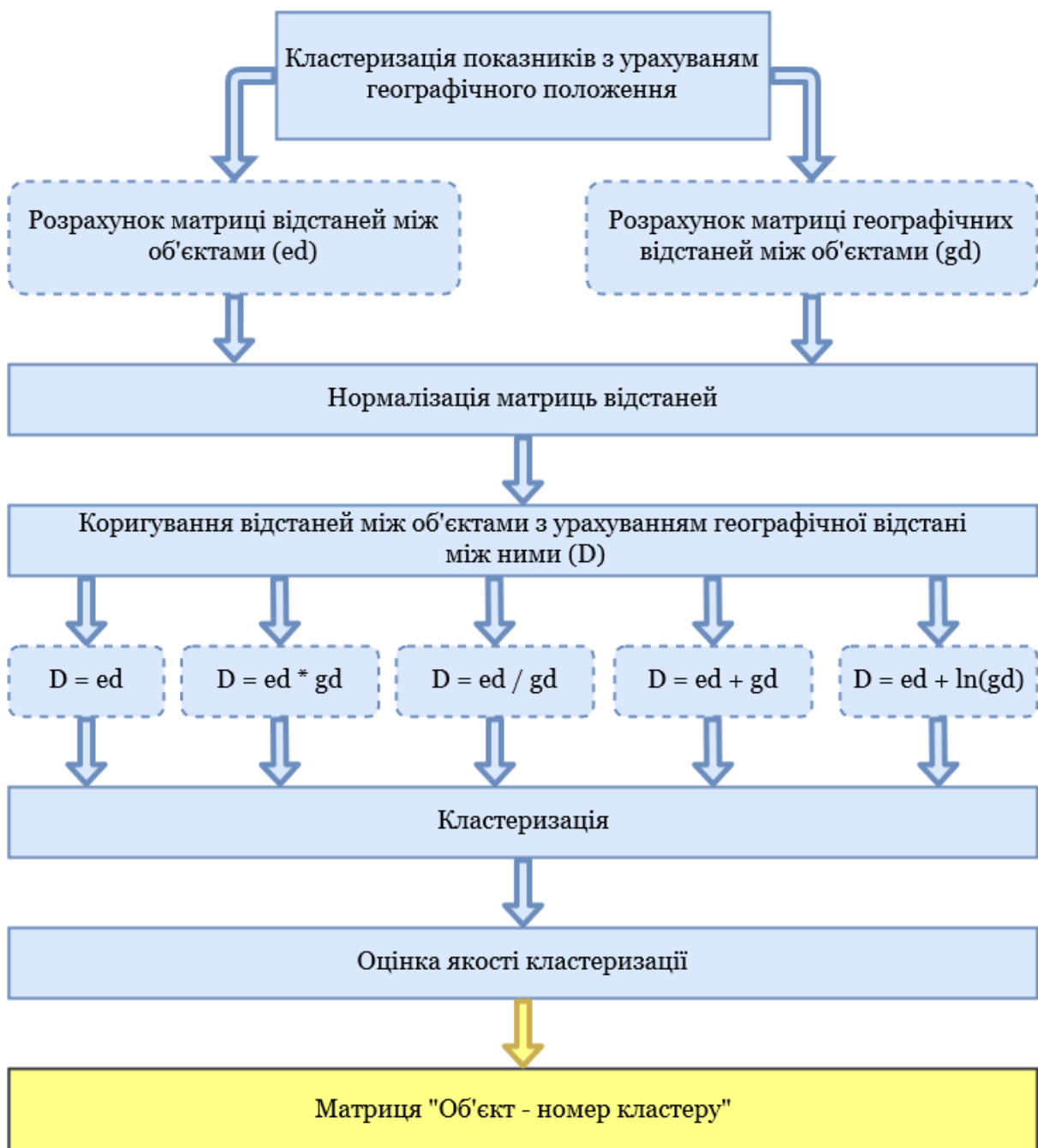


Рис. 3.2. Схема методики кластеризації з урахуванням географічного положення

Загальновідомі методи кластерного аналізу не враховують географічне положення об'єкту належним чином, навіть, якщо додати його як один із параметрів до моделі кластеризації. Тому новий підхід до кластеризації гео-даних дозволить точніше та гнучкіше розділяти сукупність об'єктів на групи за тією чи іншою ознакою.

Підхід, зображений на Рис. 3.2, базується на добре відомих алгоритмах кластеризації, що дозволяє легко інтегрувати його як до нової так і до існуючої моделі. Ідея підходу полягає не у створенні нового алгоритму розрахунку відстані між об'єктами чи кластерами, а саме у коригуванні цієї відстані.

Завдяки цьому коригуюча функція може бути будь-якою, а її вибір залишається за кінцевим користувачем методики, що додає підходу гнучкості та універсальності.

В залежності від розподілу об'єктів дослідження, що мають визначену географічну позицію, та від мети кластерного аналізу, коригуюча функція може як збільшувати чи зменшувати відстань між об'єктами, так і виконувати нелінійні перетворення, характерні для тієї чи іншої предметно області.

Існує одна умова для коригуючої функції: базова та географічна відстань повинні бути нормалізованими, тобто вимірюватись у однакових одиницях виміру.

Так само, як і для базової відстані між показниками (що не враховує географічну позицію), для відстані між кластерами можна використовувати будь-які методи розрахунку.

Тим самим методика вносить корективи саме у відстані між об'єктами, що дозволяє додати вплив географічного положення об'єкту до моделі кластеризації.

3.3. Моделювання взаємозв'язків між показниками

На першому етапі, для можливості продовження розрахунків, були відкинуті рядки таблиці, що мають хоча б один пропуск у даних.

Початкові дані виражені в різних величинах, тому було проведено нормалізацію даних для надання можливості їх порівняння.

Після цього за допомогою мови програмування R був проведений розвідувальний аналіз даних, що включає в себе розрахунки коефіцієнтів парної кореляції, багатовимірне шкалювання та аналіз головних компонент.

В табл. 3.1. наведені коефіцієнти парної кореляції показників.

Таблиця 3.1

Коефіцієнти парної кореляції показників

	OIL_CONS	GAS_CONS	COAL_CONS	NUC_GEN
OIL_CONS	1.00	0.86	0.68	0.84
GAS_CONS	0.86	1.00	0.39	0.84
COAL_CONS	0.68	0.39	1.00	0.40
NUC_GEN	0.84	0.84	0.40	1.00
NUC_CONS	0.84	0.84	0.40	1.00
HYDRO_GEN	0.69	0.45	0.89	0.45
HYDRO_CONS	0.69	0.45	0.89	0.45
SOLAR_GEN	0.84	0.59	0.88	0.61
SOLAR_CONS	0.84	0.59	0.88	0.61
WIND_GEN	0.90	0.67	0.84	0.72
WIND_CONS	0.90	0.67	0.84	0.72
OTHER_GEN	0.85	0.61	0.69	0.65
OTHER_CONS	0.85	0.61	0.69	0.65

Продовження таблиці 3.1

	NUC_CONS	HYDRO_GEN	HYDRO_CONS
OIL_CONS	0.84	0.69	0.69
GAS_CONS	0.84	0.45	0.45
COAL_CONS	0.40	0.89	0.89
NUC_GEN	1.00	0.45	0.45
NUC_CONS	1.00	0.45	0.45
HYDRO_GEN	0.45	1.00	1.00
HYDRO_CONS	0.45	1.00	1.00
SOLAR_GEN	0.61	0.81	0.81
SOLAR_CONS	0.61	0.81	0.81
WIND_GEN	0.72	0.82	0.82
WIND_CONS	0.72	0.82	0.82
OTHER_GEN	0.65	0.72	0.72

OTHER_CONS	0.65	0.72	0.72
------------	-------------	-------------	-------------

Продовження таблиці 3.1

	SOLAR_GEN	SOLAR_CONS	WIND_GEN
OIL_CONS	0.84	0.84	0.90
GAS_CONS	0.59	0.59	0.67
COAL_CONS	0.88	0.88	0.84
NUC_GEN	0.61	0.61	0.72
NUC_CONS	0.61	0.61	0.72
HYDRO_GEN	0.81	0.81	0.82
HYDRO_CONS	0.81	0.81	0.82
SOLAR_GEN	1.00	1.00	0.92
SOLAR_CONS	1.00	1.00	0.92
WIND_GEN	0.92	0.92	1.00
WIND_CONS	0.92	0.92	1.00
OTHER_GEN	0.86	0.86	0.89
OTHER_CONS	0.86	0.86	0.89

Продовження таблиці 3.1

	WIND_CONS	OTHER_GEN	OTHER_CONS
OIL_CONS	0.90	0.85	0.85
GAS_CONS	0.67	0.61	0.61
COAL_CONS	0.84	0.69	0.69
NUC_GEN	0.72	0.65	0.65
NUC_CONS	0.72	0.65	0.65
HYDRO_GEN	0.82	0.72	0.72
HYDRO_CONS	0.82	0.72	0.72
SOLAR_GEN	0.92	0.86	0.86
SOLAR_CONS	0.92	0.86	0.86
WIND_GEN	1.00	0.89	0.89
WIND_CONS	1.00	0.89	0.89
OTHER_GEN	0.89	1.00	1.00
OTHER_CONS	0.89	1.00	1.00

З огляду на таблицю 3.1. помітно дуже високу кореляцію показників між собою. Треба зазначити, що пари показників (споживання – виробництво) мають найбільшу кореляцію між собою – 100%.

Результати багатовимірного шкалювання(рис. 3.3) показують, що найбільш віддаленими країнами є США(73) та Китай(13), при цьому вони є абсолютними протилежностями. Після двох лідерів, наступними можна виділити такі країни як: Німеччина(24), Японія(35), Індія(28), Бразилія(9), а також: Франція(23), Росія(55), Великобританія(72), Канада(11), Італія(34), Південна Корея(60) та Іспанія(61)

Аналіз головних компонент дозволяє зробити наступні висновки. Вклад перших двох компонент пояснює 90%, трьох – 94%, чотирьох – 96%, загальної дисперсії(рис. 3.4). Проекція показників на фактор–план(рис. 3.5) показує існуючу кореляцію показників, що співпадає з результатами знаходження показників парної кореляції(табл. 3.1).

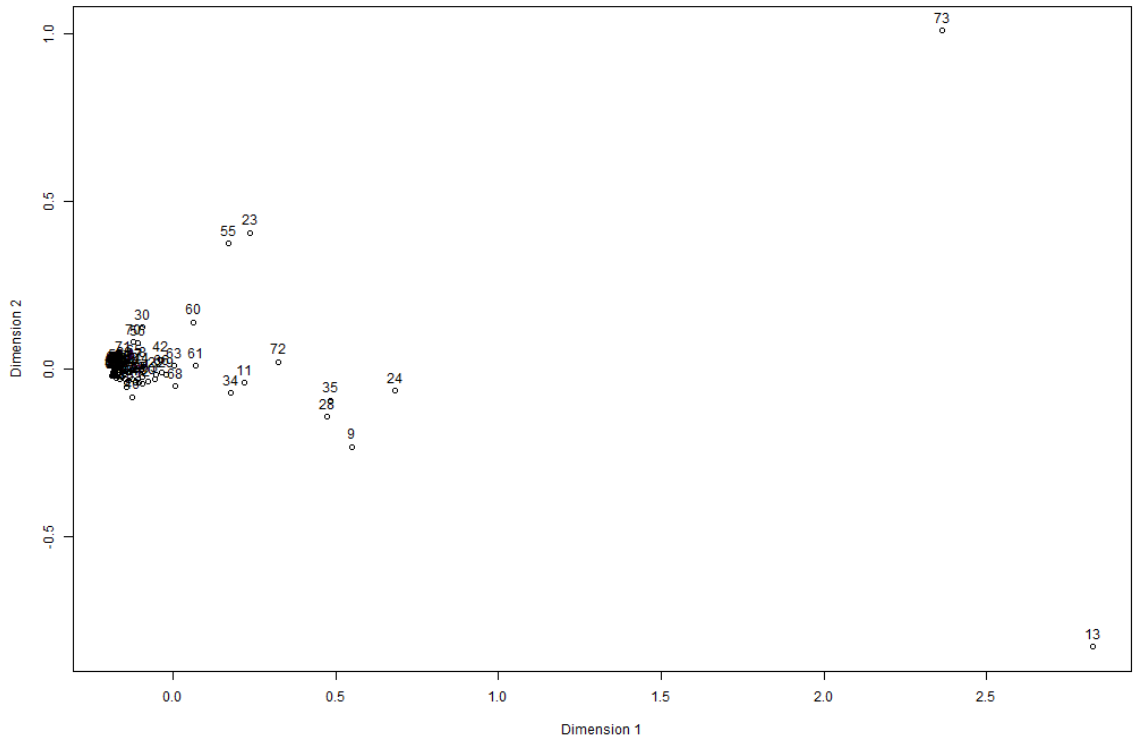


Рис. 3.3 Результати багатовимірного шкалювання

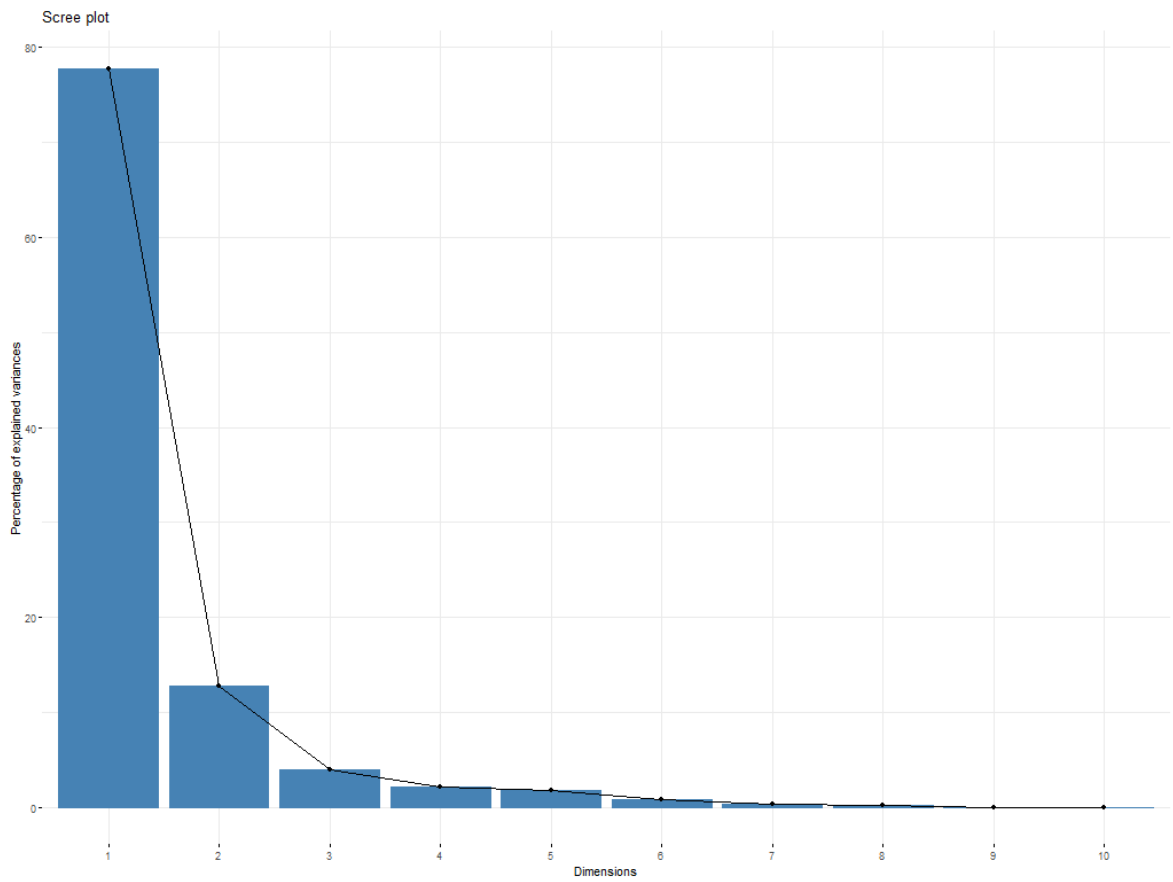


Рис. 3.4. Внесок компонент в дисперсію

Таблиця 3.2

Власні та накопиченні значення і їх вклад у дисперсію

	Власні значення	% всього	накоп. %
1	10.10	77.75	77.75
2	1.65	12.76	90.51
3	0.51	3.95	94.46
4	0.28	2.20	96.66
5	0.23	1.80	98.47
6	0.11	0.85	99.32
7	0.05	0.40	99.73
8	0.03	0.27	100
9	9.36E-21	7.20E-20	100
10	4.10E-21	3.16E-20	100
11	9.41E-22	7.24E-21	100
12	8.19E-23	6.30E-22	100
13	6.14E-23	4.72E-22	100

