

**О. С. Сердюк,**  
кандидат економічних наук,  
ORCID 0000-0003-3049-3144,

**І. П. Петрова,**  
кандидат економічних наук,  
ORCID 0000-0002-0515-5349,

*Інститут економіки промисловості НАН України, м. Київ*

## МЕТОДИЧНІ ПІДХОДИ ДО ВИЯВЛЕННЯ ГАЗОПОДІБНИХ ВІДХОДІВ ВИРОБНИЦТВА, ЩО НЕГАТИВНИМ ЧИНОМ ВПЛИВАЮТЬ НА МАКРОЕКОНОМІЧНІ ПРОЦЕСИ

Найпотужнішим фактором, що спричиняє негативні зовнішні ефекти, з точки зору наслідків, слід визнати газоподібні відходи промислового виробництва. Специфічний агрегатний стан (фізичний стан речовин) таких відходів обумовлює їхнє необмежене поширення у просторі. Крім того, агрегатний стан унеможливує виявлення безпосереднього прямого зв'язку між продуцентом та реципієнтами зовнішніх ефектів. За таких умов втрачається причино-наслідковий зв'язок, що є орієнтиром для оцінки ступеня шкоди, завданої продуцентом зовнішніх ефектів, а отже, зникають підстави для юридичного обґрунтування відповідальності суб'єкта. Частково, проблема відповідальності вирішується шляхом контролю рівня викидів газоподібних відходів на виході виробництва. Однак, з огляду на відсутність чіткого уявлення про причинно-наслідкові зв'язки, механізм подолання наслідків зовнішніх ефектів, що базується на принципі компенсації витрат продуцентом має суб'єктивний характер [1].

Із теоретичної точки зору, задля інтерналізації негативних зовнішніх ефектів необхідне підтвердження їхньої наявності. До цього часу, і в Україні, і в Європейському Союзі не існує єдиної методології оцінки впливу негативних зовнішніх ефектів, хоча в окремих країнах робляться спроби їх виявлення, і відповідно, оцінки ступеня шкоди. Так, наприклад, у роботі наведено метод розрахунку зовнішніх ефектів на основі оцінки можливих впливів в енергетичному секторі [2]. Цей метод включає в себе чотири рівні: 1) визначення джерела; 2) розсіювання забруднюючих речовин в атмосфері; 3) оцінку довгострокових впливів; 4) грошову оцінку зовнішніх витрат [2].

В іншій роботі запропоновано методику оцінки впливу зовнішніх ефектів за допомогою математичного апарату EcoSense, який поєднує принципи

висхідного аналізу та оцінки експлуатаційного ресурсу [3]. Ця модель також була адаптована в енергетичному секторі на підставі даних про стан технології та навколишнього середовища. Слід зазначити, що основна увага наведених досліджень приділялася оцінці негативних зовнішніх ефектів в енергетичному секторі. В Україні методологія оцінки впливу негативних зовнішніх ефектів перебуває на стадії розробки. Зокрема, здійснено оцінку економічного збитку життю та здоров'ю населення Донецької області, нанесеного забрудненням атмосферного повітря [4]. Проте питання оцінки впливу газоподібних відходів виробництва, що негативним чином впливають на макроекономічні процеси, в цих дослідженнях не висвітлено.

**Метою роботи** є наукове обґрунтування методичних підходів до виявлення газоподібних відходів виробництва, що негативним чином впливають на макроекономічні процеси.

Із точки зору традиційного економічного аналізу<sup>1</sup> вплив газоподібних відходів виробництва на макроекономічні процеси не є очевидним, оскільки відсутність безпосередніх зав'язків між суб'єктом та об'єктом (продуцент відходів – реципієнт) унеможливує кількісну та якісну оцінку впливу. Відтак, застосування інструментарію традиційного економічного аналізу<sup>2</sup> задля виявлення газоподібних відходів виробництва, що впливають на макроекономічні процеси, є недоречним.

Статистичний економічний аналіз (у практичній площині) також не здатен вирішити питання оцінки впливу газоподібних відходів виробництва на макроекономічні процеси, оскільки кількість соціально-економічних факторів, що впливають на об'єкт оцінки, ступінь їхньої взаємопов'язаності є неосяжною із точки зору статистики<sup>3</sup>. Однак у результаті застосування статистичного аналізу можуть бути

<sup>1</sup> Традиційний (або класичний) економічний аналіз досліджує економічні явища та процеси, між якими існує прямий зв'язок. Такий зв'язок, може бути вимірний у кількісних та якісних характеристиках.

<sup>2</sup> Балансовий метод, метод ланцюгових підстановок, метод абсолютних та відносних різниць, інтегральний метод (в умовах детермінованої моделі, коли заздалегідь

відомі показники факторів та чітко визначений зв'язок між ними).

<sup>3</sup> Збір статистичної інформації з боку офіційних державних установ та приватних господарств є обмеженим з огляду на: нерозуміння взаємопов'язаності соціально-економічних процесів; відсутність волі; неможливість здійснення об'єктивної оцінки.

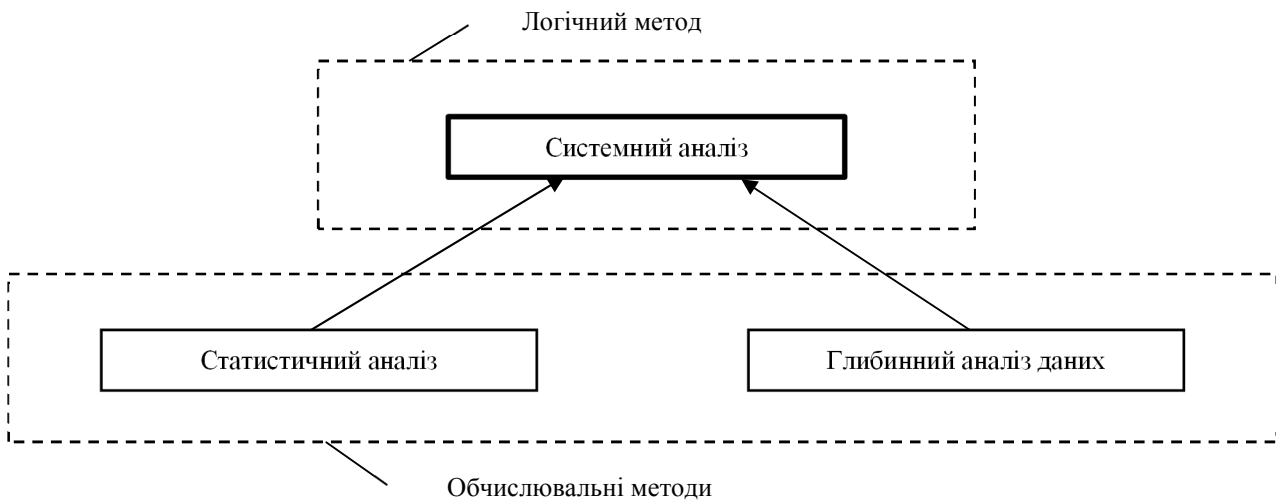
виявлені певні тенденції (не обов'язково статистично значущі), які несуть із собою корисну інформацію для інших методів аналізу.

