ЕЛЕКТРОНІКА, ЕЛЕКТРОННІ КОМУНІКАЦІЇ, ПРИЛАДОБУДУВАННЯ ТА РАДІОТЕХНІКА

УДК 528.8:528.7:528.4:528.06
DOI https://doi.org/10.32782/2521-6643-2026-2-72.44

Альперт С. І., кандидат технічних наук,
доцент кафедри аерокосмічної геодезії та землеустрою
факультету архітектури, будівництва та дизайну
Державного університету «Київський авіаційний інститут»
ORCID: 0000-0002-7284-6502

НОВІТНІ СТАТИСТИЧНІ МЕТОДИ КЛАСИФІКУВАННЯ СУПУТНИКОВИХ ЗОБРАЖЕНЬ

На даний час дистанційне зондування Землі є одним із найбільш ефективних методів одержання інформації про стан, властивості та спектральні характеристики об'єктів земної поверхні. Класифікування супутникових знімків є основною задачею у дистанційному зондуванні, що передбачає віднесення кожного пікселя зображення до конкретного тематичного класу на основі його спектральних характеристик. Результатом процедури класифікування є тематична карта, що відображає розподіл різних типів поверхні. Статистичні методи класифікування засновані на аналізі спектральних характеристик пікселів зображення та їх віднесенні до класів земного покриву. Метод максимальної правдоподібності є одним із найбільш надійних та точних методів контрольованого класифікування. Основним припущенням цього методу є те, що спектральні характеристики пікселів у кожному класі описуються нормальним (гаусівським) розподілом. Зазначено, що дане припущення є вірним тільки для однорідних природних поверхонь, таких як водойми або сільськогосподарські поля. Це припущення є невірним для змішаних типів покриву та неоднорідних міських територій. Проведено порівняльний аналіз методу максимальної правдоподібності із застосуванням трьох статистичних розподілів для представлення спектральних характеристик класів земного покриву: нормального розподілу, експоненціального розподілу та розподілу Вейбулла. Застосовано критерій максимальної правдоподібності для проведення процедури класифікування, та наведено формули логарифмів функцій правдоподібності для трьох типів розподілу. Зазначено, що нормальний розподіл ефективний для класів із однорідною структурою, таких як водні поверхні, густа рослинність та асфальтовані дороги. Експоненціальний розподіл є оптимальним для класифікування радіолокаційних зображень. Цей розподіл описує інтенсивність радіолокаційних сигналів від шорстких поверхонь. Розподіл Вейбулла ефективний для гетерогенних класів зі змінною текстурою, таких як міські території. Даний розподіл зазвичай використовується для моделювання амплітудних характеристик радіолокаційних сигналів. Вибір статистичного розподілу повинен проводитися для кожного класу земного покриву. Показано, що експоненціальний розподіл є частковим випадком розподілу Вейбулла. Подальші напрямки досліджень включають застосування ансамблевих підходів, які поєднують різні типи розподілів для різних класів та інтеграцію із методами машинного навчання.

Ключові слова: *супутникові зображення, статистичні методи класифікування, метод максимальної правдоподібності, нормальний розподіл, експоненціальний розподіл, розподіл Вейбулла.*

Alpert S. I. New statistical methods for satellite image classification

Nowadays remote sensing of the Earth is one of the most efficient methods for acquiring information about the state, properties and spectral characteristics of the Earth's surface objects. Classification of satellite images is a main task in remote sensing, that contains the assignment of each image pixel to a specific thematic class based on its spectral characteristics. The result of classification procedure is a thematic map, that shows the distribution of various surface types. Statistical classification methods are based on analysis of spectral characteristics of image pixels and their assignment to land cover classes. The maximum likelihood method is one of the most dependable and precise supervised classification methods. The main assumption of this method is that spectral characteristics of pixels within each class follow a normal (Gaussian) distribution. But it was noted, that this assumption is correct only for homogeneous natural surfaces such as water bodies or agricultural fields. This assumption is not valid for mixed cover types and heterogeneous urban territories. It was conducted a comparative analysis of maximum likelihood method applying three statistical distributions for representing spectral characteristics of land cover classes: normal distribution, exponential distribution, and Weibull distribution. The maximum likelihood criterion was applied for classification procedure and formulas of the log-likelihood functions were considered for these three distribution types. It was noted, that normal distribution is effective for classes with homogeneous structure such as, water surfaces, dense vegetation and paved roads. The exponential distribution is optimal for radar imagery classification. It describes intensities of radar signals from rough surfaces. The Weibull distribution