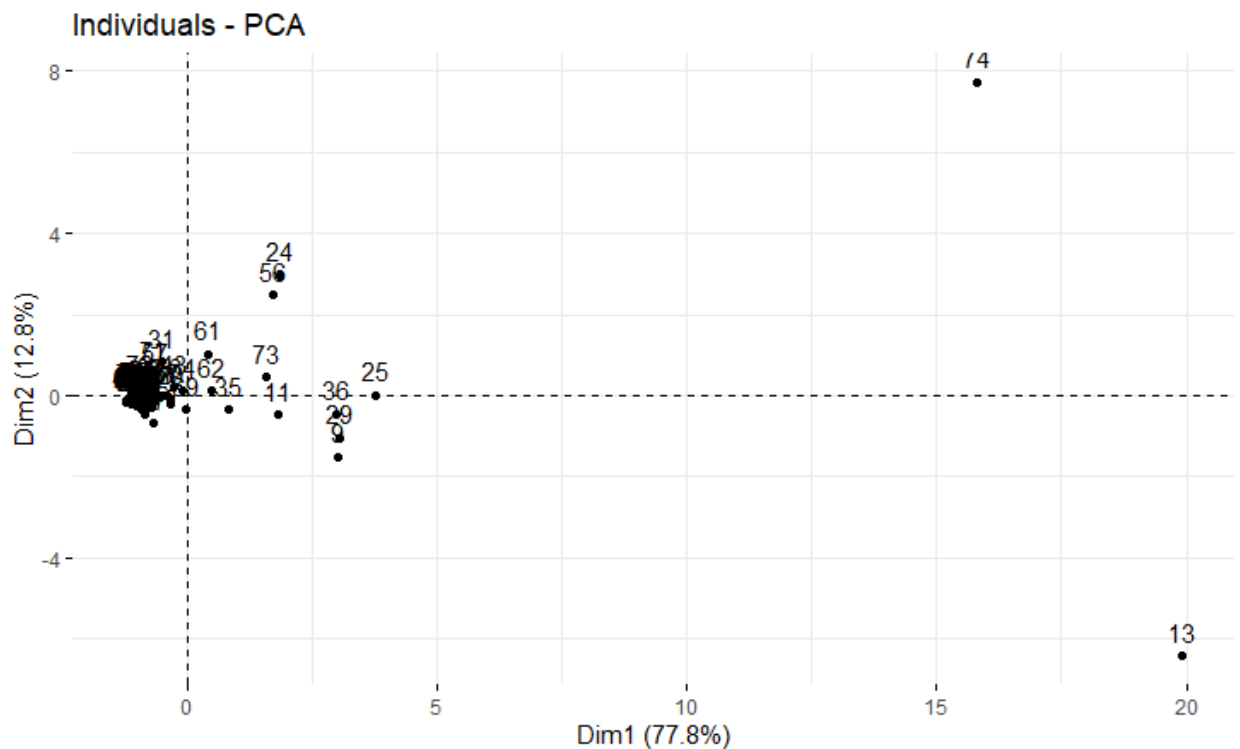


Рис.3.2. Проекція об'єктів на площину перших двох головних компонент

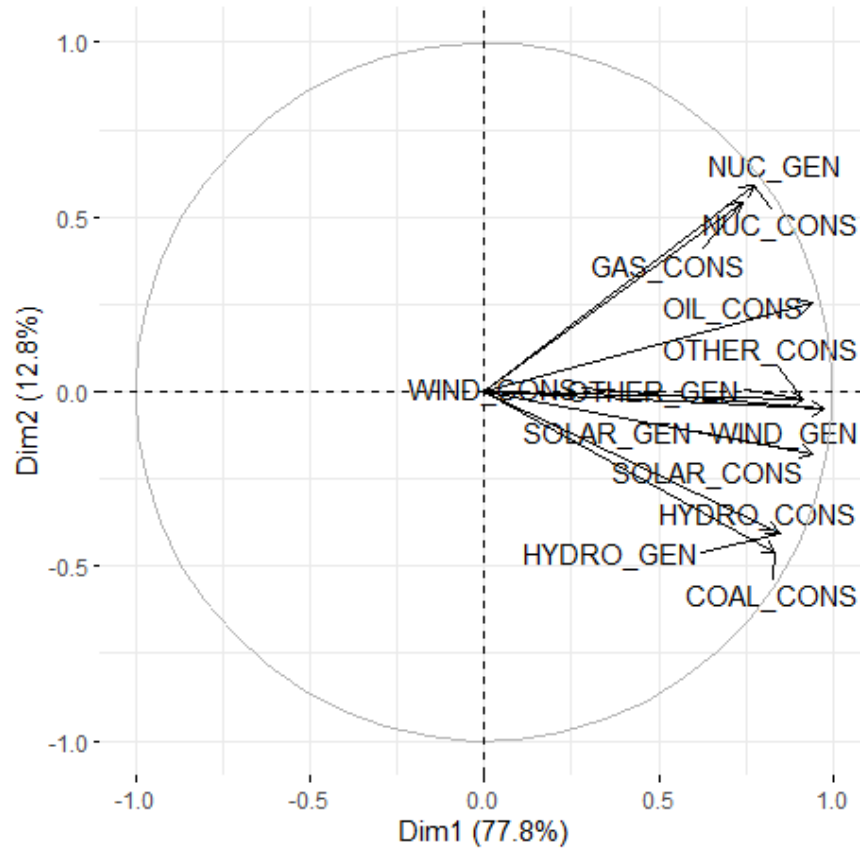


Рис. 3.3. Проекція показників на фактор–план

Розглянемо набір тринадцяти моделей функціональної залежності кожного з показників від сукупності усіх інших:

$$OIL_CONS = f_1(GAS_CONS, COAL_CONS, \dots, OTHER_CONS),$$

$$GAS_CONS = f_2(OIL_CONS, COAL_CONS, \dots, OTHER_CONS),$$

.....

$$OTHER_CONS = f_{13}(OIL_CONS, GAS_CONS, \dots, OTHER_GEN),$$

де f_i – лінійні функції.

Кожне з цих рівнянь може розглядатись окремо, а його параметри можуть бути знайдені за допомогою метода найменших квадратів. Після проведення розрахунків отримані наступні коефіцієнти при відповідних змінних(табл. 3.3). Значущі коефіцієнти виділені напівжирним.

Таблиця 3.3

Коефіцієнти моделей лінійної регресії

f	A	OIL_CONS	GAS_CONS	COAL_CONS	NUC_GEN	NUC_CONS
f_1	3.011552		0.467315	0.117473	0.198995	
f_2	6.395045	1.00899		-0.21256	0.144732	
f_3	3.933877	0.453344	-0.37991		-0.41304	
f_4	-2.37E-10	5.84E-12	-1.29E-12	-2.51E-12		0.38
f_5	6.25E-10	-1.54E-11	3.39E-12	6.60E-12	2.631579	
f_6	2.74E-10	9.09E-12	-8.25E-12	5.25E-12	2.65E-12	
f_7	-7.20E-10	-2.39E-11	2.17E-11	-1.38E-11	-6.98E-12	
f_8	-1.19E-10	-4.78E-12	5.74E-12	-5.15E-12	-8.59E-13	
f_9	3.13E-10	1.26E-11	-1.51E-11	1.36E-11	2.26E-12	
f_{10}	2.47E-10	3.43E-12	2.28E-12	-8.69E-12	-1.10E-11	
f_{11}	-6.49E-10	-9.02E-12	-6.00E-12	2.29E-11	2.91E-11	
f_{12}	-1.19E-10	-9.06E-13	4.30E-12	-3.56E-12	-2.54E-12	
f_{13}	3.12E-10	2.38E-12	-1.13E-11	9.37E-12	6.69E-12	

Продовження табл. 3.3

f	HUDRO_GEN	HYDRO_CONS	SOLAR_GEN	SOLAR_CONS	WIND_GEN
f_1	-0.06939		0.264702		0.30029
f_2	0.140011		0.399876		-0.13563
f_3	0.626136		5.067616		1.723665
f_4	-2.65E-12		1.10E-11		1.62E-11
f_5	6.96E-12		-2.89E-11		-4.26E-11
f_6		0.38	-5.68E-11		-3.53E-11
f_7	2.631579		1.50E-10		9.28E-11
f_8	1.04E-12			0.38	1.68E-11
f_9	-2.73E-12		2.631579		-4.41E-11
f_{10}	2.97E-12		6.03E-11		
f_{11}	-7.81E-12		-1.59E-10		2.631579
f_{12}	5.52E-12		6.33E-11		-8.63E-12
f_{13}	-1.45E-11		-1.67E-10		2.27E-11

Продовження табл. 3.3

f	WIND_CONS	OTHER_GEN	OTHER_CONS
f_1		1.775173	
f_2		-2.09232	
f_3		-5.80697	
f_4		-2.01E-11	
f_5		5.28E-11	
f_6		6.19E-11	
f_7		-1.63E-10	
f_8		-1.06E-10	
f_9		2.79E-10	
f_{10}	0.38	-6.40E-11	
f_{11}		1.68E-10	
f_{12}			0.38
f_{13}		2.631579	

Оцінка множинного коефіцієнту кореляції R , коефіцієнту детермінації R^2 , критеріїв Фішера F і вірогідності прийняття гіпотези $H_0: R = 0$ про незначущість множинного коефіцієнту кореляції приведені в табл. 3.4. Гіпотеза H_0 для усіх тринадцяти моделей відкидається і множинні коефіцієнти кореляції значно відрізняються від нуля. Усі моделі адекватно описують дані.

Таблиця 3.4

Характеристики моделей лінійної регресії

Модель	R	R^2	F	p
f_1	0.97325	0.947215	174.3194	0
f_2	0.923123	0.852156	55.99192	0
f_3	0.967069	0.935222	140.2475	0
f_4	1	1	6.38E+22	0
f_5	1	1	6.38E+22	0
f_6	1	1	4.82E+22	0
f_7	1	1	4.82E+22	0
f_8	1	1	1.02E+21	0
f_9	1	1	1.02E+21	0
f_{10}	1	1	4.40E+21	0
f_{11}	1	1	4.40E+21	0
f_{12}	1	1	4.48E+20	0
f_{13}	1	1	4.48E+20	0

Можна виділити наступні існуючі залежності показників:

$$OIL_CONS = f_1(GAS_CONS + NUC_GEN + OTHER_GEN),$$

$$GAS_CONS = f_2(OIL_CONS + COAL_CONS + OTHER_GEN)$$

$$COAL_CONS = f_3(GAS_CONS + NUC_GEN + \\ + HYDRO_GEN + WIND_GEN + OTHER_GEN)$$

$$NUC_GEN = f_4(NUC_CONS)$$

$$NUC_CONS = f_5(NUC_GEN)$$

$$HYDRO_GEN = f_6(HYDRO_CONS + WIND_GEN)$$

$$HYDRO_CONS = f_7(HYDRO_GEN + WIND_GEN)$$

$$SOLAR_GEN = f_8(SOLAR_CONS + OTHER_GEN)$$

$$SOLAR_CONS = f_9(SOLAR_GEN + OTHER_GEN)$$

$$WIND_GEN = f_{10}(WIND_CONS + COAL_CONS + NUC_GEN)$$

$$WIND_CONS = f_{11}(WIND_GEN + COAL_CONS + NUC_GEN)$$

$$OTHER_GEN = f_{12}(OTHER_CONS)$$

$$OTHER_CONS = f_{13}(OTHER_GEN)$$

З Рис. 3.6. видно, що виробництво та споживання електроенергії, добутої з традиційних джерел енергії (нафта, газ, вугілля, ядерна енергія) на сьогоднішні день залишається основною складовою частиною всесвітньої енергетики. Відновлювальні джерела енергії набирають популярність, що особливо помітно серед гідроенергетики та вітряної енергетики, які добре інтегрувалися у всесвітнє виробництво та споживання електроенергії. В той же час сонячна та інша альтернативна енергетика (геотермальні джерела, біомаса та інші) знаходиться на стадії розвитку, що частково пов'язано з низьким ККД технологічних пристроїв, що генерують електроенергії з цих відновлювальних джерел.