Наряду із статистичними методами, задля потреб вилучення корисної інформації із масивів статистичних даних можуть бути також використані методи глибинного аналізу даних (англ. Data Mining)<sup>1</sup>. Відмінність цього методу від стандартного статистичного полягає у абстрагуванні від наявних гіпотез<sup>2</sup>. Тобто, методи глибинного аналізу даних відкидають апріорні теоретичні конструкції щодо зв'язків між даними, вибудовуючи при цьому власну модель міжфакторної залежності. Такий підхід робить методи глибинного аналізу даних корисним інструментом для дослідження моделей із неповною інформацією<sup>3</sup> та неочевидністю зв'язків.

Попри відсутність жорсткої прив'язки до структури даних, методи глибинного аналізу даних самі собою не здатні визначити, які саме газоподібні відходи виробництва впливають на макроекономічні

процеси. Причиною цьому є високий рівень абстрагування від взаємозалежних процесів.

Аналізуючи вищенаведені методи економічного аналізу, слід звернути увагу на те, що їхні слабкі місця (у контексті поставленої задачі) є різновекторними за своєю суттю. Для статистичних методів, слабким місцем є детермінація, тоді як для методів глибинного аналізу даних – абстракція. Виходячи із зазначених концептуальних розбіжностей, задля виявлення газоподібних відходів виробництва, що впливають на макроекономічні процеси, доцільно застосувати аналітичний підхід, який узагальнюватиме інформацію отриману за допомогою статистичних методів та методів глибинного аналізу даних. З огляду на суб'єктивність інформації (кожен метод інтерпретує інформацію у властивий для нього спосіб), доцільно застосувати аналітичний підхід, що спирається на логічні висновки. Зробити відповідні логічні висновки можна за допомогою системного аналізу інформації (рис. 1).



**Рис. 1. Методологія виявлення газоподібних відходів виробництва, що негативним чином впливають на макроекономічні процеси**

Джерело: складено авторами.

У рамках даного дослідження, задля опробування методології виявлення газоподібних відходів виробництва, що негативним чином впливають на макроекономічні процеси, як цільовий показник, було обрано врожайність зернових та соняшнику. Вхідними факторами (що впливають на цільовий показник) визначено найпоширеніші елементи газоподібних відходів виробництва, а саме: оксид вуглецю (CO), метан (CH<sub>4</sub>), діоксид азоту (NO<sub>2</sub>), оксиди азоту (NO), сажа, діоксид сірки (SO<sub>2</sub>), неметалеві легкі органічні сполуки, діоксид вуглецю (CO<sub>2</sub>).

Крім того, задля підсилення інформативності вибірки, були додані такі вхідні фактори, як використання мінеральних та органічних добрив. Апріорі відомо, що ці фактори мають позитивний вплив на врожайність, а отже мають більш-менш збалансувати статистичне обчислювання.

Вибірку даних для дослідження складено із фактичних показників по Харківській [6], Дніпропетровській [7], Запорізькій [8], Полтавській [9] та Сумській [10] областях України за період 2008-2019 рр. Задля підвищення інформативності вибірки

<sup>1</sup> Процес напівавтоматичного аналізу великих баз даних з метою пошуку корисної інформації [5].

<sup>2</sup> Слід зауважити, що у певних випадках, у процесі глибинного аналізу даних використовуються статистичні методи. Однак таке використання є допоміжним (прикладним) за своєю суттю.

<sup>3</sup> Неповна інформація може бути наслідком: фактичної нестачі інформації; неусвідомлення зв'язків між факторами, через що певна інформація виключається із аналізу.

розділено на два кластери за видами ґрунтів, які є характерними для певних районів представлених областей. Для першого кластеру – це чорноземи типові, для другого – чорноземи звичайні (табл. 1). Деякі райони було виключено із аналізу, з огляду на недостатню репрезентативність вибірки<sup>1</sup>. Вхідні дані, що відповідають вимогам, було розділено на

чотири групи (таблиці): врожайність зернових – вхідні фактори, 1 кластер; врожайність зернових – вхідні фактори, 2 кластер; врожайність соняшнику – вхідні фактори, 1 кластер; врожайність соняшнику – вхідні фактори, 2 кластер. Об'єм вибірки дорівнює 405 випадкам на кожену групу.

Таблиця 1

Розподілення районів на кластери за видами ґрунтів\*

Кластер I <sup>2</sup>		Кластер II <sup>3</sup>	
<b>Райони Харківської області</b>			
Богодухівський	Коломацький	Балаклійський	Кегичівський
Валківський	Краснокутський	Барвінківський	Красноградський
Великобурлуцький	Куп'янський	Близнюківський	Лозівський
Вовчанський	Печенізький	Борівський	Нововодолазький
Дергачівський	Харківський	Зачепилівський	Первомайський
Дворічанський	Чугувський	Зміївський	Сахновщинський
Золочівський	Шевченківський	Ізюмський	
<b>Райони Дніпропетровської області</b>			
–		всі райони	
<b>Райони Запорізької області</b>			
–		всі райони, за винятком Бердянського, Веселівського, Мелітопольського, Приазовського, Якимівського	
<b>Райони Полтавської області</b>			
решта районів		Новосанжарський	Кобеляцький
<b>Райони Сумської області</b>			
всі райони		–	

\* Розраховано авторами.