is effective for heterogeneous classes with variable texture such as urban territories. This distribution is commonly used to model the amplitude characteristics of radar signals. The choice of statistical distribution should be justified for each specific land cover class. It was shown that exponential distribution is a special case of Weibull distribution. Future research directions include applying ensemble approaches combining various distribution types for different classes and integration with machine learning methods.
Key words: *satellite images, statistical classification methods, maximum likelihood method, normal distribution, exponential distribution, Weibull distribution.*

Постановка проблеми. Як відомо, дистанційне зондування Землі (ДЗЗ) є одним із самих ефективних методів отримання даних про стан та характеристики об'єктів земної поверхні, що, в свою чергу, має досить широке застосування у таких сферах науки, як: картографія, моніторинг навколишнього середовища, лісове і сільське господарство, геологія, екологія, кліматологія, містобудування, військові задачі. Супутникові зображення надають величезний об'єм інформації про спектральні характеристики об'єктів земної поверхні.

Класифікування супутникових зображень є ключовим завданням ДЗЗ, яке базується на віднесенні кожного пікселя супутникового зображення до певного тематичного класу (категорії) на основі його спектральних характеристик. Слід зазначити, що клас – це певний тип земного покриву. При цьому результатом класифікування є тематична карта, яка відображає просторовий розподіл різноманітних типів поверхні: лісів, сільськогосподарських угідь, урбанізованих територій, водних об'єктів, тощо. Статистичні методи класифікування супутникових даних базуються на аналізі спектральних характеристик пікселів зображення та їх належності до певних класів земного покриву.

Класифікування поділяється на контрольоване та неконтрольоване. Контрольоване класифікування використовує еталонні зразки (навчальні вибірки), для яких відома належність до певних класів. При неконтрольованому класифікуванні виділяються групи пікселів із подібними спектральними характеристиками без попереднього визначення класів [1, 2].

Метод максимальної правдоподібності (Maximum Likelihood Classification) є одним із найбільш достовірних та ефективних методів контрольованого класифікування. Даний метод базується на застосуванні даних про розподіл спектральних значень у межах кожного класу, що дає можливість досягти досить точних результатів класифікування при коректному виборі навчальних даних.

Слід зазначити, що вибір типу статистичного розподілу для моделювання спектральних характеристик класів земної поверхні має критичне значення для достовірності класифікування методом максимальної правдоподібності. Різні типи земного покриву характеризуються різними статистичними властивостями, що, в свою чергу, потребує застосування відповідних моделей розподілу. Тобто амплітудні характеристики сигналів (інтенсивності сигналів) різних класів природних об'єктів супутникового зображення можуть бути описані різними статистичними розподілами. Із-за цього виникає необхідність залучення альтернативних статистичних моделей для опису спектральних характеристик.

Отже, актуальність дослідження методу максимальної правдоподібності зумовлена необхідністю підвищення достовірності тематичного картування земної поверхні на основі супутникових зображень.

Аналіз останніх досліджень і публікацій. У даній було проаналізовані сучасні статистичні підходи для обробки космічних даних, зокрема, процедура класифікування супутникових зображень для вирішення задач дистанційного зондування Землі.

У статтях Штралера А., Мьона І., Вана Дж., Гонга П. були розглянуті різні методики класифікування супутникових зображень, зокрема, метод максимальної правдоподібності. Були проаналізовані основні переваги (висока точність, відсутність некласифікованих пікселів) та недоліки (потреба у великому об'ємі машинної пам'яті та великих часових затратах) даного методу. Зазначалося, що зазвичай при застосуванні методу максимальної правдоподібності припускається, що спектральні характеристики пікселів у межах кожного класу підпорядковуються нормальному розподілу [5, 7, 9].