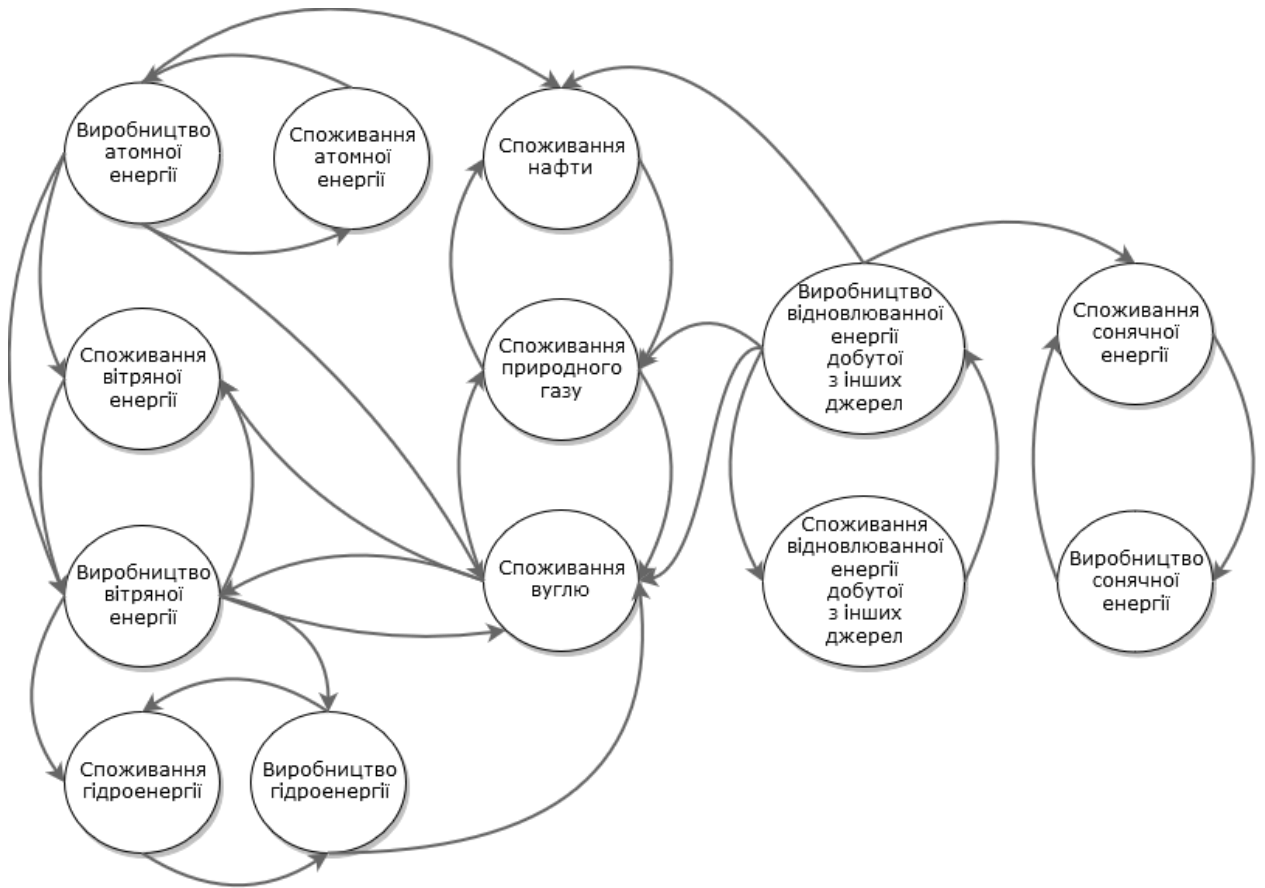


Рис. 3.6. Когнітивна мапа взаємозв'язків між показниками

3.4. Початкові дані та показники

3.4.1. Показники

За даними 2018 року для аналізу були обрані показники споживання природних копалин:

1. Споживання нафти (мільйонів тонн);
2. Споживання природного газу (мільярдів кубічних метрів);
3. Споживання вуглю (мільйонів тонн);

Показники споживання та виробництва електроенергії, добутої з різних джерел:

4. Споживання ядерної електроенергії (ТВт–год.);
5. Виробництво ядерної електроенергії (ТВт–год.);
6. Споживання гідро–електроенергії (ТВт–год.);
7. Виробництво гідро–електроенергії (ТВт–год.);
8. Споживання сонячної електроенергії (ТВт–год.);
9. Виробництво сонячної електроенергії (ТВт–год.);
10. Споживання вітряної електроенергії (ТВт–год.);
11. Виробництво вітряної електроенергії (ТВт–год.);
12. Споживання електроенергії, добутої з інших відновлювальних джерел (ТВт–год.);
13. Виробництво електроенергії, добутої з інших відновлювальних джерел (ТВт–год.);

Показники зростання виробництва та споживання електроенергії добутої з різних джерел у відсотках, порівняно з 2013 роком:

14. Зростання споживання ядерної електроенергії;
15. Зростання виробництва ядерної електроенергії;
16. Зростання споживання гідро–електроенергії;
17. Зростання виробництва гідро–електроенергії;

- 18.Зростання споживання сонячної електроенергії;
- 19.Зростання виробництва сонячної електроенергії;
- 20.Зростання споживання вітряної електроенергії;
- 21.Зростання виробництва вітряної електроенергії;
- 22.Зростання споживання електроенергії, добутої з інших відновлювальних джерел;
- 23.Зростання виробництва електроенергії, добутої з інших відновлювальних джерел.

3.4.2. Початкові дані

Початкові дані моделювання – чисельні показники, що представляють собою таблицю «об’єкт–ознака» розмірністю 78x23. Позначимо:

1. OIL_CONS – споживання нафти;
2. GAS_CONS – споживання природного газу;
3. COAL_CONS – споживання вугілля;
4. NUC_CONS – споживання ядерної енергії;
5. NUC_GEN – виробництво ядерної енергії;
6. HYDRO_CONS – споживання гідро–електроенергії;
7. HYDRO_GEN – виробництво гідро–електроенергії;
8. SOLAR_CONS – споживання сонячної електроенергії;
9. SOLAR_GEN – виробництво сонячної електроенергії;
- 10.WIND_CONS – споживання вітряної електроенергії;
- 11.WIND_GEN – виробництво вітряної електроенергії;
- 12.OTHER_CONS – споживання електроенергії, добутої з інших відновлювальних джерел
- 13.OTHER_GEN – виробництво електроенергії, добутої з інших відновлювальних джерел;

- 14.NUC_CONS_GROW – зростання споживання ядерної електроенергії;
- 15.NUC_GEN_GROW – зростання виробництва ядерної електроенергії;
- 16.HYDRO_CONS_GROW – зростання споживання гідро–електроенергії;
- 17.HYDRO_GEN_GROW – зростання виробництва гідро–електроенергії;
- 18.SOLAR_CONS_GROW – зростання споживання сонячної електроенергії;
- 19.SOLAR_GEN_GROW – зростання виробництва сонячної електроенергії;
- 20.WIND_CONS_GROW – зростання споживання вітряної електроенергії;
- 21.WIND_GEN_GROW – зростання виробництва вітряної електроенергії;
- 22.OTHER_CONS_GROW – зростання споживання електроенергії, добутої з інших відновлювальних джерел;
- 23.OTHER_GEN_GROW – зростання виробництва електроенергії, добутої з інших відновлювальних джерел.

3.5. Моделювання угруповань серед показників

Моделювання угруповань серед показників виконано за допомогою метода ієрархічної кластеризації для показників зростання виробництва відновлювальної енергії протягом п'яти років (2013 – 2018). Обрано наступні показники: зростання виробництва гідро–електроенергії, зростання виробництва сонячної електроенергії, зростання виробництва вітряної електроенергії та зростання виробництва електроенергії, добутої з інших відновлювальних джерел.

На Рис. 3.7 – 3.11 представлені дендрограми кластеризацій для різних варіацій коригування відстаней між об'єктами на значення географічної

відстані між цими об'єктами. Червоною лінією обведено кластери об'єктів, що утворилися в результаті кластеризації.

З Рис. 3.7, що відображає кластеризацію, яка не враховує географічну відстань між об'єктами, добре видно, що утворені кластери не моделюють реальну ситуацію – лише деякі країни виділені в окремі кластери, а усі інші – згруповані в одному великому кластері.

Інша ситуація на продемонстрована Рис. 3.8, де початкова відстань між об'єктами була помножена на географічну відстань. За таких умов добре виділяється чотири кластери країн та окремий кластер з Марокко (44) – країною, яка продемонструвала колосальне зростання виробництва відновлювальної електроенергії протягом періоду 2013 – 2018 років. Така ситуація можлива завдяки успішному будівництву та введенню в експлуатацію СЕС Уарзаат[12] протягом 2013 – 2016 років, що значно підвищило виробництво відновлювальної електроенергії у цьому регіоні.

Якщо спробувати розділити початкову відстань між об'єктами на стандартизовану географічну відстань (Рис. 3.9), то ситуація стає схожою на модель з Рис 3.7 – утворюється один великий кластер об'єктів, що включає в себе більшу частину усієї вибірки країн.

Краща ситуація відображена на Рис 3.10, 3.11, де до початкової відстані між об'єктами додається стандартизована географічна відстань та натуральний логарифм стандартизованої географічної відстані відповідно. Добре виділяються чотири кластери країн та окремий кластер з Марокко (44), як і на Рис. 3.8, однак відстані між об'єктами всередині кластерів більш помітні, що дозволяє розділити вибірку на більшу кількість кластерів, якщо це необхідно. При додаванні натурального логарифму стандартизованої географічної відстані до початкової відстані між об'єктами, результуюча відстань між об'єктами всередині кластерів стає меншою, що утворює більш згруповані кластери.

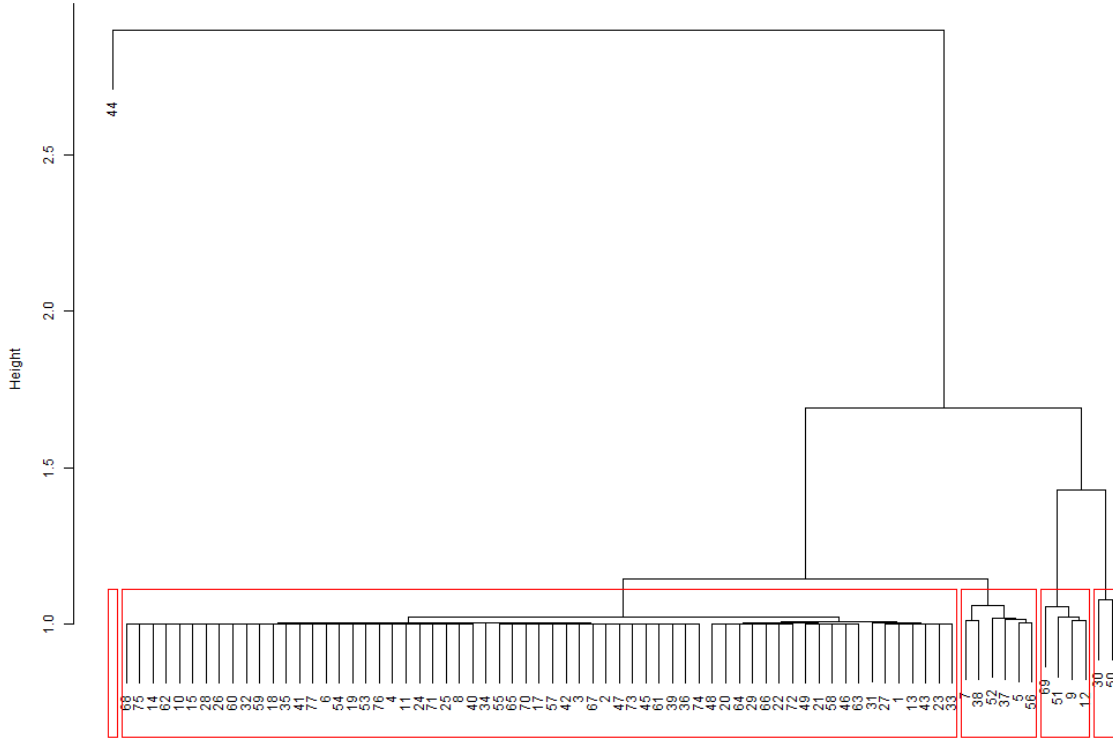


Рис. 3.7. Дендрограма ієрархічної кластеризації для матриці відстаней, що не враховує географічну позицію об'єктів

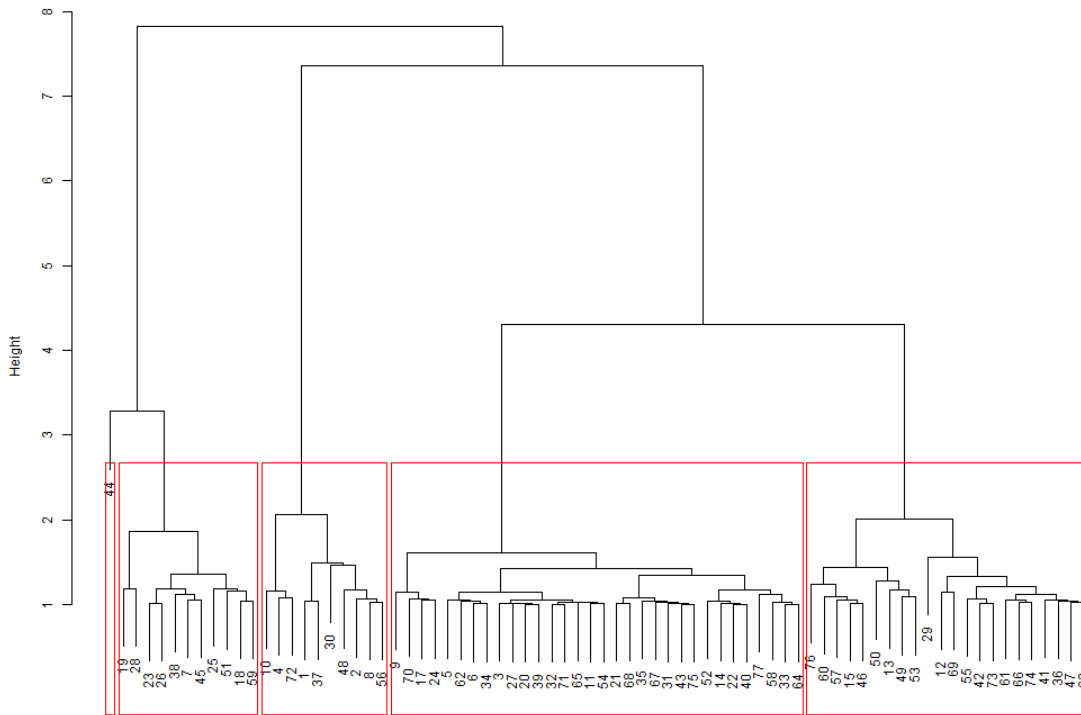


Рис. 3.8. Дендрограма ієрархічної кластеризації для відстані між об'єктами помноженої на стандартизовану географічну відстань

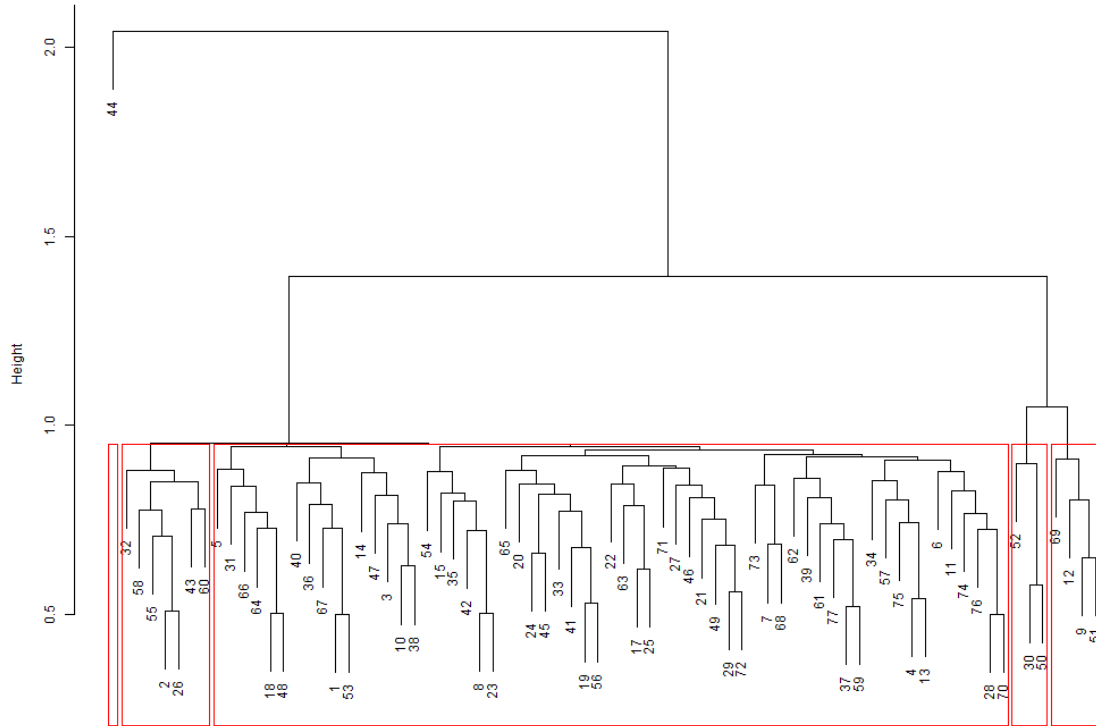


Рис. 3.9. Дендрограма ієрархічної кластеризації для відстані між об'єктами поділеної на стандартизовану географічну відстань