**Застосування методу лінійної регресії задля визначення впливу газоподібних відходів на врожайність зернових та соняшнику**

Маючи в наявності спектр даних щодо обсягів викидів газоподібних відходів виробництва при певному рівні врожайності, слід визначити, чи підтверджують наявні дані теоретичні положення щодо впливу перших на останні. Задля цієї мети можна застосувати метод лінійної регресії (метод найменших квадратів). В рамках даного дослідження реалізацію методу лінійної регресії виконано на мові програмування Python із використанням бібліотек Sklearn та Pandas. Слід зауважити, що у контексті підтвердження теоретичних положень щодо впливу газоподібних відходів виробництва на врожайність зернових та соняшнику, коефіцієнт детермінації ( $R^2$ ) регресії не має значення, оскільки мета аналізу даних полягає не у прогнозуванні змін цільового показника, а лише у визначенні орієнтовних закономірностей у спектрі даних, які характеризують направленість впливу (позитивну чи негативну) вхідних факторів на цільовий показник. Тобто, із точки зору методології досліджень, важливим є знак (мінус або його відсутність), що знаходиться перед ваговими коефіцієнтами регресії.

За результатами реалізації методу лінійної регресії (табл. 2) встановлено, що у контексті наявних даних очевидним (за всіма проаналізованими спектрами даних) є негативний зв'язок між врожайністю та викидами: метану ( $CH_4$ ), оксиду азоту ( $NO$ ), сажі. Менш очевидним (3 негативні результати із 4) є негативний зв'язок із викидами оксиду вуглецю ( $CO$ ) та діоксиду сірки ( $SO_2$ ). Позитивний зв'язок врожайності має із: викидами діоксиду азоту ( $NO_2$ ) та легких неметалевих органічних сполук; використанням мінеральних та органічних добрив. Зв'язок врожайності із викидами діоксиду вуглецю ( $CO_2$ ) є невизначеним, оскільки два проаналізованих спектри даних свідчать про позитивний зв'язок, тоді як інші два про негативний (табл. 3).

Із теоретичної точки зору, вплив метану ( $CH_4$ ) на врожайність, є нейтральним. Цей газ майже у два рази легший за повітря, через що накопичується у верхніх шарах атмосфери<sup>4</sup>, не чинячи безпосереднього впливу на рослинність. Щодо оксиду азоту ( $NO$ ), діоксиду сірки ( $SO_2$ ) та сажі, їхній негативний вплив на рослинність є доведеним. Оксид азоту ( $NO$ ) та діоксид сірки ( $SO_2$ ), впливають на рослини, як через прямий контакт, так і через опади (деталь-

<sup>1</sup> Деякі райони Запорізької області мають південні чорноземи. З огляду на те, що таких районів лише декілька, не можна сформулювати репрезентативну вибірку окремого кластеру на їхній основі.

<sup>2</sup> Чорноземи типові.

<sup>3</sup> Чорноземи звичайні.

<sup>4</sup> Негативний вплив метану, полягає у створенні оболонки (у верхніх шарах атмосфери), що перешкоджає віддачі тепла Землі у космос. Як наслідок, виникає парниковий ефект, який несе із собою загрозу глобального потепління.

Таблиця 2

## Вагові коефіцієнти регресії\*

	CO	CH <sub>4</sub>	NO <sub>2</sub>	NO	Сажа
Врожайність зернових, кластер 1	2.24414238e-03	-1.36406613e-03	6.13529163e-03	-1.52658720e-01	-3.09037156e-03
Врожайність зернових, кластер 2	-8.75313356e-04	-1.42847556e-04	3.07197002e-03	-1.82813634e-03	-5.33330655e-06
Врожайність соняшнику, кластер 1	-1.10825357e-03	-9.83824015e-04	3.34415677e-03	-1.67289101e-03	-1.55342248e-03
Врожайність соняшнику, кластер 2	-2.51152011e-03	-7.23805857e-05	2.23340287e-03	-1.08841567e-03	-1.33308310e-04

Продовження табл. 2

	SO <sub>2</sub>	Неметалеві сполуки	CO <sub>2</sub>	Мінеральні добрива	Органічні добрива
Врожайність зернових, кластер 1	-4.70316538e-04	2.58861538e-03	-9.42927297e-05	1.92288806e-01	6.51019926e-02
Врожайність зернових, кластер 2	-2.90020539e-04	1.01821315e-02	1.57840629e-06	1.39295297e-01	-1.04401556e-01
Врожайність соняшнику, кластер 1	4.23670046e-04	-1.44966920e-03	-3.47460151e-06	3.87247645e-02	2.19868202e-02
Врожайність соняшнику, кластер 2	-1.85002393e-04	2.16061717e-03	8.51828959e-07	8.14354900e-02	5.47901453e-02

\* Розраховано авторами.

Таблиця 3

## Зв'язок вхідних факторів із врожайністю, згідно лінійної регресії\*\*

Позитивний	Негативний	Невизначений
Діоксид азоту (NO <sub>2</sub> )	Метан (CH <sub>4</sub> )	Діоксид вуглецю (CO <sub>2</sub> )
Легких неметалевих органічних сполуки	Оксид азоту (NO)	
Мінеральні добрива	Сажа	
Органічні добрива	Оксид вуглецю (CO)	
	Діоксид сірки (SO <sub>2</sub> )	

\*\* Розраховано авторами.

ніше у роботі [11]). Сажа, осідаючи на ґрунті, змінює його хімічний склад, що як правило негативним чином відображається на родючості. Щодо оксиду вуглецю (CO), його негативний вплив на рослинність можливий лише за умови значної концентрації в повітрі<sup>1</sup>. Однак в умовах відкритого простору такої концентрації бути не може.