У роботах Чжана С., Вейбулла В., Гао Г., Фрері А. наголошувалося на тому, що при класифікуванні знімків припущення про нормальність розподілу спектральних характеристик пікселів є вірним тільки для однорідних природних поверхонь, а для гетерогенних територій та змішаних типів покриву слід використовувати інші розподіли [3, 4, 10, 11]. Тому, у даній роботі пропонується застосовувати альтернативні статистичні розподіли, а саме, експоненціальний розподіл та розподіл Вейбулла для опису спектральних характеристик об'єктів земної поверхні при класифікуванні зображень методом максимальної правдоподібності.

Метою статті є проведення глибокого порівняльного аналізу методів класифікування аерокосмічних зображень на основі критерію максимальної правдоподібності з використанням нормального розподілу, експоненціального розподілу та розподілу Вейбула для спектральних характеристик класів.

Виклад основного матеріалу. Метод максимальної правдоподібності базується на теорії прийняття статистичних рішень. Основний принцип цього методу полягає у віднесенні кожного пікселя супутникового зображення до того класу, для якого ймовірність приналежності даного пікселя є максимальною. Слід зазначити, що у випадку, коли апріорні ймовірності класів вважаються невідомими чи рівними, критерій зводиться до максимізації функції правдоподібності. Зазвичай для спрощення розрахунків використовується критерій максимума логарифму функції правдоподібності.

При цьому основним припущенням методу максимальної правдоподібності є те, що спектральні характеристики пікселів у межах кожного класу (категорії) підпорядковуються нормальному (Гаусівському) розподілу [6, 7].

Нормальний розподіл широко застосовується для обробки оптичних багатоспектральних зображень, визначення вегетаційних індексів та коефіцієнтів відбиття поверхні. Даний розподіл застосовується для класів із однорідною структурою: густої рослинності, водних поверхонь, асфальтованих доріг, тобто для класів об'єктів, які описуються симетричною гістограмою без значних викидів.

Нормальний (Гаусівський) розподіл

При застосуванні методу максимальної правдоподібності спектральні характеристики пікселів (інтенсивності сигналів пікселів) конкретного класу можуть бути представлені у вигляді вибірки $X_1, X_2, \ldots, X_n$ незалежних однаково розподілених випадкових величин, що підпорядковуються нормальному розподілу з математичним очікуванням μ та дисперсією δ^2

При цьому густина розподілу одного спостереження X_i обчислюється за наступною формулою:

$$f\left(x_i \middle| \mu, \delta^2\right) = \frac{1}{\sqrt{2\pi\delta^2}} \exp\left(-\frac{\left(x_i - \mu\right)^2}{2\delta^2}\right). \tag{1}$$

Оскільки усі спостереження X_i, $i = 1, \ldots, n$ (в даному випадку значення інтенсивностей сигналів пікселів, що відносяться до певного класу) є незалежними, то функція правдоподібності для всієї вибірки дорівнює добутку густин, а саме:

$$L\left(\mu, \delta^2 \middle| x_1, \ldots, x_n\right) = \prod_{i=1}^{n} f\left(x_i \middle| \mu, \delta^2\right). \tag{2}$$

Підставляючи вираз (1) у вираз (2) маємо:

$$L\left(\mu, \delta^2\right) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\delta^2}} \exp\left(-\frac{\left(x_i - \mu\right)^2}{2\delta^2}\right). \tag{3}$$

Далі, використовуючи основні властивості експоненти, вираз (3) записуємо у наступному вигляді (4):

$$L\left(\mu, \delta^2\right) = \left(\frac{1}{\sqrt{2\pi\delta^2}}\right)^n \exp\left(-\frac{1}{2\delta^2} \sum_{i=1}^{n} \left(x_i - \mu\right)^2\right). \tag{4}$$

Перепишемо вираз (4) у вигляді:

$$L\left(\mu, \delta^2\right) = \left(2\pi\delta^2\right)^{-n/2} \exp\left(-\frac{1}{2\delta^2} \sum_{i=1}^{n} \left(x_i - \mu\right)^2\right). \tag{5}$$

Для спрощення розрахунків слід використовувати не саму функцію правдоподібності, а її натуральний логарифм.