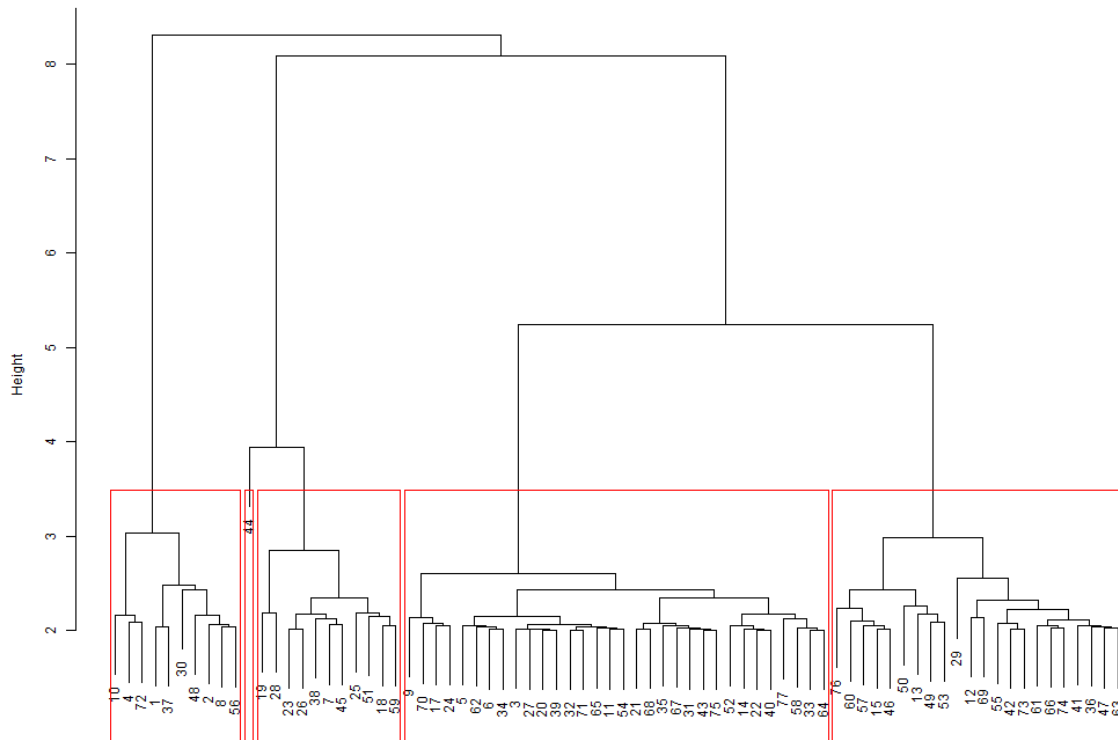


Рис. 3.10. Дендрограма ієрархічної кластеризації для відстані між об'єктами з додаванням стандартизованої географічної відстані

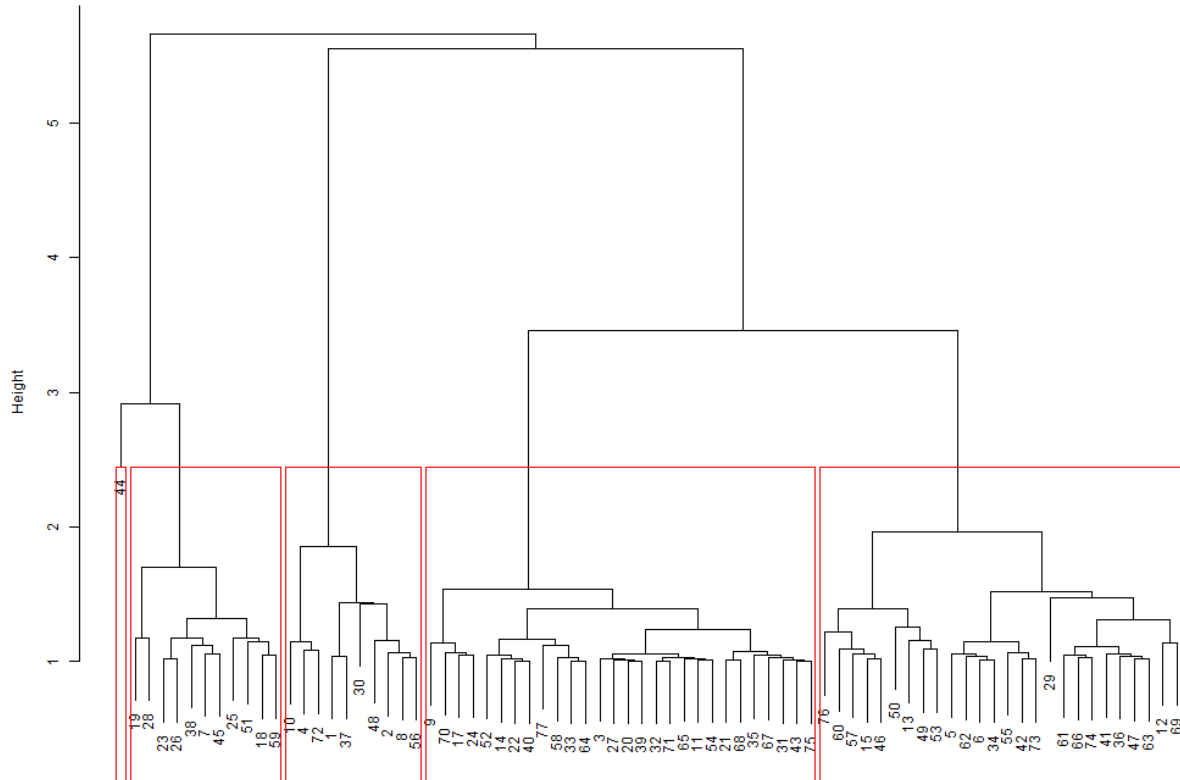


Рис. 3.11. Дендрограма ієрархічної кластеризації для відстані між об'єктами з додаванням натурального логарифма від стандартизованої географічної відстані

За результатами кластеризації (Рис. 3.12 – 3.16), що відображені на карті світу можна виділити найбільш вдалі моделі:

- Модель з додаванням стандартизованої географічної відстані між об'єктами (Рис. 3.15). Індекс Дунна = 0.785.
- Модель з додаванням натурального логарифма стандартизованої географічної відстані між об'єктами (Рис. 3.16). Індекс Дунна = 0.60.
- Модель з добутком стандартизованої географічної відстані та початковою відстанню між об'єктами (Рис. 3.13). Індекс Дунна = 0.62.

На прикладі моделі з Рис. 3.15 можна виділити наступні кластери:

1. Країни з найбільшим приростом виробництва сонячної та вітряної електроенергії;

2. Країни в яких спостерігається стабільне (помірне) зростання виробництва відновлювальної електроенергії;
3. Країни з найбільшим рівнем зростання виробництва відновлювальної електроенергії;
4. Країни з розвиненим виробництвом вітряної електроенергії, які успішно розвивають сонячну енергетику та експериментують з видобуванням електроенергії з інших джерел відновлювальної енергії;
5. Країни, які знаходяться поза світовими тенденціями, в даному випадку – це Марокко, де у 2013–му році почали будівництво СЕС, а у 2016–му ввели в експлуатацію одну з найбільших у світі СЕС.

Кластеризація країн світу за рівнем зростання виробництва відновлювальної енергії без урахування географічного положення країни

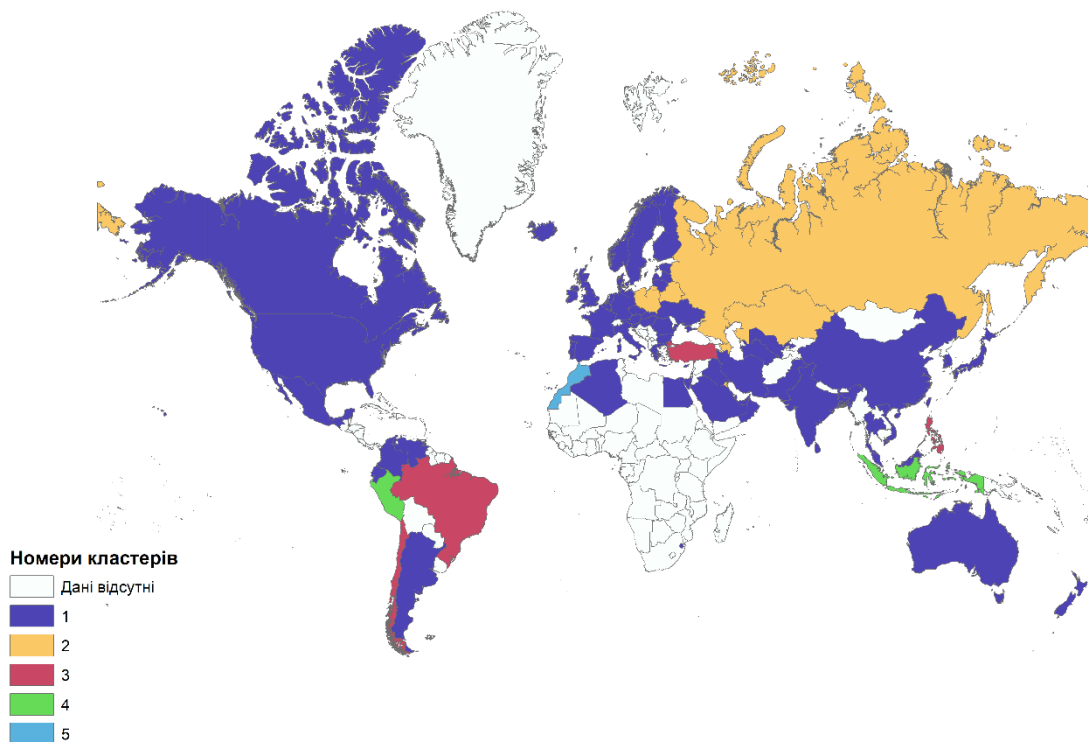


Рис. 3.12. Кластеризація країн світу за рівнем зростання виробництва відновлювальної енергії без урахування географічного положення країни

Кластеризація країн світу за рівнем зростання виробництва відновлювальної енергії з урахуванням географічного положення країни (добуток)

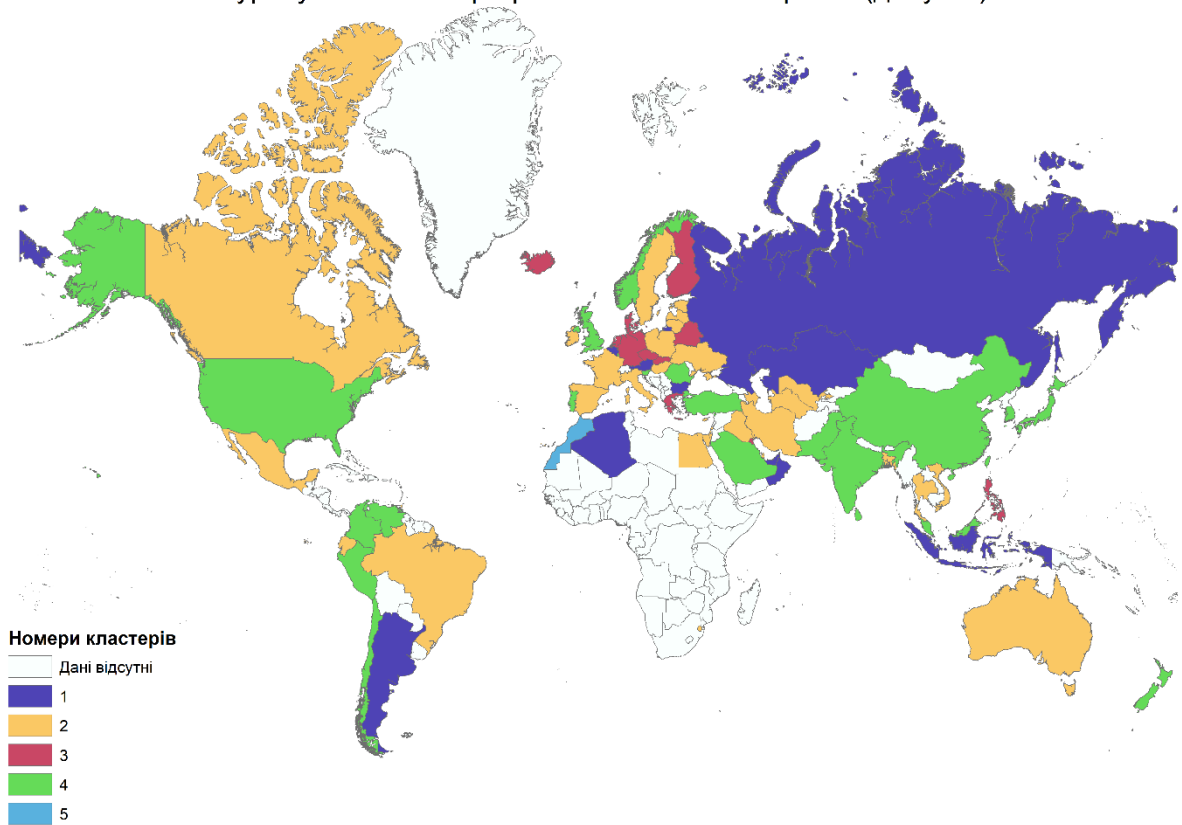


Рис. 3.13. Кластеризація країн світу за рівнем зростання виробництва відновлювальної енергії з урахуванням географічного положення країни (добуток)

Кластеризація країн світу за рівнем зростання виробництва відновлювальної енергії з урахуванням географічного положення країни (ділення)