Потрапляння мінеральних та органічних добрив до позитивних факторів, свідчить на користь інформативності вибірки. При цьому, потрапляння діоксиду азоту (NO<sub>2</sub>) до позитивних факторів, певним чином дискредитує інформативність вибірки, оскільки негативний вплив останнього на рослинність є доведеним. Щодо легких неметалевих органічних сполук, із теоретичної точки зору їх важко

охарактеризувати, як виключно позитивний або негативний фактор, оскільки до цієї категорії потрапляє багато елементів, які мають різнонаправлений вплив.

Підсумовуючи результати реалізації методу лінійної регресії, слід звернути увагу на те, що п'ять із десяти факторів узгоджуються із теоретичними положеннями щодо їхнього впливу на рослинність, один фактор суперечить цій теорії, три фактори є нейтральними до впливу, один – невизначеним. Виходячи з цього (п'ять проти одного), вибірку можна вважати інформативною для потреб виявлення газоподібних відходів виробництва, що впливають на врожайність.

Отже, маючи підтвердження того, що наявна вибірка даних об'єктивно репрезентує напрямок впливу (позитивний або негативний) п'яти факторів на врожайність, три із яких є газоподібними відходами виробництва, доцільно сконцентрувати подальшу увагу дослідження саме на них. Решту факторів, три із яких є нейтральними до врожайності, один суперечить теоретичним положенням, ще один, який є невизначеним з огляду на неоднозначність його змісту, доцільно залишити у вибірці задля збалансування аналізу даних.

**Застосування статистичного методу дисперсійного аналізу (ANOVA) задля визначення впливу газоподібних відходів на врожайність зернових та соняшнику**

Мета аналізу даних полягає у встановленні значимості факторів для цільового показника. Задля

<sup>1</sup> Негативний вплив оксиду вуглецю (CO) на рослинність, проявляється за рівня концентрації, що перевищує 1%.

встановлення значимості факторів, із точки зору лінійної залежності, в рамках дослідження використано статистичний метод дисперсійного аналізу (ANOVA). Метод реалізовано на мові програмування Python із використанням бібліотек Sklearn та Pandas.

За результатами дисперсійного аналізу складено чотири матриці значимості для кожного спек-

тру даних (рис. 2, 3, 4, 5). Стовпчики матриці відповідають вхідним факторам, які розташовані у послідовності, визначеній у табл. 2. Рядки, характеризують задану для аналізу кількість факторів (зверху, від 1 до 10). Тобто, кількість значимих факторів, що мають бути виявлені для даного випадку. Показник «True», характеризує значимість фактору, «False» – навпаки.

```
percentile=10 [False True False False False False False False False False]
percentile=20 [False True False False False True False False False False]
percentile=30 [ True True False False False True False False False False]
percentile=40 [ True True True False False True False False False False]
percentile=50 [ True True True False False True True False False False]
percentile=60 [ True True True False False True True False True False]
percentile=70 [ True True True False True True True False True False]
percentile=80 [ True True True False True True True False True True]
percentile=90 [ True True True True True True True False True True]
percentile100 [ True True True True True True True True True True]
```

**Рис. 2. Матриця значимості факторів, для набору даних «Врожайність зернових, 1 кластер» (дисперсійний аналіз)**

Джерело: розраховано авторами.

```
percentile=10 [False False False False True False False False False False]
percentile=20 [False False True False True False False False False False]
percentile=30 [False False True False True False False False True False]
percentile=40 [False False True False True True False False True False]
percentile=50 [ True False True False True True False False True False]
percentile=60 [ True True True False True True False False True False]
percentile=70 [ True True True False True True False False True True]
percentile=80 [ True True True False True True True False True True]
percentile=90 [ True True True False True True True True True True]
percentile100 [ True True True True True True True True True True]
```

**Рис. 3. Матриця значимості факторів, для набору даних «Врожайність зернових, 2 кластер» (дисперсійний аналіз)**

Джерело: розраховано авторами.

```
percentile=10 [False False False False False True False False False False]
percentile=20 [False False True False False True False False False False]
percentile=30 [ True False True False False True False False False False]
percentile=40 [ True True True False False True False False False False]
percentile=50 [ True True True False True True False False False False]
percentile=60 [ True True True False True True False False False True]
percentile=70 [ True True True False True True False False True True]
percentile=80 [ True True True False True True True False True True]
percentile=90 [ True True True True True True True False True True]
percentile100 [ True True True True True True True True True True]
```

**Рис. 4. Матриця значимості факторів, для набору даних «Врожайність соняшнику, 1 кластер» (дисперсійний аналіз)**

Джерело: розраховано авторами.

```

percentile=10 [False False False False True False False False False False]
percentile=20 [False False False False True False True False False False]
percentile=30 [False False False False True False True True False False]
percentile=40 [False False False False True True True True False False]
percentile=50 [ True False False False True True True True False False]
percentile=60 [ True False False False True True True True True False]
percentile=70 [ True False True False True True True True True False]
percentile=80 [ True True True False True True True True True False]
percentile=90 [ True True True False True True True True True True]
percentile100 [ True True True True True True True True True True]
    
```

**Рис. 5. Матриця значимості факторів, для набору даних «Врожайність соняшнику, 2 кластер» (дисперсійний аналіз)**

Джерело: розраховано авторами.

За результатами аналізу чотирьох матриць складено узагальнюючу таблицю умовних оцінок, де кожен умовний бал відповідає одному визначенню значимості фактору, тобто, одному показнику «True» у матриці (табл. 4). Сума умовних балів, характеризує значимість фактору для показника врожайності (в цілому).

Таблиця 4

**Умовна оцінка значимості факторів, складена на основі дисперсійного аналізу\***

CO	CH <sub>4</sub>	NO <sub>2</sub>	NO	Сажа	SO <sub>2</sub>	Н.О.С.	CO <sub>2</sub>	М.Д	О.Д
28	25	29	6	30	33	21	12	22	14

\* Розраховано авторами.

Відповідно до результатів умовного оцінювання, виконаного на основі дисперсійного аналізу, найбільшу значимість для показника врожайності, має: діоксид сірки (SO<sub>2</sub>), сажа, діоксид азоту (NO<sub>2</sub>), оксид вуглецю (CO). Посередня значимість, є характерною для метану (CH<sub>4</sub>), мінеральних добрив та неметалевих легких органічних сполук. Незначимими, із точки зору дисперсійного аналізу, є органічні добрива, діоксид вуглецю (CO<sub>2</sub>) та оксиди азоту (NO).