Обчислюємо натуральний логарифм функції правдоподібності, враховуючи, що логарифм добутку дорівнює сумі логарифмів:

$$\ln L\left(\mu, \delta^2\right) = \ln\left[\left(2\pi\delta^2\right)^{-n/2}\right] + \ln\left[\exp\left(-\frac{1}{2\delta^2} \sum_{i=1}^{n} \left(x_i - \mu\right)^2\right)\right]. \tag{6}$$

На наступному кроці застосовуємо наступні властивості натуральних логарифмів, а саме: \ln(e^x) = x; ln(a^b) = b ln(a).

Маємо:

$$\ln L\left(\mu, \delta^2\right) = -\frac{n}{2} \ln\left(2\pi\delta^2\right) - \frac{1}{2\delta^2} \sum_{i=1}^{n} \left(x_i - \mu\right)^2. \tag{7}$$

Знов, використовуючи основні властивості натуральних логарифмів, маємо:

$$\ln L\left(\mu, \delta^2\right) = -\frac{n}{2} \ln\left(2\pi\right) - \frac{n}{2} \ln\left(\delta^2\right) - \frac{1}{2\delta^2} \sum_{i=1}^{n} \left(x_i - \mu\right)^2. \tag{8}$$

Аналогічним чином розраховуємо натуральні логарифми функції максимальної правдоподібності для інших класів.

Далі застосовуємо критерій максимума функції правдоподібності, згідно якого піксель супутникового зображення має бути віднесений до того класу, якому відповідає максимальне значення логарифму функції правдоподібності, а саме максимальне значення виразу (8). Тобто ми розв'язуємо задачу максимізації функціоналу [7, 9].

Основні переваги нормального розподілу:

1) густина розподілу описується лише двома параметрами;

2) стійкість оцінок параметрів за умови достатнього об'єму вибірки;

3) ефективність для опису даних оптичних сенсорів;

4) простота обчислень.

Основні недоліки нормального розподілу:

1) симетричність розподілу може не відповідати реальним даним, що характеризуються асиметричними розподілами;

2) чутливість до викидів;

3) припущення про необмеженість діапазону значень не завжди виконується для фізичних величин.

Також зазначимо, що при проведенні класифікування припущення про нормальність розподілу є виправданим тільки для однорідних природних поверхонь (сільськогосподарських угідь чи водних об'єктів). Однак для гетерогенних міських територій або змішаних типів покриву це припущення може не виконуватися.

Тому постає задача, що полягає у залученні альтернативних статистичних розподілів для опису спектральних характеристик об'єктів земної поверхні для вирішення задач класифікування супутникових зображень методом максимальної правдоподібності [8, 9, 10].

Далі розглянемо застосування експоненціального розподілу та розподілу Вейбулла для опису спектральних характеристик земних об'єктів при проведенні процедури класифікування на основі критерію максимума функції правдоподібності.

Експоненціальний розподіл

При вирішенні задач ДЗЗ, а саме при обробці та класифікуванні супутникових зображень експоненціальний розподіл характеризує інтенсивності (квадрати амплітуд) радіолокаційних сигналів від шорстких поверхонь [3, 4].

При застосуванні методу максимальної правдоподібності інтенсивності сигналів пікселів класу можуть бути представлені у вигляді вибірки $X_1, X_2, \ldots, X_n$ незалежних однаково розподілених випадкових величин, що підпорядковуються експоненціальному розподілу з параметром інтенсивності $\lambda > 0$.

Густина розподілу одного спостереження X_i для $x_i \geq 0$ обчислюється наступним чином:

$$f\left(x_i \middle| \lambda\right) = \lambda e^{-\lambda x_i}, \tag{9}$$

де $\lambda > 0$ – параметр інтенсивності.