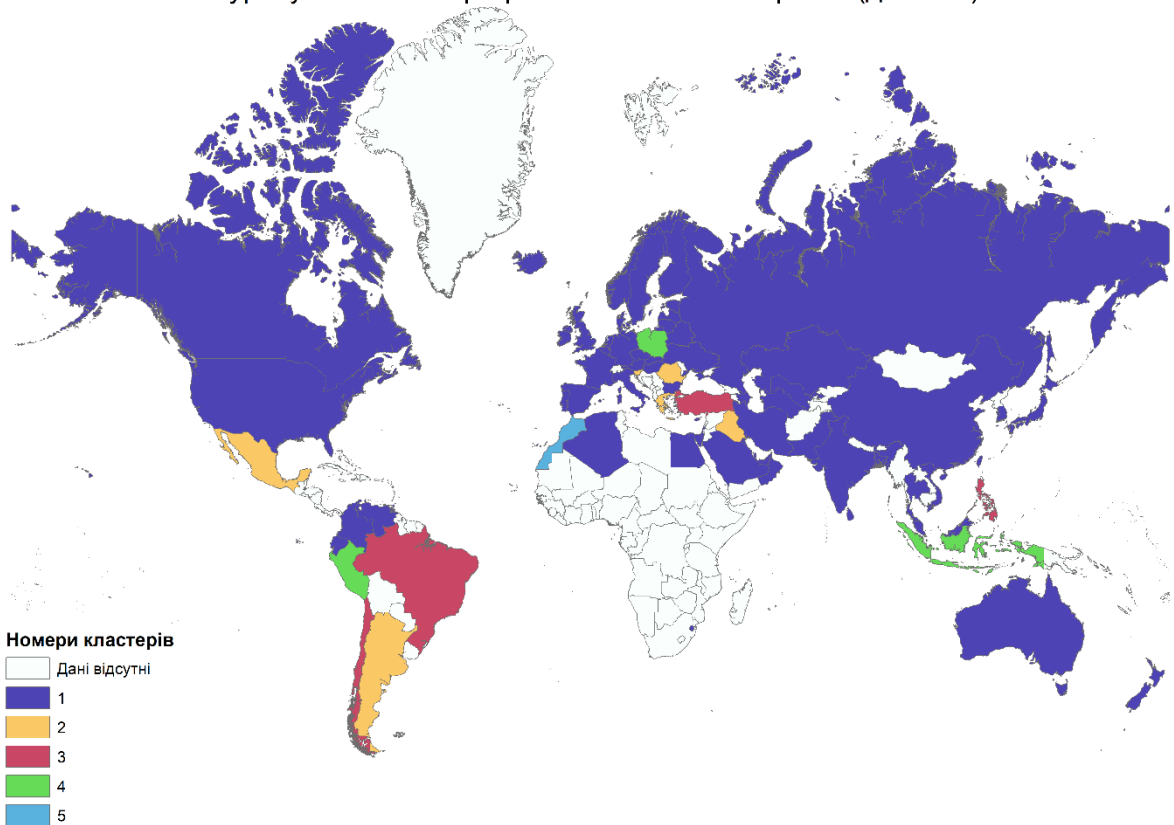


Рис. 3.14. Кластеризація країн світу за рівнем зростання виробництва відновлювальної енергії з урахуванням географічного положення країни (ділення)

Кластеризація країн світу за рівнем зростання виробництва відновлювальної енергії з урахуванням географічного положення країни (додавання)

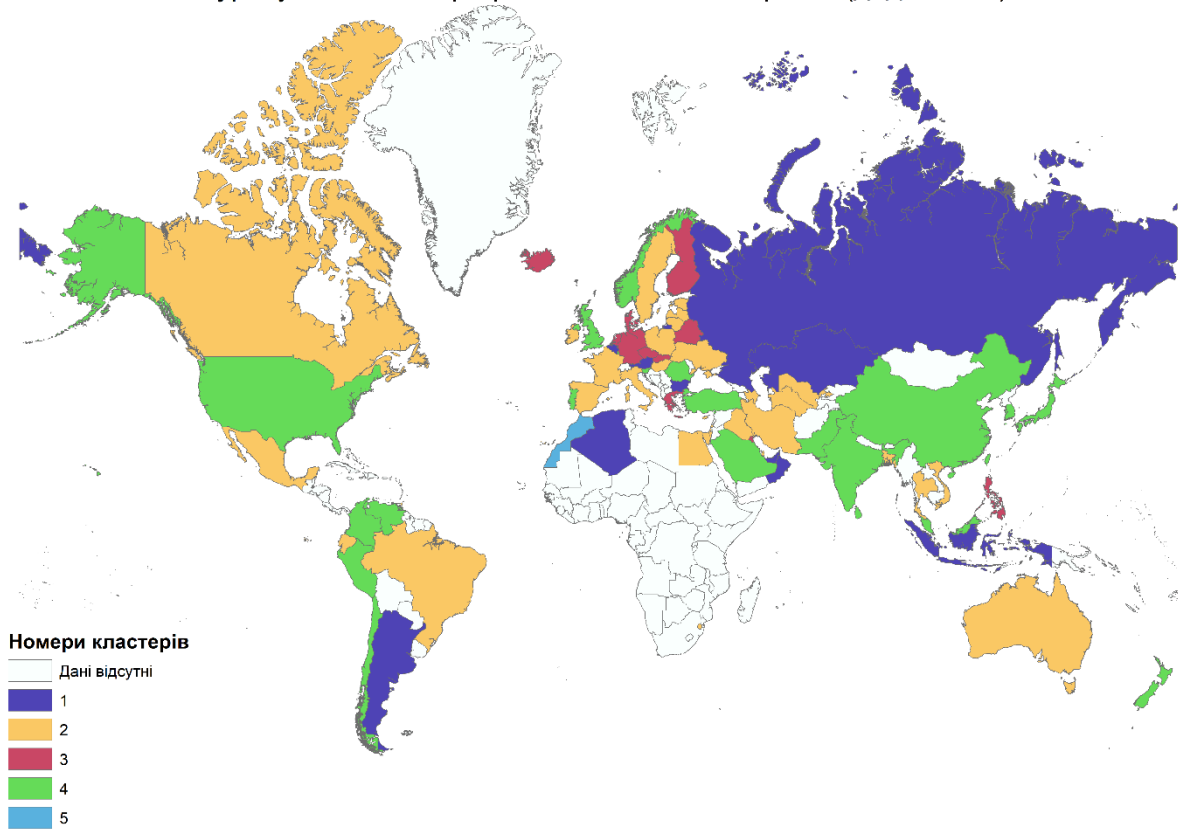


Рис. 3.15. Кластеризація країн світу за рівнем зростання виробництва відновлювальної енергії з урахуванням географічного положення країни (додавання)

Кластеризація країн світу за рівнем зростання виробництва відновлювальної енергії з урахуванням географічного положення країни (додавання логарифму)

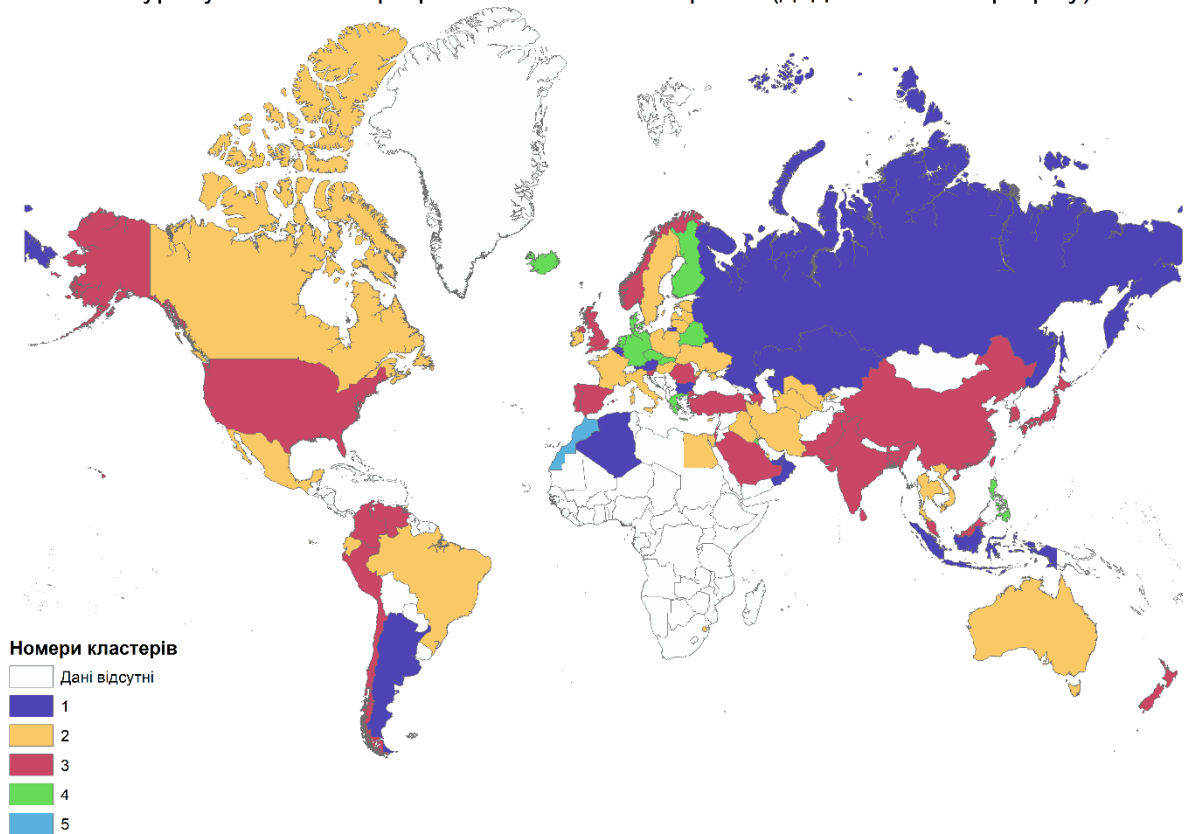


Рис. 3.16. Кластеризація країн світу за рівнем зростання виробництва відновлювальної енергії з урахуванням географічного положення країни (додавання логарифму)

3.6. Відображення показників зростання виробництва та споживання електроенергії на карті світу

На Рис 3.17 – 3.22 відображені дані зростання виробництва та споживання електроенергії, добутої з різних джерел, за період у п'ять років (2013 – 2018) у відсотках.

З Рис. 3.17 та 3.18 добре помітно, що більшість країн поступово відмовляються від атомної електроенергії на користь альтернативних джерел.

Сектор гідроенергетики (Рис. 3.19) залишається більш–менш стабільним. Деякі країни – Білорусь, Китай, Казахстан – укріпили свої позиції в цьому секторі, а в таких країнах, як Марокко, Іран, Ірак, Україна та деяких країнах Європи, виробництво гідро–електроенергії зменшилось, що може бути

пов'язано з великими витратами на обслуговування гідроелектростанцій, побудованих у минулому сторіччі.

Сонячна енергетика (Рис 3.20) демонструє стрімке зростання протягом минулих п'яти років. Такі показники пов'язані як і з розробкою сучасних фотоелементів з високим ККД, так і з форм-фактором сонячних панелей, який дозволяє розміщувати електростанції будь-яких розмірів у майже бідь-якому куточку світу. Не менш важливим є і політичний аспект проблеми: все більше і більше країн впроваджує пільги на «зелену» енергетику, що добре позначається на розвитку цього сектору.

Вітряна енергетика (Рис. 3.21) добре закріпилася майже у всіх країнах світу. Спостерігається значне зростання потужностей виробництва в цьому секторі.

Інша альтернативна енергетика (Рис. 3.22) демонструє помірне зростання, проте все більше країн світу звертає увагу на розробку альтернативних відновлювальних джерел енергії. Цей сектор залишається менш дослідженим, ніж усі інші, однак інтерес країн світу до нього підштовхує його розвиток.

Загальносвітові тенденції демонструють поступову відмову від традиційних джерел енергії на користь нових, відновлювальних. Крім цього, відновлювальні джерела енергії є екологічно чистими і перехід до використання саме таких джерел добре позначається на екологічній ситуації у світі.