Екстраполюючи результати умовного оцінювання на теоретичні положення (щодо впливу факторів на врожайність) та направленість впливу факторів (визначену за допомогою лінійної регресії), слід зауважити, що із точки зору статистичних методів, наявна вибірка даних об'єктивно репрезентує негативний вплив діоксиду сірки (SO<sub>2</sub>) та сажі на врожайність. Щодо решти факторів, які відповідають теоретичними положенням щодо напрямку впливу, але із точки зору дисперсійного аналізу не є значимими, доцільно застосувати метод глибинного аналізу даних. Такий підхід дасть змогу, по-перше, остаточно підтвердити вплив діоксиду сірки (SO<sub>2</sub>)

та сажі на врожайність, по-друге, перевірити значимість інших факторів з позиції абстрагування від наявних гіпотез.

**Застосування методу рекурсивного виключення ознак задля визначення впливу газоподібних відходів на врожайність зернових та соняшнику**

У рамках дослідження, задля виявлення значимих факторів на основі глибинного аналізу даних, використано метод рекурсивного виключення ознак (recursive feature elimination, RFE). Сутність даного методу полягає у побудові моделі (що включає у себе всі фактори), із якої виключається найменш значимий фактор (ознака) із токи зору моделі. Після цього, строюється нова модель, яка містить у собі всі фактори за винятком виключеного на попередньому етапі, і так далі, допоки не залишиться заздалегідь визначена кількість факторів [12].

Метод рекурсивного виключення ознак виконано на мові програмування Python із використанням бібліотек Sklearn та Pandas.

Як у випадку дисперсійного аналізу, за результатами реалізації методу рекурсивного виключення ознак складено чотири матриці значимості<sup>1</sup> (рис. 6, 7, 8, 9).

Таблиця 5

**Умовна оцінка значимості факторів, складена на основі методу рекурсивного виключення ознак\***

CO	CH <sub>4</sub>	NO <sub>2</sub>	NO	Сажа	SO <sub>2</sub>	Н.О.С.	CO <sub>2</sub>	М.Д	О.Д
40	32	29	7	13	21	23	17	32	6

\* Розраховано авторами.

Аналізуючи таблицю умовних оцінок значимості факторів, складену на основі методу рекурсивного виключення ознак, слід звернути увагу на більшу збалансованість оцінок у порівнянні із дисперсійним аналізом. Це свідчить про відносну ідентичність результатів глибинного аналізу за різними

<sup>1</sup> Відмінність у метаданих матриці дисперсійного аналізу та методу рекурсивного виключення ознак полягає у тому, що у першому випадку кількість факторів для

визначення задавалися в перцентилях (percentile), у другому випадку у кількості (n\_features\_to\_select).

спектрами даних. Щодо оцінок наведених у таблиці, наочною є відмінність від результатів дисперсійного аналізу у значимості таких показників, як: сажа, діоксид сірки (SO<sub>2</sub>), оксид вуглецю (CO), мінеральні

добрива. У відносних величинах<sup>1</sup>, суттєво відрізняється значимість органічних добрив, сажі, діоксиду сірки (SO<sub>2</sub>) та мінеральних добрив (табл. 6).

```

n_features_to_select=1 [ True False False False False False False False False False]
n_features_to_select=2 [ True True False False False False False False False False]
n_features_to_select=3 [ True True False False False False False False True False]
n_features_to_select=4 [ True True True False False False False False True False]
n_features_to_select=5 [ True True True False False False False True True False]
n_features_to_select=6 [ True True True False False True True False True False]
n_features_to_select=7 [ True True True False True True True False True False]
n_features_to_select=8 [ True True True False True True True True True False]
n_features_to_select=9 [ True True True False True True True True True True]
n_features_to_select=10[ True True True True True True True True True True]
    
```

**Рис. 6. Матриця значимості факторів, для набору даних «Врожайність зернових, 1 кластер» (метод рекурсивного виключення ознак)**

Джерело: розраховано авторами.

```

n_features_to_select=1 [ True False False False False False False False False False]
n_features_to_select=2 [ True False False False False False True False False False]
n_features_to_select=3 [ True True True False False False False False False False]
n_features_to_select=4 [ True False True False False False True False True False]
n_features_to_select=5 [ True True True False False False False True True False]
n_features_to_select=6 [ True True True False False False True True True False]
n_features_to_select=7 [ True True True False False True True True True False]
n_features_to_select=8 [ True True True False True True True True True False]
n_features_to_select=9 [ True True True True True True True True True False]
n_features_to_select=10[ True True True True True True True True True True]
    
```

**Рис. 7. Матриця значимості факторів, для набору даних «Врожайність зернових, 2 кластер» (метод рекурсивного виключення ознак)**

Джерело: розраховано авторами.

```

n_features_to_select=1 [ True False False False False False False False False False]
n_features_to_select=2 [ True True False False False False False False False False]
n_features_to_select=3 [ True True False False False False False False True False]
n_features_to_select=4 [ True True True False False False False False True False]
n_features_to_select=5 [ True True True False False False True False True False]
n_features_to_select=6 [ True True True False False True True False True False]
n_features_to_select=7 [ True True True False True True True False True False]
n_features_to_select=8 [ True True True False True True True True True False]
n_features_to_select=9 [ True True True False True True True True True True]
n_features_to_select=10[ True True True True True True True True True True]
    
```

**Рис. 8. Матриця значимості факторів, для набору даних «Врожайність соняшнику, 1 кластер» (метод рекурсивного виключення ознак)**

Джерело: розраховано авторами.