Графік експоненціального розподілу характеризується монотонно спадаючою формою. При цьому максимальне значення функції густини розподілу досягається у точці $x = 0$ та дорівнює параметру λ, після чого крива графіку експоненціально спадає до 0.

Слід зазначити, що даний розподіл характеризується дуже великою асиметрією. Швидкість спадання графіка визначається параметром λ, а саме: чим більше значення λ, тим швидший спад кривої.

Всі спостереження X_i, $i = 1, \ldots, n$ (значення інтенсивностей сигналів пікселів, які належать певного класу) є незалежними, то функція правдоподібності обчислюється як добуток густин, а саме:

$$L\left(\lambda \middle| x_1, \ldots, x_n\right) = \prod_{i=1}^{n} f\left(x_i \middle| \lambda\right). \tag{10}$$

Вираз (9) підставляємо у вираз (10), отримуємо:

$$L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda x_i}. \tag{11}$$

Далі вираз (11) перепишемо у вигляді:

$$L(\lambda) = \lambda^n \prod_{i=1}^{n} e^{-\lambda x_i}. \tag{12}$$

Використовуючи властивості показникової функції, отримуємо функцію правдоподібності:

$$L(\lambda) = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} x_i\right). \tag{13}$$

Для спрощення обчислень застосовуємо не саму функцію правдоподібності, а її натуральний логарифм.

Знаходимо натуральний логарифм функції правдоподібності, враховуючи, що логарифм добутку є сума логарифмів:

$$\ln L(\lambda) = \ln\left(\lambda^n\right) + \ln\left[\exp\left(-\lambda \sum_{i=1}^{n} x_i\right)\right]. \tag{14}$$

Застосовуючи властивості логарифмів, отримуємо:

$$\ln L(\lambda) = n\ln(\lambda) - \lambda\sum_{i=1}^{n} x_i. \quad (15)$$

Аналогічно, як у випадку з нормальним розподілом, застосовуємо критерій максимума функції правдоподібності, згідно якого піксель супутникового зображення має бути класифікований до того класу, якому відповідає максимальне значення логарифму функції правдоподібності, тобто найбільше значення виразу (15). Тобто ми розв'язуємо задачу максимізації функціоналу.

Основні переваги експоненціального розподілу:

1) нескладні розрахунки;

2) проста параметризація (густина розподілу описується одним параметром);

3) досконало досліджені статистичні властивості та методи оцінки параметрів;

4) застосування для моделювання радіолокаційних даних при вирішенні задач ДЗЗ;

5) ефективність при моделюванні класів земних об'єктів з високою варіабельністю.

Основні недоліки експоненціального розподілу:

1) монотонно спадаючі графіки функції густини розподілу не завжди відповідають реальним гістограмам оптичних даних;

2) із-за високої асиметрії даний розподіл не завжди точно апроксимує розподіл інтенсивностей сигналів пікселів деяких класів земних об'єктів.

При вирішенні задач ДЗЗ експоненціальний розподіл є оптимальним саме для радіолокаційних зображень. При проведенні класифікування знімків даний розподіл може бути застосований за наявності природних класів об'єктів із шорсткою поверхнею, а саме: піщаних поверхонь, води у вітряну погоду, деяких типів рослинності [3, 4].

Розподіл Вейбула

Розподіл Вейбула є узагальненням експоненціального розподілу і характеризується досить великою гнучкістю при моделюванні різноманітних типів даних [10, 11].

Аналогічно, як у попередньо розглянутих випадках при проведенні класифікування методом максимальної правдоподібності інтенсивності сигналів пікселів класу представляються у вигляді вибірки x_1, x_2, …, x_n незалежних однаково розподілених випадкових величин, які підпорядковуються розподілу Вейбулла з параметром форми $k > 0$ та параметром масштабу $\lambda > 0$.