Рівень зростання споживання атомної електроенергії

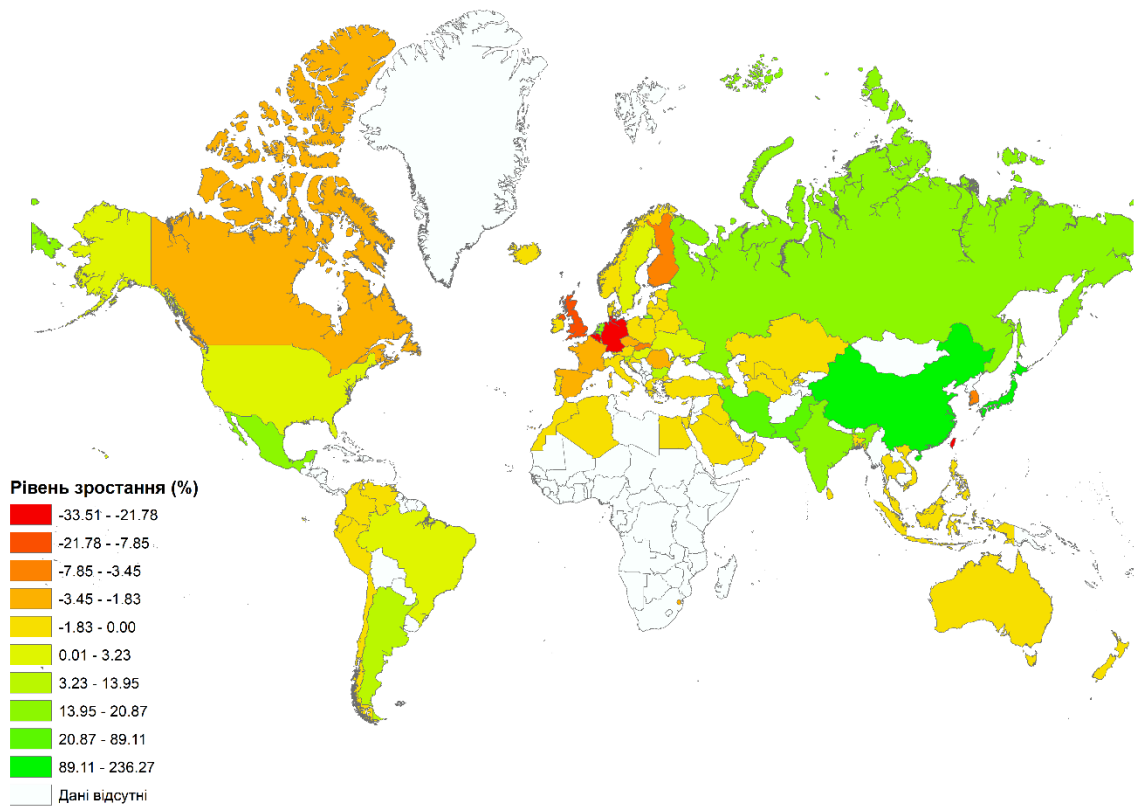


Рис. 3.17. Рівень зростання споживання атомної електроенергії у відсотках

Рівень зростання виробництва атомної електроенергії

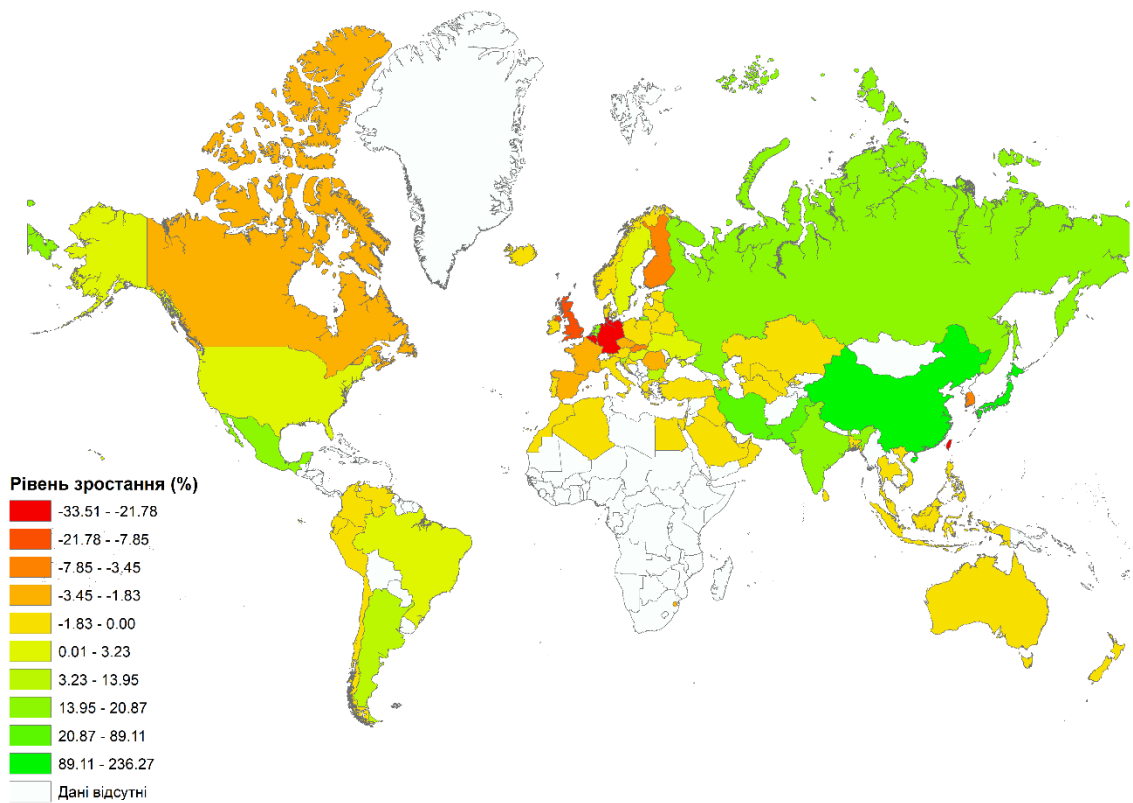


Рис. 3.18. Рівень зростання виробництва атомної електроенергії у відсотках

Рівень зростання виробництва гідроелектроенергії

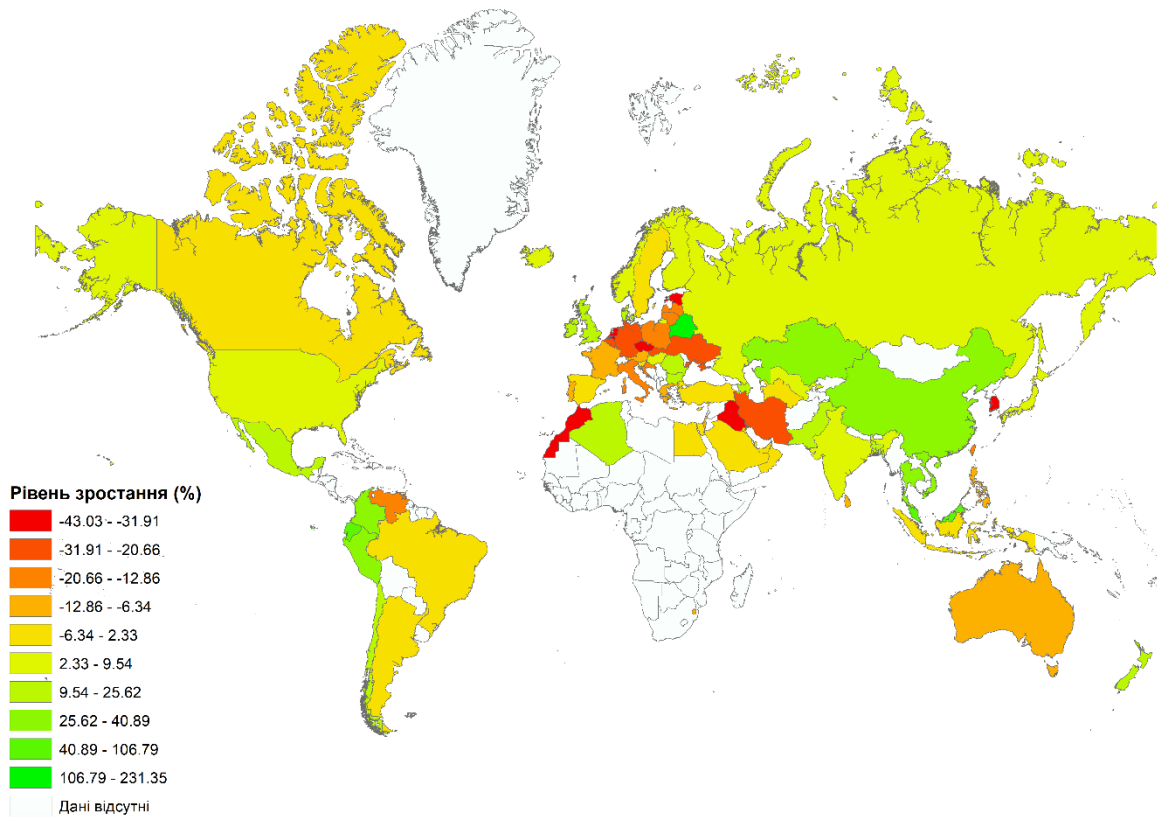


Рис. 3.19. Рівень зростання виробництва гідро–електроенергії у відсотках

Рівень зростання виробництва сонячної електроенергії

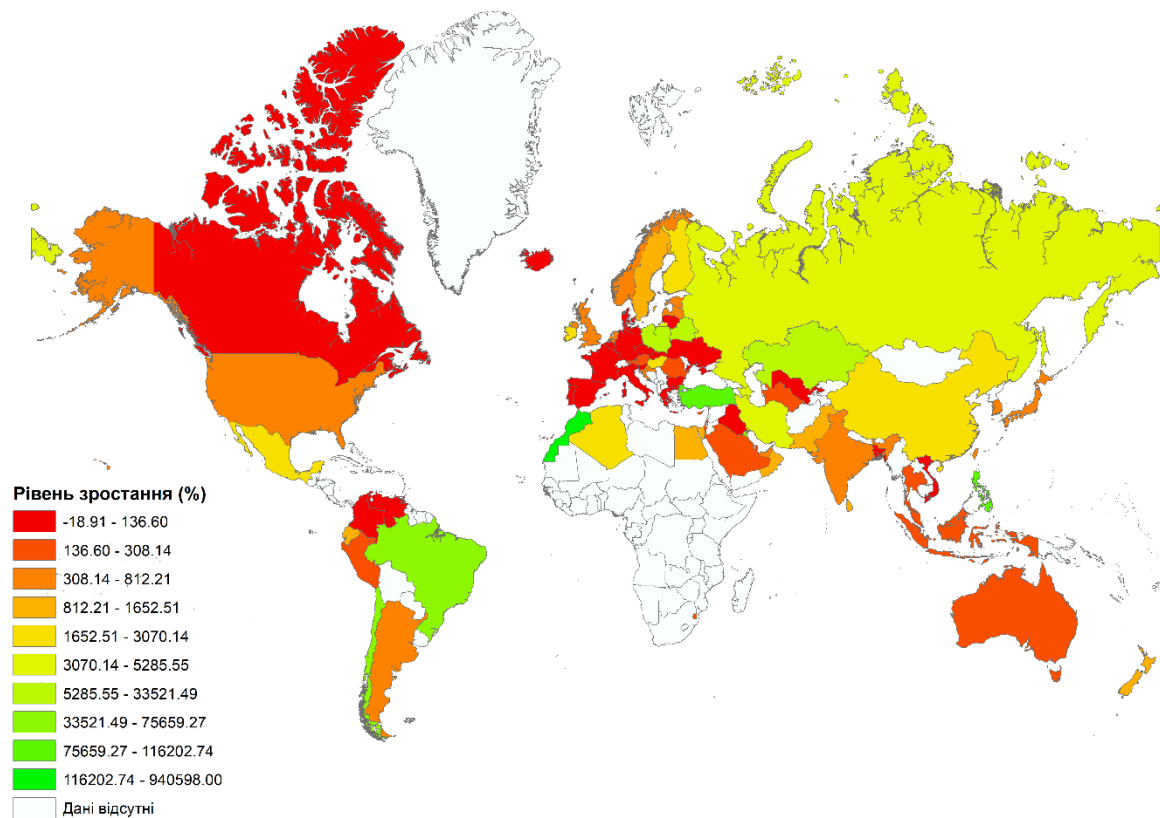


Рис. 3.20. Рівень зростання виробництва сонячної електроенергії у відсотках

Рівень зростання виробництва вітряної електроенергії

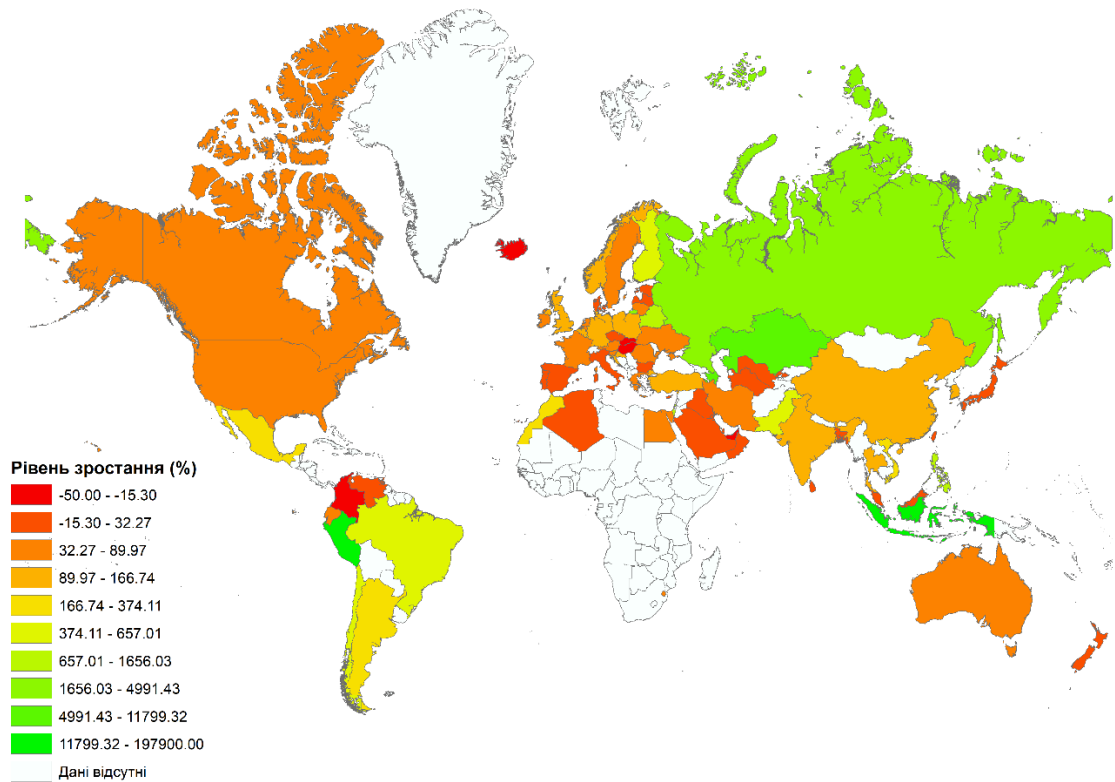


Рис. 3.21. Рівень зростання виробництва вітряної електроенергії у відсотках

Рівень зростання виробництва електроенергії, добутої з інших відновлювальних джерел

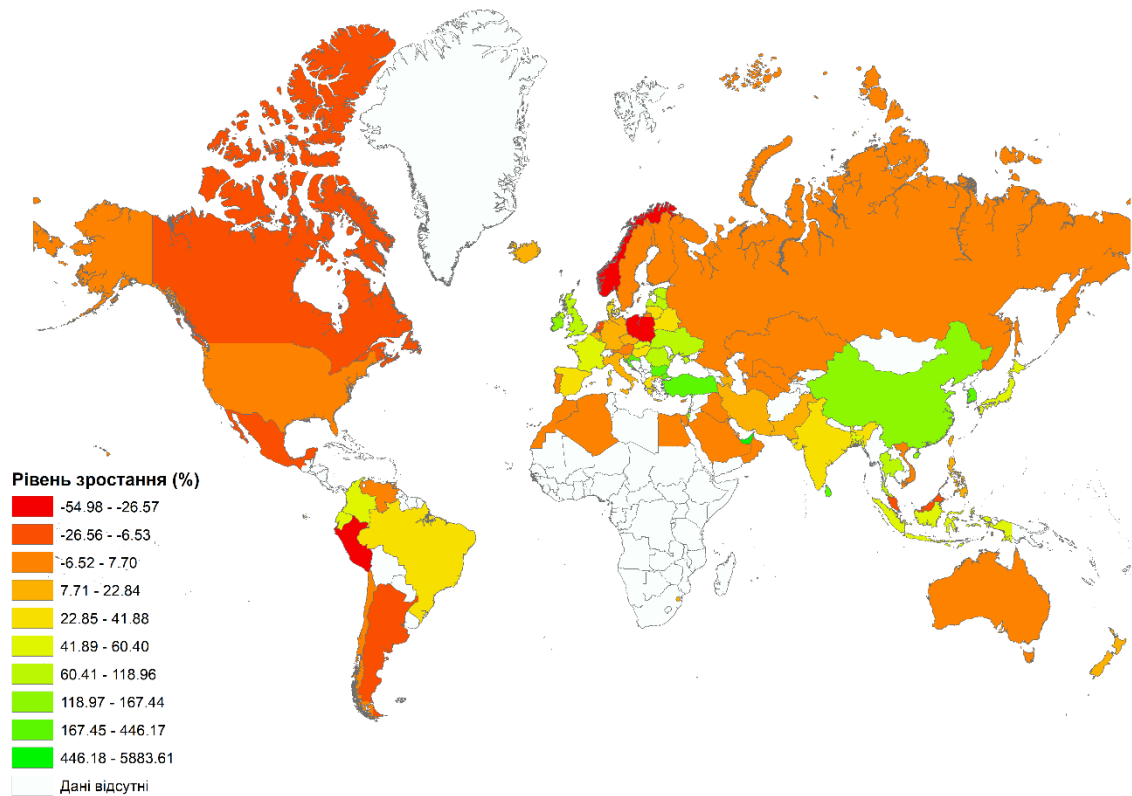


Рис. 3.22. Рівень зростання виробництва електроенергії, добутої з інших відновлювальних джерел, у відсотках

Висновки

Проведено аналіз сукупності показників енергетичної безпеки, який вказує на утворення світових тенденцій з відмови від традиційних джерел енергії на користь відновлювальних.

Розроблено нову методику кластеризації об'єктів з урахуванням їх географічної позиції.

Проведено моделювання взаємозв'язків між показниками, за результатами помітна поступова інтеграція відновлювальних джерел енергії у світову енергетику.