<sup>1</sup> Коефіцієнт (від одиниці), що характеризує відмінність показників.

```
n_features_to_select=1 [ True False False False False False False False False False]
n_features_to_select=2 [ True False False False False False False False True False]
n_features_to_select=3 [ True False False False False True False False True False]
n_features_to_select=4 [ True True True False False False False False True False]
n_features_to_select=5 [ True True True False False True False False True False]
n_features_to_select=6 [ True True True False False True True False True False]
n_features_to_select=7 [ True True True False False True True True True False]
n_features_to_select=8 [ True True True True False True True True True False]
n_features_to_select=9 [ True True True True True True True True True False]
n_features_to_select=10 [ True True True True True True True True True True]
```

**Рис. 9. Матриця значимості факторів, для набору даних «Врожайність соняшнику, 2 кластер» (метод рекурсивного виключення ознак)**

Джерело: розраховано авторами.

Таблиця 6

**Розбіжності умовних оцінок, отриманих на основі дисперсійного аналізу та методу рекурсивного виключення ознак\***

	CO	CH <sub>4</sub>	NO <sub>2</sub>	NO	Сажа	SO <sub>2</sub>	Н.О.С.	CO <sub>2</sub>	М.Д	О.Д
Дисперсійний аналіз	28	25	29	6	30	33	21	12	22	14
Метод рекурсивного виключення ознак	40	32	29	7	13	21	23	17	32	6
Відмінність у кількісних величинах	12	7	0	1	17	12	2	5	10	8
Відмінність у відносних величинах	0.3	0.21	0	0.14	0.56	0.36	0.09	0.29	0.31	0.57

\* Розраховано авторами.

Таблиця 7

**Розбіжності у позиціях за умовним рейтингом значимості, складеним на основі дисперсійного аналізу та методу рекурсивного виключення ознак\***

	CO	CH <sub>4</sub>	NO <sub>2</sub>	NO	Сажа	SO <sub>2</sub>	Н.О.С.	CO <sub>2</sub>	М.Д	О.Д
Дисперсійний аналіз	4	5	3	10	2	1	7	9	6	8
Метод рекурсивного виключення ознак	1	2	3	8	7	5	4	6	2	9
Відмінність у позиціях	3	3	0	2	3	4	3	3	4	1

\* Розраховано авторами.

За умовним рейтингом значимості<sup>1</sup>, складеним на основі дисперсійного аналізу та методу рекурсивного виключення ознак, розбіжності є більш-менш однотипними, що може свідчити про певні концептуальні подібності використаних методів.

Абсолютну подібність за всіма напрямками має показник значимості діоксид азоту (NO<sub>2</sub>). Однак, з огляду те, що наявні дані об'єктивно не репрезентують напрямок впливу цієї речовини на рослинність (суперечність фактичних результатів аналізу теоретичним положенням), брати діоксид азоту (NO<sub>2</sub>) до уваги не доцільно. Близьким до абсолютної ідентичності результатів аналізу є показник оксиду

азоту (NO). Однак, попри свою репрезентативність, оксид азоту (NO) має низьку значимість, що виключає його із числа факторів (принаймні із токи зору наявної інформації), які суттєво впливають на макроекономічні процеси.

Попри помірну розбіжність у результатах аналізу (оціночну), можна зробити висновок, що використання мінеральних добрив має безпосередній позитивний вплив на врожайність. Про це свідчать високі показники значимості фактору, отримані за допомогою методу рекурсивного виключення ознак, та посередні оцінки дисперсійного аналізу. Щодо органічних добрив, їхній вплив на врожайність не є

<sup>1</sup> Фактору, що має найвищий показник значимості, присвоєно 1 позицію у рейтингу, і так ділі (2, 3, 4...).



очевидним, з огляду на низькі показники значимості цього фактору за обома методами оцінки. Розбіжність у один пункт в умовних рейтингах значимості, підтверджує цю тезу.

Особливої уваги заслуговують результати оцінки значимості оксиду вуглецю (CO) та метану (CH<sub>4</sub>). Як зазначалось раніше, із теоретичної точки зору ці фактори є нейтральними до врожайності. Однак лінійна регресія визначає їхній негативний вплив, а дисперсійний аналіз та метод рекурсивного виключення ознак (попри помірну розбіжність оцінок) підтверджують високу значимість цих факторів. Можна припустити, що такі результати є наслідком тісного зв'язку оксиду вуглецю (CO) та метану (CH<sub>4</sub>) із елементами, які безпосередньо негативним чином впливають на родючість ґрунтів. Так, наприклад, супутніми елементами викидів оксиду вуглецю (CO) та метану (CH<sub>4</sub>) в атмосферу є оксид азоту (NO), діоксид сірки (SO<sub>2</sub>) та сажа, чий негативний вплив на врожайність є доведеним. Більше того, можна припустити, що супутніми викидами оксиду вуглецю (CO) та метану (CH<sub>4</sub>) є елементи не представлені у вибірці даних, які суттєво впливають на рослинність.

Із точки зору виявлення газоподібних відходів виробництва, що негативним чином впливають на врожайність, інформативними є результати аналізу за діоксидом сірки (SO<sub>2</sub>) та сажею. Попри суттєву розбіжність оцінок, можна зробити висновок щодо очевидності впливу цих факторів на врожайність. Об'єктивність висновку обумовлюється тим, що за результатами дисперсійного аналізу, ці фактори отримали найвищі показники значимості. Помірковану значимість діоксиду сірки (SO<sub>2</sub>) та сажі, засвідчили результати застосування методу рекурсивного виключення ознак, що у сумі дає підстави для висновку щодо очевидності негативного впливу цих факторів на врожайність.

Наведена у дослідженні методика є інструментарієм для виявлення газоподібних відходів виробництва, що негативним чином впливають на макроекономічні процеси. Ефективність застосування цієї методики, залежить від: об'єктивності даних зібраних у вибірці; врахування чинників, що є спільними для певної групи даних (у нашому випадку, це особливості ґрунту); диференціації цільових показників (в рамках дослідження використано показники врожайності зернових та соняшнику); максимізації факторів, що пов'язані з цільовим показником.

Наведений у рамках дослідження приклад застосування методики не є спробою надати вичерпну оцінку факторам газоподібних відходів виробництва, які із теоретичної точки зору впливають на макроекономічні процеси. Наведений приклад, це

лише локальний аналіз (вплив газоподібних відходів на врожайність двох культур), об'єктивність якого залежить від достовірності використаних даних<sup>1</sup>.