Параметр масштабу λ визначає масштаб розподілу Вейбулла вздовж осі x. А графічні характеристики даного розподілу залежать від параметра форми kЯкщо $k < 0$то функція густини монотонно спадає. При $k = 1$ розподіл Вейбула співпадає з експоненціальним розподілом. За умови $k > 1$ положення максимума кривої зміщується вправо зі зростанням цього ж параметра k. Якщо параметр форми $k \approx 3{,}6$, то розподіл Вейбулла наближається до нормального.

Асиметрія розподілу Вейбулла також залежить від параметра форми k, що, в свою чергу, дозволяє моделювати як симетричні, так і асиметричні розподіли.

Густина розподілу Вейбулла розраховується за формулою (16):

$$f(x;k,\lambda) = \begin{cases} \dfrac{k}{\lambda}\left(\dfrac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}, & x \ge 0, \\ 0, & x < 0 \end{cases}. \quad (16)$$

$$L(k,\lambda \mid x_1,...,x_n) = \prod_{i=1}^{n} f(x_i;k,\lambda) = \prod_{i=1}^{n} \frac{k}{\lambda}\left(\frac{x_i}{\lambda}\right)^{k-1} e^{-(x_i/\lambda)^k}. \quad (17)$$

Вираз (17) перепишемо у наступній формі (18):

$$L(k,\lambda) = \left(\frac{k}{\lambda}\right)^n \left(\prod_{i=1}^{n} x_i^{k-1}\right) \lambda^{-n(k-1)} e^{-\sum_{i=1}^{n}(x_i/\lambda)^k}. \quad (18)$$

Знаходимо натуральний логарифм функції правдоподібності, спираючись на допоміжні додаткові обчислення та основні властивості натуральних логарифмів:

1) $\ln\left(\dfrac{k}{\lambda}\right)^n = \ln\left(\dfrac{k^n}{\lambda^n}\right) = n\ln k - n\ln\lambda;$

2) $\ln\left(\prod_{i=1}^{n} x_i^{k-1}\right) = (k-1)\ln\left(\prod_{i=1}^{n} x_i\right) = (k-1)\sum_{i=1}^{n}\ln x_i;$

3) $\ln[\lambda^{-n(k-1)}] = -n(k-1)\ln;$

4) $\ln e^{-\sum_{i=1}^{n}(x_i/\lambda)^k} = -\sum_{i=1}^{n}(x_i/\lambda)^k.$

Отже, натуральний логарифм функції правдоподібності розраховується за наступною формулою:

$$\ln L(k,\lambda) = n\ln k - nk\ln\lambda + (k-1)\sum_{i=1}^{n}\ln x_i - \sum_{i=1}^{n}\left(\frac{x_i}{\lambda}\right)^k. \quad (19)$$

Далі так само, як і у попередніх випадках, застосовуємо критерій максимума функції правдоподібності, згідно якого піксель знімку має бути віднесений до того класу об'єктів, якому відповідає максимальне значення логарифму функції правдоподібності, тобто найбільше значення виразу (19).

У задачах ДЗЗ розподіл Вейбулла зазвичай застосовується для моделювання амплітудних характеристик радіолокаційних сигналів. При чому параметр форми *k* вказує на ступінь однорідності класу земних об'єктів. Чим менші значення параметру *k*, тим даний клас є більш неоднорідним.

Слід зазначити, що розподіл Вейбулла є універсальним для радіолокаційних даних різних типів поверхонь та є найбільш оптимальним для неоднорідних класів зі змінною текстурою, а саме: мозаїчних ландшафтів, міських територій та перехідних зон [10, 11].

Основні переваги розподілу Вейбулла:

1) гнучкість та універсальність завдяки параметру форми *k*;

2) ефективність для класів із різною однорідністю і текстурою;

3) можливість застосування для задач обробки радіолокаційних даних різної природи;

4) можливість застосування до моделювання як симетричних, так і асиметричних розподілів інтенсивностей сигналів пікселів супутникових зображень;

5) високий ступінь відповідності реальним даним.