За результатами кластеризації з використанням нової методики можна зазначити, що модель з додаванням стандартизованої географічної відстані між об'єктами найбільш вдало описує ситуацію у світі на сьогоднішній день і дозволяє виділити декілька угруповань країн.

ВИСНОВКИ

В ході дипломної роботи створено новий підхід до кластеризації об'єктів із географічним положенням.

Для початкових даних за 2018 рік проведено моделювання взаємозв'язків між показниками, що відповідають складовим енергетичної безпеки. В ході моделювання проведено розвідувальний аналіз даних, що включає в себе розрахунки коефіцієнтів парної кореляції, багатовимірне шкалювання та аналіз головних компонент. Побудована лінійна регресійна модель, що описує залежність кожного з показників.

На основі результатів моделювання взаємозв'язків між показниками, побудовано когнітивну мапу взаємозв'язків між показниками.

Для початкових даних за 2013 – 2018 роки проведено моделювання угруповань серед показників. Розраховано значення зростання кожного з показників за період у п'ять років. З використанням нового підходу до кластеризації з урахуванням географічного положення, побудовано кластерні моделі для різних варіацій коригування відстані між об'єктами.

За результатами кластеризації можна зазначити, що модель з додаванням стандартизованої географічної відстані між об'єктами найбільш вдало описує ситуацію у світі на сьогоднішній день.

Для цієї моделі можна виділити наступні п'ять кластерів: країни з найбільшим зростанням виробництва сонячної та вітряної електроенергії; країни в яких спостерігається стабільне (помірне) зростання виробництва відновлювальної електроенергії; країни з найбільшим рівнем зростання виробництва відновлювальної електроенергії; країни з розвиненим виробництвом вітряної електроенергії, які успішно розвивають сонячну енергетику та експериментують з видобуванням електроенергії з інших джерел відновлювальної енергії; країни, які знаходяться поза світовими тенденціями – з аномальним зростанням у секторі відновлювальної енергетики.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Renewables 2019 Global Status Report [Електрон. ресурс]. – Спосіб доступу: URL: https://www.ren21.net/wp-content/uploads/2019/05/gsr_2019_full_report_en.pdf
2. Бобылев С. Н., Гирусов Э. В., Перелет Р. А. Экономика устойчивого развития. Учебное пособие. Изд-во Ступени, Москва, 2004, 303 сс.
3. Белоусов К. Ю. Современный этап эволюции концепции устойчивого развития и формирование парадигмы корпоративной устойчивости // Проблемы современной экономики – СПб., 2013, N1 (45). – стр. 47–50
4. Основы устойчивого развития: Учеб. пособие / Под общ. ред. д.е.н., проф. Л. Г. Мельника. – Сумы: ИТД «Университетская книга», 2005. – 654 с.
5. 32-я Генеральная конференция ЮНЕСКО, Доклад Комиссии III с. 17 [Електрон. ресурс]. – Спосіб доступу: URL: <http://unesdoc.unesco.org/images/0013/001320/132068r.pdf>
6. Michael J. Crawley (2007). The R Book. John Wiley & Sons. ISBN 978-0-470-51024-7.
8. Метод наименьших квадратов [Електрон. ресурс]. – Спосіб доступу: URL: https://uk.wikipedia.org/wiki/Метод_наименьших_квадратов
9. Estivill-Castro, Vladimir (20 June 2002). Why so many clustering algorithms – A Position Paper. ACM SIGKDD Explorations Newsletter 4 (1): 65–75. [Електрон. ресурс]. – Спосіб доступу: URL: <http://dx.doi.org/10.1145%2F568574.568575>
10. Practical Guide To Cluster Analysis in R. Alboukadel Kassambara, 2017. [Електрон. ресурс]. – Спосіб доступу: URL: https://www.datanovia.com/en/wp-content/uploads/dn-tutorials/book-preview/clustering_en_preview.pdf

11. Peter J. Rousseeuw (1987). "Silhouettes: a Graphical Aid to the Interpretation and Validation of Cluster Analysis". *Computational and Applied Mathematics*. 20: 53–65. [Електрон. ресурс]. – Спосіб доступу: URL: <https://doi.org/10.1016%2F0377-0427%2887%2990125-7>
12. Сторінка СЕС Уарзазат на Вікіпедії. [Електрон. ресурс]. – Спосіб доступу: URL: https://en.wikipedia.org/wiki/Ouarzazate_Solar_Power_Station
13. Official website of the International Energy Agency . [Електрон. ресурс]. – Спосіб доступу: URL: <http://www.iea.org/topics/energysecurity/>
14. Contribution of renewables to Energy Security. Archived from the original (PDF) on 2009-03-18. Retrieved 2010-06-01. [Електрон. ресурс]. – Спосіб доступу: URL: https://web.archive.org/web/20090318231652/http://www.iea.org/textbase/papers/2007/so_contribution.pdf
15. Rokach, Lior, and Oded Maimon. «Clustering methods.» *Data mining and knowledge discovery handbook*. Springer US, 2005. 321—352.
16. «Жамбю М.» Ієрархічний кластер-аналіз та відповідності. — М.: Финансы и статистика, 1988. — 345 с.
17. "The DISTANCE Procedure: Proximity Measures". *SAS/STAT 9.2 Users Guide*. SAS Institute. Retrieved 2009-04-26. [Електрон. ресурс]. – Спосіб доступу: URL: https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_distance_sect016.htm
18. Dunn, J. C. (1973-09-17). "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters". *Journal of Cybernetics*. 3 (3): 32–57. [Електрон. ресурс]. – Спосіб доступу: URL: <https://doi.org/10.1080%2F01969727308546046>
19. Pearson, K. (1901). "On Lines and Planes of Closest Fit to Systems of Points in Space". *Philosophical Magazine*. 2 (11): 559–572 [Електрон. ресурс]. – Спосіб доступу: URL: <https://doi.org/10.1080%2F14786440109462720>

ДОДАТОК А

Відомості матеріалів дипломного проекту

		Позначення	Найменування	Кіль- кість аркушів	Примітки			
1								
2			Документація					
3								
4	A4	ІСТ.КР 19.01.ДА.ПЗ	Пояснювальна записка	86				
5								
6			Електронний носії інформації					
7								
8			Диск DVD-R з презентацією	2				
Зм	Ли ст	№ докум	Підпис	Дата	ІСТ.КР 19.01.ДА.ПЗ			
Розроб.		Радіонов Є.Д.			Матеріали дипломного проекти			
						Літ.	Лист	Арку- шів
						Н	1	1
Керівник		Сергєєва К.Л.				НТУ «ДП», 12; 122М-18-1		
Н.контр.		Коротенко Г.М.						
Зав.каф.		Бусигін Б.С.						

ДОДАТОК Б

Код програми

```
# Create required folders to store output data
dir.create("Tables", showWarnings = FALSE)
dir.create("Images", showWarnings = FALSE)

# Read csv file with header, ';' – cell separator, '.' – decimal separator
table <- na.omit(read.csv2("dataset_2018_selected_csv.csv", header = T, sep = ";", dec = "."))
table.data <- table[,3:ncol(table)]

#

# Normalize data and add it to the new data frame
#

maxs <- apply(table.data, 2, max)
mins <- apply(table.data, 2, min)
norm.table <- data.frame(scale(table.data, center=mins, scale=maxs-mins))
write.csv2(norm.table, file = "Tables/Нормализованная ТАБЛИЦА.csv")

#

# Correlation
#

pair_correlation_coefficients <- data.frame(cor(norm.table))
write.csv2(pair_correlation_coefficients, file = "Tables/Коэффициенты парной корреляции ТАБЛИЦА.csv")

#

# 2D multidimensional scaling
#

dist.norm.table <- dist(norm.table)
mds <- cmdscale(dist.norm.table)

require("png")

png("Images/Multidimension_scale_countries_names.png", width=3840, height=2160, units='px')
plot(mds, xlab='Dimension 1', ylab='Dimension 2')
```

```

text(mds, labels=table$COUNTRY, pos=3, cex=1)
dev.off()

png('Images/Multidimension_scale_countries_indexes.png', width=3840, height=2160, units='px')
plot(mds, xlab='Dimension 1', ylab='Dimension 2')
text(mds, labels=seq_along(table$COUNTRY), pos=3, cex=1)
dev.off()

png('Images/Multidimension_scale_countries_indexes_720.png', width=1024, height=720, units='px')
plot(mds, xlab='Dimension 1', ylab='Dimension 2')
text(mds, labels=seq_along(table$COUNTRY), pos=3, cex=1)
dev.off()

plot(mds, xlab='Dimension 1', ylab='Dimension 2')
text(mds, labels=seq_along(table$COUNTRY), pos=3, cex=1)

#
# Principal Component Methods
#

library(factoextra)

norm.table.pca <- prcomp(table.data, scale = T)
summary(norm.table.pca)

png("Images/Personal_vals_and_theirs_contrib_to_disp.png", width=1024, height=768, units='px')
fviz_eig(norm.table.pca)
dev.off()

eigenvalues <- get_eigenvalue(norm.table.pca)
write.csv2(eigenvalues, file = "Tables/Собственные значения их вклад в дисперсию ТАБЛИЦА.csv")

png('Images/Indicators_projection_on_factor_pan.png', width=1024, height=768, units='px')
fviz_pca_var(norm.table.pca, repel = TRUE)
dev.off()

pca_vars <- get_pca_var(norm.table.pca)

```

```

write.csv2(pca_vars$contrib/100, file = "Tables/Вклад признаков в главные компоненты ТАБЛИЦА.csv")

png('Images/Indicators_representation_in_first_two_main_components_dims.png', width=1024, height=768, units='px')
fviz_pca_ind(norm.table.pca)
dev.off()

# Prepare tables to write data
variables <- 13

# Table with regression coefficients
regression.table.names <- colnames(table.data)
regression.table <- data.frame(A=rep(NA, variables),
                              OIL_CONS=rep(NA, variables),
                              GAS_CONS=rep(NA, variables),
                              COAL_CONS=rep(NA, variables),
                              NUC_GEN=rep(NA, variables),
                              NUC_CONS=rep(NA, variables),
                              HUDRO_GEN=rep(NA, variables),
                              HYDRO_CONS=rep(NA, variables),
                              SOLAR_GEN=rep(NA, variables),
                              SOLAR_CONS=rep(NA, variables),
                              WIND_GEN=rep(NA, variables),
                              WIND_CONS=rep(NA, variables),
                              OTHER_GEN=rep(NA, variables),
                              OTHER_CONS=rep(NA, variables),
                              row.names = regression.table.names)

# Table with characteristics for models
models.stats <- data.frame(R=rep("", variables),
                          RSQUARE=rep("", variables),
                          F=rep("", variables),
                          p=rep("", variables),
                          stringsAsFactors = F)

```

```

regression.model.x1 <- lm(OIL_CONS ~ GAS_CONS + COAL_CONS + NUC_GEN + NUC_CONS + HYDRO_GEN +
HYDRO_CONS + SOLAR_GEN + SOLAR_CONS + WIND_GEN + WIND_CONS + OTHER_GEN + OTHER_CONS,
table.data)

regression.model.x2 <- lm(GAS_CONS ~ OIL_CONS + COAL_CONS + NUC_GEN + NUC_CONS + HYDRO_GEN +
HYDRO_CONS + SOLAR_GEN + SOLAR_CONS + WIND_GEN + WIND_CONS + OTHER_GEN + OTHER_CONS,
table.data)

regression.model.x3 <- lm(COAL_CONS ~ OIL_CONS + GAS_CONS + NUC_GEN + NUC_CONS + HYDRO_GEN +
HYDRO_CONS + SOLAR_GEN + SOLAR_CONS + WIND_GEN + WIND_CONS + OTHER_GEN + OTHER_CONS,
table.data)

regression.model.x4 <- lm(NUC_GEN ~ OIL_CONS + GAS_CONS + COAL_CONS + NUC_CONS + HYDRO_GEN +
HYDRO_CONS + SOLAR_GEN + SOLAR_CONS + WIND_GEN + WIND_CONS + OTHER_GEN + OTHER_CONS,
table.data)

regression.model.x5 <- lm(NUC_CONS ~ OIL_CONS + GAS_CONS + COAL_CONS + NUC_GEN + HYDRO_GEN +
HYDRO_CONS + SOLAR_GEN + SOLAR_CONS + WIND_GEN + WIND_CONS + OTHER_GEN + OTHER_CONS,
table.data)

regression.model.x6 <- lm(HYDRO_GEN ~ OIL_CONS + GAS_CONS + COAL_CONS + NUC_GEN + NUC_CONS +
HYDRO_CONS + SOLAR_GEN + SOLAR_CONS + WIND_GEN + WIND_CONS + OTHER_GEN + OTHER_CONS,
table.data)

regression.model.x7 <- lm(HYDRO_CONS ~ OIL_CONS + GAS_CONS + COAL_CONS + NUC_GEN + NUC_CONS +
HYDRO_GEN + SOLAR_GEN + SOLAR_CONS + WIND_GEN + WIND_CONS + OTHER_GEN + OTHER_CONS,
table.data)

regression.model.x8 <- lm(SOLAR_GEN ~ OIL_CONS + GAS_CONS + COAL_CONS + NUC_GEN + NUC_CONS +
HYDRO_GEN + HYDRO_CONS + SOLAR_CONS + WIND_GEN + WIND_CONS + OTHER_GEN + OTHER_CONS,
table.data)

regression.model.x9 <- lm(SOLAR_CONS ~ OIL_CONS + GAS_CONS + COAL_CONS + NUC_GEN + NUC_CONS +
HYDRO_GEN + HYDRO_CONS + SOLAR_GEN + WIND_GEN + WIND_CONS + OTHER_GEN + OTHER_CONS,
table.data)

regression.model.x10 <- lm(WIND_GEN ~ OIL_CONS + GAS_CONS + COAL_CONS + NUC_GEN + NUC_CONS +
HYDRO_GEN + HYDRO_CONS + SOLAR_GEN + SOLAR_CONS + WIND_CONS + OTHER_GEN + OTHER_CONS,
table.data)

regression.model.x11 <- lm(WIND_CONS ~ OIL_CONS + GAS_CONS + COAL_CONS + NUC_GEN + NUC_CONS +
HYDRO_GEN + HYDRO_CONS + SOLAR_GEN + SOLAR_CONS + WIND_GEN + OTHER_GEN + OTHER_CONS,
table.data)

regression.model.x12 <- lm(OTHER_GEN ~ OIL_CONS + GAS_CONS + COAL_CONS + NUC_GEN + NUC_CONS +
HYDRO_GEN + HYDRO_CONS + SOLAR_GEN + SOLAR_CONS + WIND_GEN + WIND_CONS + OTHER_CONS,
table.data)

regression.model.x13 <- lm(OTHER_CONS ~ OIL_CONS + GAS_CONS + COAL_CONS + NUC_GEN + NUC_CONS +
HYDRO_GEN + HYDRO_CONS + SOLAR_GEN + SOLAR_CONS + WIND_GEN + WIND_CONS + OTHER_GEN,
table.data)

regression.models <- list(regression.model.x1,
                           regression.model.x2,
                           regression.model.x3,
                           regression.model.x4,
                           regression.model.x5,
                           regression.model.x6,
                           regression.model.x7,
                           regression.model.x8,