Задля вичерпної оцінки, дана методика має бути застосована по відношенню до всіх секторів, які є потенційними реципієнтами негативних ефектів, що виникають унаслідок викидів газоподібних відходів виробництва. Результати оцінювання мають стати основою для розробки моделі інтерналізації негативних зовнішніх ефектів.

### Література

1. Сердюк О. С. Зовнішні ефекти продукування газоподібних відходів виробництва. *Економічний вісник Донбасу*. 2018. № 3 (53). С. 36-48.
2. Streimikienė D., Ciegis R., Jankauskas V. (2007). Sustainable Energy Development. Vilnius.
3. Чегис Р., Пусинайте Р. Отрицательные внешние эффекты и устойчивое развитие в сфере энергетики. *Балтийский регион*. 2010. № 1. С. 22–40.
4. Гаркушенко О. М. Оцінка економічного збитку життю та здоров'ю населення Донецької області, нанесеного забрудненням атмосферного повітря. *Стратегія і механізми регулювання промислового розвитку*: зб. наук. праць / НАН України, Ін-т економіки пром-сті; редкол.: Булеєв І.П. (відп. ред.) та ін. Донецьк, 2010. С. 252-265.
5. Silberschatz A., Sudarshan S. (2011). Database system concepts. New York: McGraw-Hill.
6. Статистична інформація. Головне управління статистики у Харківській області. URL: <http://kh.ukrstat.gov.ua/stat-informatsiya>.
7. Статистична інформація. Головне управління статистики у Дніпропетровській області. URL: <http://www.dnprstat.gov.ua/statinfo/>.
8. Статистична інформація. Головне управління статистики у Запорізькій області. URL: <http://www.zp.ukrstat.gov.ua/index.php/statystychna-informatsiya>.
9. Статистична інформація. Головне управління статистики у Полтавській області. URL: <http://www.pl.ukrstat.gov.ua>.
10. Статистична інформація. Головне управління статистики у Сумській області. URL: <http://sumy.ukrstat.gov.ua/?menu=610>.
11. Бояршинов А. В., Асафова Е. В. Стрессовые реакции листьев пшеницы на обезвоживание: участие эндогенного NO и эффект нитропрусида натрия. *Физиология растений*. 2011. № 6. С. 891-897.
12. Мюллер А., Гвидо С. Введение в машинное обучение с помощью Python. Руководство для специалистов по работе с данными / пер. с англ. СПб.: ООО «Альфа-книга», 2017. 480 с.

### References

1. Serdiuk O. S. (2018). Zovnishni efekty prodrukuvannia hazopodibnykh vidkhodiv vyrobnytstva [External effects of production of gaseous industrial waste]. *Ekonomichnyi visnyk Donbasu – Economic Herald of the Donbas*, 3 (53), pp. 36-48 [in Ukrainian].

<sup>1</sup> У рамках дослідження, було подано запит на отримання необхідних для дослідження даних до Харківського, Дніпропетровського, Запорізького, Полтавського та Сумського обласних відділень статистики.

2. Streimikiene D., Ciegis R., Jankauskas V. (2007). *Sustainable Energy Development*. Vilnius.
3. Chegis R., Pusinaite R. (2010). Negative externalities and sustainable development in the energy sector. *Baltic region*, 1, pp. 22–40 [in Russian].
4. Harkushenko O. M. (2010). Otsinka ekonomichnoho zbytku zhyttiu ta zdoroviu naseleння Donetskoї oblasti, nanesenoho zabrudnenniam atmosferного povitria [Estimation of economic damage to life and health of the population of Donetsk region caused by air pollution]. *Stratehiia i mekhanizmy rehulivannia promyslovoho rozvytku – Strategy and mechanisms for regulating industrial development*, pp. 252-265. Donetsk, IIE of NAS of Ukraine [in Ukrainian].
5. Silberschatz A., Sudarshan S. (2011). *Database system concepts*. New York: McGraw-Hill.
6. Statystychna informatsiia. Holovne upravlinnia statystyky u Kharkivskii oblasti [Statistical information. Main Department of Statistics in Kharkiv Region]. *kh.ukrstat.gov.ua*. Retrieved from <http://kh.ukrstat.gov.ua/stat-informatsiya> [in Ukrainian].
7. Statystychna informatsiia. Holovne upravlinnia statystyky u Dnipropetrovskii oblasti [Statistical information. Main Department of Statistics in Dnipropetrovsk region]. *dneprstat.gov.ua*. Retrieved from <http://www.dneprstat.gov.ua/statinfo> [in Ukrainian].
8. Statystychna informatsiia. Holovne upravlinnia statystyky u Zaporizkii oblasti [Statistical information. Department of Statistics in Zaporizhia region]. *zp.ukrstat.gov.ua*. Retrieved from <http://www.zp.ukrstat.gov.ua/index.php/statystychna-informatsiia> [in Ukrainian].
9. Statystychna informatsiia. Holovne upravlinnia statystyky u Poltavskii oblasti [Statistical information. Main Department of Statistics in Poltava region]. *pl.ukrstat.gov.ua*. Retrieved from <http://www.pl.ukrstat.gov.ua> [in Ukrainian].
10. Statystychna informatsiia. Holovne upravlinnia statystyky u Sumskii oblasti [Statistical information. Main Department of Statistics in Sumy region]. *sumy.ukrstat.gov.ua*. Retrieved from <http://sumy.ukrstat.gov.ua/?menu=610> [in Ukrainian].
11. Boyarshinov A. V., Asafova E. V. (2011). Stress reactions of wheat leaves to dehydration: participation of endogenous NO and the effect of sodium nitroprusside. *Plant physiology*, 6, pp. 891-897 [in Russian].
12. Mueller A., Guido S. (2017). *An Introduction to Machine Learning with Python. Data Scientist's Guide*. Transl. from English. SPb., LLC "Alfa-book" [in Russian].

**Сердюк О. С., Петрова І. П. Методичні підходи до виявлення газоподібних відходів виробництва, що негативним чином впливають на макроекономічні процеси**

Негативні зовнішні ефекти, що спричинені газоподібними відходами виробництва, це явище досить динамічне, тому потребують додаткового вивчення з метою пошуку інструментів їх інтерналізації. Негативні зовнішні ефекти в даний час недостатньо вивчені, залишаються мало вивченими питання оцінки ступеня впливу газоподібних відходів виробництва на макроекономічні процеси.