Основні недоліки розподілу Вейбулла:

1) потреба у вибірках великого об'єму із-за наявності двох параметрів (параметр форми та параметр масштабу);

2) потреба у більш складних розрахунках у порівнянні із однопараметричними розподілами;

3) чутливість оцінки параметра форми до викидів.

Порівняльний аналіз розподілів

Графік нормального розподілу має дзвоноподібну форму, симетричний та необмежений з двох боків. Графік експоненціального розподілу монотонно спадає. Графіки розподілу Вейбула можуть мати різні форми залежно від значень параметру форми *k*. Експоненціальний розподіл та розподіл Вейбулла визначені лише для невід'ємних значень.

Нормальний розподіл має один параметр, а експоненціальний розподіл та розподіл Вейбулла характеризуються двома параметрами. При цьому менше число параметрів зменшує потребу у навчальних даних та спрощує розрахунки, але обмежує гнучкість моделі. Двопараметричні розподіли (нормальний та Вейбула) краще адаптуються до реальних даних.

Нормальний розподіл характеризується помірною складністю. Експоненціальний розподіл найпростіший для обчислень із-за наявності тільки одного параметра. Розподіл Вейбула найскладніший через наявність двох параметрів та необхідність чисельної оптимізації при їх оцінці.

Нормальний розподіл є найчутливіший до викидів через квадратичну залежність у показнику експоненти. При цьому при обробці супутникових знімків графік нормального розподілу зображує викиди, як значення інтенсивностей сигналів пікселів, які розташовані далеко від центрального середнього. Іншими словами, викиди – це аномальні точки на графіку, що знаходяться на «хвостах розподілу» та з'являються з дуже низькою ймовірністю. Експоненціальний розподіл не є стійким до викидів, а великі викиди (екстремальні значення інтенсивностей сигналів пікселів) зображаються у вигляді асиметричного графіка з «важким хвостом». Розподіл Вейбулла має проміжну стійкість до викидів залежно від значення параметра форми *k*. При цьому стійкість до викидів забезпечується можливістю моделювати «хвости» розподілу [3, 10].

При $k = 1$ розподіл Вейбулла збігається з експоненціальним розподілом. Тобто експоненціальний розподіл є частинним випадком розподілу Вейбулла. А при $k \approx 3{,}6$ розподіл Вейбулла є майже ідентичним нормальному розподілу. У даному випадку форма кривої густини розподілу стає дзвіноподібною та симетричною, що є характерним саме для нормального розподілу. Дані взаємозв'язки між розглянутими розподілами є важливими для розуміння фізики різноманітних природних процесів та для правильного вибору розподілу для точного моделювання розподілу спектральних характеристик при вирішенні задач ДЗЗ [4, 11].

Висновки. Статистичні методи класифікування на основі критерію максимальної правдоподібності є одними із найбільш надійних підходів до обробки супутникових зображень. Основною перевагою методу максимальної правдоподібності є можливість врахування статистичної природи спектральних даних. Вибір статистичного розподілу має бути обґрунтованим для кожного конкретного класу земного покриву. Так, однорідна рослинність та водні об'єкти досить точно описуються нормальним розподілом. Але припущення про нормальність розподілу може не підходити для опису гетерогенних класів (змішаних типів покриву та забудов). Тому для гетерогенних територій слід використовувати більш гнучкі статистичні розподіли, а саме: експоненціальний розподіл та розподіл Вейбулла, оскільки вони надають більш достовірні результати.

Також у даній роботі було проведено детальний порівняльний аналіз нормального розподілу, експоненціального розподілу та розподілу Вейбулла, визначено їх основні переваги та оптимальні сфери застосування, враховуючи класи земного покриву та типи даних.

Було зазначено, що у випадку, коли апріорні ймовірності класів є рівними чи невідомими, то критерій максимальної правдоподібності зводиться до критерію максимума логарифму функції правдоподібності.

У подальших дослідженнях розвиток даної методології може включати застосування ансамблевих підходів із комбінуванням різних типів розподілів для різних класів та інтеграцію із методами машинного навчання для автоматичного вибору оптимальних статистичних моделей.