```

```

        regression.model.x9,
        regression.model.x10,
        regression.model.x11,
        regression.model.x12,
        regression.model.x13)

i<-1
for (reg.model in regression.models) {
  summary <- summary(reg.model)
  r.square <- summary$r.squared
  r <- sqrt(r.square)
  fstatistic <- summary$fstatistic
  f <- unname(fstatistic['value'])
  p <- round(unname(pf(fstatistic[1], fstatistic[2], fstatistic[3], lower.tail = F)), 10)
  models.stats[i,] <- c(r, r.square, f, p)

  # Collect coefficients data for model
  coefs <- unname(reg.model$coefficients)
  if (i < variables) {
    row <- c(coefs[1:i], NA, coefs[(i+1):length(coefs)])
  } else {
    row <- c(coefs[1:i], NA)
  }
  regression.table[i, ] <- row

  print(paste(c("Regression Model",i,"Summary"), collapse = " "))
  print("=====")
  print(summary)
  print("=====")

  i<-i+1
}

write.csv2(regression.table, file = "Tables/Коэффициенты линейной регрессии ТАБЛИЦА.csv", na="")
write.csv2(models.stats, file = "Tables/Характеристики моделей ТАБЛИЦА.csv", na="")

#

```

```

# Build regression models

#

# OIL_CONS = f1(GAS_CONS + NUC_GEN + OTHER_GEN)
# GAS_CONS = f2(OIL_CONS + COAL_CONS + OTHER_GEN)
# COAL_CONS = f3(GAS_CONS + NUC_GEN + HYDRO_GEN + WIND_GEN + OTHER_GEN)
# NUC_GEN = f4(NUC_CONS)
# NUC_CONS = f5(NUC_GEN)
# HYDRO_GEN = f6(HYDRO_CONS + WIND_GEN)
# HYDRO_CONS = f7(HYDRO_GEN + WIND_GEN)
# SOLAR_GEN = f8(SOLAR_CONS + OTHER_GEN)
# SOLAR_CONS = f9(SOLAR_GEN + OTHER_GEN)
# WIND_GEN = f10(WIND_CONS + COAL_CONS + NUC_GEN)
# WIND_CONS = f11(WIND_GEN + COAL_CONS + NUC_GEN)
# OTHER_GEN = f12(OTHER_CONS)
# OTHER_CONS = f13(OTHER_GEN)

oil.cons.lm <- lm(OIL_CONS ~ GAS_CONS + NUC_GEN + OTHER_GEN, table.data)
gas.cons.lm <- lm(GAS_CONS ~ OIL_CONS + COAL_CONS + OTHER_GEN, table.data)
coal.cons.lm <- lm(COAL_CONS ~ GAS_CONS + NUC_GEN + HYDRO_GEN + WIND_GEN + OTHER_GEN, table.data)
nuc.gen.lm <- lm(NUC_GEN ~ NUC_CONS, table.data)
nuc.cons.lm <- lm(NUC_CONS ~ NUC_GEN, table.data)
hydro.gen.lm <- lm(HYDRO_GEN ~ HYDRO_CONS + WIND_GEN, table.data)
hydro.cons.lm <- lm(HYDRO_CONS ~ HYDRO_GEN + WIND_GEN, table.data)
solar.gen.lm <- lm(SOLAR_GEN ~ SOLAR_CONS + OTHER_GEN, table.data)
solar.cons.lm <- lm(SOLAR_CONS ~ SOLAR_GEN + OTHER_GEN, table.data)
wind.gen.lm <- lm(WIND_GEN ~ WIND_CONS + COAL_CONS + NUC_GEN, table.data)
wind.cons.lm <- lm(WIND_CONS ~ WIND_GEN + COAL_CONS + NUC_GEN, table.data)
other.gen.lm <- lm(OTHER_GEN ~ OTHER_CONS, table.data)
other.cons.lm <- lm(OTHER_CONS ~ OTHER_GEN, table.data)

plot(oil.cons.lm)
plot(gas.cons.lm)
plot(coal.cons.lm)
plot(nuc.gen.lm)
plot(nuc.cons.lm)
plot(hydro.gen.lm)

```

```
plot(hydro.cons.lm)
plot(solar.gen.lm)
plot(solar.cons.lm)
plot(wind.gen.lm)
plot(wind.cons.lm)
plot(other.gen.lm)
plot(other.cons.lm)

table.result <- table
table.result$OIL_CONS_PREDICTED <- oil.cons.lm$fitted.values
table.result$OIL_CONS_RESIDUALS <- oil.cons.lm$residuals
table.result$GAS_CONS_PREDICTED <- gas.cons.lm$fitted.values
table.result$GAS_CONS_RESIDUALS <- gas.cons.lm$residuals
table.result$COAL_CONS_PREDICTED <- coal.cons.lm$fitted.values
table.result$COAL_CONS_RESIDUALS <- coal.cons.lm$residuals
table.result$NUC_GEN_PREDICTED <- nuc.gen.lm$fitted.values
table.result$NUC_GEN_RESIDUALS <- nuc.gen.lm$residuals
table.result$NUC_CONS_PREDICTED <- nuc.cons.lm$fitted.values
table.result$NUC_CONS_RESIDUALS <- nuc.cons.lm$residuals
table.result$HYDRO_GEN_PREDICTED <- hydro.gen.lm$fitted.values
table.result$HYDRO_GEN_RESIDUALS <- hydro.gen.lm$residuals
table.result$HYDRO_CONS_PREDICTED <- hydro.cons.lm$fitted.values
table.result$HYDRO_CONS_RESIDUALS <- hydro.cons.lm$residuals
table.result$SOLAR_GEN_PREDICTED <- solar.gen.lm$fitted.values
table.result$SOLAR_GEN_RESIDUALS <- solar.gen.lm$residuals
table.result$SOLAR_CONS_PREDICTED <- solar.cons.lm$fitted.values
table.result$SOLAR_CONS_RESIDUALS <- solar.cons.lm$residuals
table.result$WIND_GEN_PREDICTED <- wind.gen.lm$fitted.values
table.result$WIND_GEN_RESIDUALS <- wind.gen.lm$residuals
table.result$WIND_CONS_PREDICTED <- wind.cons.lm$fitted.values
table.result$WIND_CONS_RESIDUALS <- wind.cons.lm$residuals
table.result$OTHER_GEN_PREDICTED <- other.gen.lm$fitted.values
table.result$OTHER_GEN_RESIDUALS <- other.gen.lm$residuals
table.result$OTHER_CONS_PREDICTED <- other.cons.lm$fitted.values
table.result$OTHER_CONS_RESIDUALS <- other.cons.lm$residuals

library(openxlsx)
```

```

write.xlsx(table.result, "Tables/Result_Table.xlsx")

#
# Clusterization
#

library(raster)
library(rgdal)
library(openxlsx)

grow_table <- na.omit(read.csv("dataset_grow__formatted_2013_2018.csv", header = T, sep = ",", dec = "."))
grow.data <- grow_table[, c(11,14,17,19)]

countries <- readOGR(dsn = "C:/Users/yevenh/Dropbox/University/5th_year/Diploma/r_module/countries_shp", layer =
"countries")
countries <- spTransform(countries, "+init=epsg:4326")

countries@data <- merge(countries@data, grow_table, by.x="COUNTRY_IS", by.y="COUNTRY_ISO")
select.countries <- countries$COUNTRY!="
select.countries[is.na(select.countries)] <- FALSE

coords <- coordinates(countries[select.countries,])
coords.dist <- as.dist(pointDistance(coords, lonlat = T))

dist.grow <- dist(grow.data)

norm.dist.grow <- (dist.grow - min(dist.grow))/(max(dist.grow) - min(dist.grow))+1
norm.coords.dist <- (coords.dist - min(coords.dist))/(max(coords.dist) - min(coords.dist))+1

dist.grow <- norm.dist.grow
dist.grow.mult <- norm.dist.grow * norm.coords.dist
dist.grow.div <- norm.dist.grow / norm.coords.dist
dist.grow.add <- norm.dist.grow + norm.coords.dist
dist.grow.log <- norm.dist.grow + log(norm.coords.dist)

grow.clust <- hclust(dist.grow, method = "ward.D")
png('Images/Clust_Basic.png', width=1024, height=768, units='px')
plot(grow.clust)

```



```

rect.hclust(grow.clust, k=5, border = "red")
dev.off()
grow.groups <- cutree(grow.clust, k=5)
grow_table$CLUSTER_DEFAULT <- grow.groups

grow.clust.mult <- hclust(dist.grow.mult, method = "ward.D")
png('Images/Clust_Mult.png', width=1024, height=768, units='px')
plot(grow.clust.mult)
rect.hclust(grow.clust.mult, k=5, border = "red")
dev.off()
grow.groups.mult <- cutree(grow.clust.mult, k=5)
grow_table$CLUSTER_MULT <- grow.groups.mult

grow.clust.div <- hclust(dist.grow.div, method = "ward.D")
png('Images/Clust_Div.png', width=1024, height=768, units='px')
plot(grow.clust.div)
rect.hclust(grow.clust.div, k=5, border = "red")
dev.off()
grow.groups.div <- cutree(grow.clust.div, k=5)
grow_table$CLUSTER_DIV <- grow.groups.div

grow.clust.add <- hclust(dist.grow.add, method = "ward.D")
png('Images/Clust_Add.png', width=1024, height=768, units='px')
plot(grow.clust.add)
rect.hclust(grow.clust.add, k=5, border = "red")
dev.off()
grow.groups.add <- cutree(grow.clust.add, k=5)
grow_table$CLUSTER_ADD <- grow.groups.add

grow.clust.log <- hclust(dist.grow.log, method = "ward.D")
png('Images/Clust_Add_Log.png', width=1024, height=768, units='px')
plot(grow.clust.log)
rect.hclust(grow.clust.log, k=5, border = "red")
dev.off()
grow.groups.log <- cutree(grow.clust, k=5)
grow_table$CLUSTER_LOG <- grow.groups.log

```

```

# write.xlsx(grow_table, "Tables/Clusters.xlsx")
write.xlsx(grow_table, "Tables/Clusters_norm.xlsx", sheetName = "CLUST_NORM")

#
# Clusterization validation
#

library(fpc)

clust.stats.add <- cluster.stats(dist.grow.add, grow.groups.add)
clust.stats.add$dunn # => 0.7850008
clust.stats.add$dunn2 # => 0.9898722
clust.stats.add$pearsongamma # => 0.5492046 (0 – the same cluster, 1 – different clusters)
clust.stats.add$entropy # => 1.326665
clust.stats.add$ch # => 5.404674
clust.stats.add$clus.avg.silwidths # => Silhouette. (closer to 1 is better); 1: 0.08859622, 2: 0.07576626, 3: 0.07445905, 4:
0.01902806, 5: 0.00000000

clust.stats.mult <- cluster.stats(dist.grow.mult, grow.groups.mult)
clust.stats.mult$dunn # => 0.6193354
clust.stats.mult$dunn2 # => 0.9788105
clust.stats.mult$clus.avg.silwidths # => 0.14693870 0.13840542 0.13051182 0.03379362 0.00000000

clust.stats.log <- cluster.stats(dist.grow.log, grow.groups.log)
clust.stats.log$dunn # => 0.6022099
clust.stats.log$dunn2 # => 0.7891844
clust.stats.log$clus.avg.silwidths # => 0.04637030 -0.07051094 0.02868095 -0.11724916 0.00000000

```

ДОДАТОК В

ВІДГУК

на дипломну роботу магістра

"Створення інформаційної технології

аналізу показників енергетичної безпеки регіонів"

студента групи 122м-18-1 Радіонова Євгена Дмитровича

1. Метою дипломної роботи є дослідження часових рядів показників моніторингу енергетичної безпеки країн світу за допомогою методів та інформаційної технології інтелектуального аналізу даних.

2. Завдання та зміст дипломної роботи відповідає основній меті – оцінці знань і ступеня підготовленості студента за спеціальністю 122 "Комп'ютерні науки".

3. Тема роботи представляється актуальною, оскільки створення автоматизованого інструментарію інформаційної технології аналізу показників енергетичної безпеки спрямоване на визначення тенденцій переходу до відновлюваних джерел енергії.

4. Для досягнення мети дипломної роботи Радіоновим Є.Д. запропоновано метод просторово-часової кластеризації регіонів, що враховує часові тенденції до зміни показників енергетичної безпеки та розташування регіонів у географічному просторі.

5. Оформлення пояснювальної записки виконано, в основному, відповідно до діючих стандартів і нормативних вимог.

6. У якості зауваження слід відзначити недостатнє порівняння отриманих результатів дослідження з існуючими аналогами.

Незважаючи на зазначений недолік, дипломна робота заслуговує оцінки "відмінно".

Керівник,

доцент кафедри ІСТ

К.Л. Сергєєва

ДОДАТОК Г
РЕЦЕНЗІЯ

на дипломну роботу магістра

"Створення інформаційної технології

аналізу показників енергетичної безпеки регіонів"

студента групи 122м-18-1 Радіонова Євгена Дмитровича

1. Тема дипломної роботи, присвячена дослідженню показників моніторингу енергетичної безпеки країн світу, є актуальною і спрямована на вивчення проблеми сталого розвитку регіонів, зокрема, тенденцій використання відновлюваних джерел енергії.

2. У рецензованій роботі досліджено методи інтелектуального та геоінформаційного аналізу показників моніторингу: кластерного, регресійного, кореляційного, факторного аналізу та аналізу головних компонент, багатомірного шкалювання та ін. У роботі запропоновано метод кластеризації, який, на відміну від існуючих, дозволяє здійснювати розбиття регіонів на кластери з урахуванням їх географічного розташування.

3. Практична значимість результатів дипломної роботи полягає в автоматизації етапів аналізу показників моніторингу засобами запропонованої дипломником інформаційної технології з використанням мови програмування R та ГІС ESRI ArcGIS. За даними 2018 року – показниками споживання природних копалин, споживання та виробництва електроенергії, зростання виробництва (у порівнянні з базовим роком) побудовано геоіконічні моделі, що відображають тенденції розвитку країн світу. Зроблено змістовні висновки.

Дипломна робота цілком відповідає вимогам, що пред'являються до кваліфікаційних робіт рівня магістра.

Недолік: недостатня повнота опису результатів практичних досліджень.

Незважаючи на зазначений недолік, дипломна робота в цілому може бути відзначена оцінкою "_____".

Рецензент,

доцент кафедри електронних засобів

телекомунікацій ДНУ імені Олеся Гончара

Д.К. Мозговий

Довідка

про результати перевірки тексту кваліфікаційної роботи магістра
на тему: Створення інформаційної технології аналізу показників
енергетичної безпеки регіонів

12 "Інформаційні технології".

(назва за наказом ректора)

студента групи 122М-18-1

(шифр групи)

Радіонова Євгена Дмитровича

(прізвище, ім'я, по батькові)

Зазначена робота перевірялася коп'ютерною програмою перевірки
плагіату «ЕТХТ».

За результатами перевірки порушень не знайдено.

Керівник кваліфікаційної роботи:

(П.І.Б.)

Нормоконтролер

Коротенко Г.М.
(П.І.Б.)

Завідувач кафедри ІСТ

Бусигін Б.С.
(П.І.Б.)

23.12.2019 р.