Метою роботи є наукове обґрунтування методичних підходів до виявлення газоподібних відходів виробництва, що негативним чином впливають на макроекономічні процеси.

Запропоновано методологію виявлення газоподібних відходів виробництва, що негативним чином впливають на макроекономічні процеси. В основі лежить аналітичний підхід, який узагальнюватиме інформацію, отриману за допомогою статистичних методів та методів глибинного аналізу даних. Цільовим показником для виявлення впливу газоподібних відходів було обрано врожайність зернових та соняшнику. Вхідними факторами визначено найпоширеніші елементи газоподібних відходів виробництва, а саме: оксид вуглецю (CO), метан (CH<sub>4</sub>), діоксид азоту (NO<sub>2</sub>), оксиди азоту (NO), сажа, діоксид сірки (SO<sub>2</sub>), неметалеві легкі органічні сполуки, діоксид вуглецю (CO<sub>2</sub>). Апробовано методологію визначення ступеня впливу відходів виробництва на макроекономічні процеси в Харківській, Дніпропетровській, Запорізькій, Полтавській та Сумській областях України.

За результатами аналітичного узагальнення оцінено ступінь впливу найпоширеніших елементів газоподібних відходів виробництва. Інформативними є результати аналізу за діоксидом сірки (SO<sub>2</sub>) та сажею. Отримані результати оцінювання мають стати основою для розробки моделі інтерналізації негативних зовнішніх ефектів. Перспективним напрямом подальших досліджень є апробація даної методології по відношенню до всіх секторів, які є потенційними реципієнтами негативних ефектів, що виникають внаслідок викидів газоподібних відходів виробництва.

*Ключові слова:* негативні зовнішні ефекти, газоподібні відходи, промисловість, макроекономічні процеси, аналітичний підхід, статистичний метод, метод глибинного аналізу даних.

**Serdiuk O., Petrova I. Methodical Approaches to the Identification of Gaseous Production Waste that in a Negative Effect Affect Macroeconomic Processes**

The negative externalities caused by gaseous production waste are quite dynamic, so they need additional study in order to find tools for their internalization. Negative externalities are currently not sufficiently studied, the question of assessing the degree of impact of gaseous production waste on macroeconomic processes remains little studied.

The aim of the work is to scientifically substantiate methodological approaches to the detection of gaseous production waste that negatively affects macroeconomic processes.

A methodology for identifying production wastes that negatively affect macroeconomic processes is proposed. It is based on an analytical approach that will summarize the information obtained using statistical methods and methods of in-depth data analysis. Grain and sunflower yields were chosen as a target for detecting the impact of gaseous waste. The input factors are the most common elements of gaseous production waste, namely: carbon monoxide (CO), methane (CH<sub>4</sub>), nitrogen dioxide (NO<sub>2</sub>), nitrogen oxides (NO), soot, sulfur dioxide (SO<sub>2</sub>), non-metallic light organic compounds, carbon dioxide (CO<sub>2</sub>). The methodo-

logy for determining the degree of impact of production waste on macroeconomic processes in Kharkiv, Dnipropetrovsk, Zaporizhia, Poltava and Sumy regions of Ukraine has been tested.

According to the results of analytical generalization, the degree of influence of the most common elements of gaseous production waste is estimated. The results of the analysis for sulfur dioxide (SO<sub>2</sub>) and soot are informative. The obtained evaluation results should become the basis for the development of a model of internalization of negative externalities. A promising area of further research is the testing of this methodology in relation to all sectors that are potential recipients of negative effects arising from emissions of gaseous production waste.

*Keywords:* negative externalities, gaseous waste, industry, macroeconomic processes, analytical approach, statistical method, method of in-depth data analysis.

**Сердюк А. С., Петрова И. П. Методические подходы к выявлению газообразных отходов производства, которые негативно влияют на макроэкономические процессы**

Отрицательные внешние эффекты, вызванные газообразными отходами производства, это явление довольно динамичное, поэтому требуют дополнительного изучения с целью поиска инструментов их интернализации. Отрицательные внешние эффекты в настоящее время недостаточно изучены. Остается мало изученным вопрос оценки степени влияния газообразных отходов производства на макроэкономические процессы.

Целью работы является научное обоснование методических подходов к выявлению газообразных отходов производства, которые негативно влияют на макроэкономические процессы.

Предложена методология выявления отходов производства, которые негативно влияют на макроэкономические процессы. В основе лежит аналитический подход, который обобщает информацию, полученную с помощью статистических методов и методов глубинного анализа данных. Целевым показателем для выявления влияния газообразных отходов выбрана урожайность зерновых и подсолнечника. Входными факторами определены наиболее распространенные элементы газообразных отходов производства, а именно: оксид углерода (CO), метан (CH<sub>4</sub>), диоксид азота (NO<sub>2</sub>), оксиды азота (NO), сажа, диоксид серы (SO<sub>2</sub>), неметаллические легкие органические соединения, диоксид углерода (CO<sub>2</sub>). Апробирована методология определения степени влияния отходов производства на макроэкономические процессы в Харьковской, Днепропетровской, Запорожской, Полтавской и Сумской областях Украины.

По результатам аналитического обобщения оценена степень влияния наиболее распространенных элементов газообразных отходов производства. Информативными являются результаты анализа по диоксиду серы (SO<sub>2</sub>) и саже. Полученные результаты оценки могут стать основой для разработки модели интернализации отрицательных внешних эффектов. Перспективным направлением дальнейших исследований является апробация данной методологии в отношении всех секторов, которые являются потенциальными реципиентами негативных эффектов, возникающих в результате выбросов газообразных отходов производства.

*Ключевые слова:* отрицательные внешние эффекты, газообразные отходы, промышленность, макроэкономические процессы, аналитический подход, статистический метод, метод глубинного анализа данных.

Стаття надійшла до редакції 02.09.2020  
Прийнято до друку 15.09.2020