Список використаних джерел:

1. Alpert S. I. Data combination method in Remote Sensing tasks in case of conflicting information sources. *Ukrainian Journal of Remote Sensing*. 2021. Vol. 8(3). P. 44–48.

2. Chang C. I. Hyperspectral data processing. *Algorithm design and analysis. Hoboken, NJ: John Wiley and Sons*. 2013. 1164 p.

3. Frery A. C., Müller H. J., Yanasse C. C. F., Sant'Anna S. J. S. A model for extremely heterogeneous clutter. *IEEE Transactions on Geoscience and Remote Sensing*. 1997. Vol. 35(3). P. 648–659.

4. Gao G. Statistical modeling of SAR images: A survey. *Sensors*. 2010. Vol. 10(1). P. 775–795.

5. Li C., Wang J., Wang L., Hu L., Gong P. Comparison of Classification Algorithms and Training Sample Sizes in Urban Land Classification with Landsat Thematic Mapper Imagery. *Remote Sensing*. 2014. Vol. 6. P. 964–983.

6. Lu D., Weng Q. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*. 2007. Vol. 28(5). P. 823–870.

7. Myung J. Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*. 2003. Vol. 47(1). P. 90–100. DOI: 10.1016/S0022-2496(02)00028-7

8. Prokopenko I. G., Alpert S. I., Alpert M. I., Dmytruk A. Y. Classification of Sentinel-2 imagery using Rayleigh distribution modeling. *Electronics and Control Systems*. 2025. Vol. 2(84). P. 92–97.

9. Strahler A. H. The use of prior probabilities in maximum likelihood classification of remotely sensed data. *Remote Sensing of Environment*. 1980. Vol. 10. P. 135–163.

10. Weibull W. A statistical distribution function of wide applicability. *Journal of Applied Mechanics*. 1951. Vol. 18. P. 293–297.

11. Zhang C. W. Weibull parameter estimation and reliability analysis with zero-failure data from high-quality products. *Reliability Engineering and System Safety*. 2021. Vol. 207. 107321. DOI: 10.1016/j.ress.2020.107321

References:

1. Alpert S. I. (2021). Data combination method in Remote Sensing tasks in case of conflicting information sources. *Ukrainian Journal of Remote Sensing*, 8(3), 44–48.

2. Chang C. I. (2013). Hyperspectral data processing. Algorithm design and analysis. *Hoboken, NJ: John Wiley and Sons*. 1164 p.

3. Frery A. C., Müller H. J., Yanasse C. C. F., Sant'Anna S. J. S. (1997). A model for extremely heterogeneous clutter. *IEEE Transactions on Geoscience and Remote Sensing*, 35(3), 648–659.

4. Gao G. (2010). Statistical modeling of SAR images: A survey. *Sensors*, 10(1), 775–795.

5. Li C., Wang J., Wang L., Hu L., Gong P. (2014). Comparison of Classification Algorithms and Training Sample Sizes in Urban Land Classification with Landsat Thematic Mapper Imagery. *Remote Sensing*, 6, 964–983.

6. Lu D., Weng Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28(5), 823–870.

7. Myung J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, 47(1), 90–100. DOI: 10.1016/S0022-2496(02)00028-7

8. Prokopenko I. G., Alpert S. I., Alpert M. I., Dmytruk A. Y. (2025). Classification of Sentinel-2 imagery using Rayleigh distribution modeling. *Electronics and Control Systems*, 2(84), 92–97.

9. Strahler A. H. (1980). The use of prior probabilities in maximum likelihood classification of remotely sensed data. *Remote Sensing of Environment*, 10, 135–163.

10. Weibull W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, 18, 293–297.

11. Zhang C. W. (2021). Weibull parameter estimation and reliability analysis with zero-failure data from high-quality products. *Reliability Engineering and System Safety*, 207, 107321. DOI: 10.1016/j.ress.2020.107321

Дата першого надходження статті до видання: 03.03.2026
Дата прийняття статті до друку після рецензування: 26.03.2026
Дата публікації (оприлюднення) статті: 30.05.2026