



<https://doi.org/10.36023/ujrs.2023.10.2.232>

УДК 528.8.04:519.248

Застосування інформаційних дивергенцій до аналізу геосистем і процесів за даними дистанційного зондування

М. В. Артюшенко*, А. Ю. Порушкевич

ДУ "Науковий центр аерокосмічних досліджень Землі ІГН НАН України", вул. Олеся Гончара, 55-Б, Київ, 01054, Україна

Візуальне спостереження за змінами, що відбуваються в станах об'єктів та процесів на поверхні Землі, успішно вирішується засобами аерокосмічного моніторингу. Подальше вдосконалення інформаційних технологій моніторингу пов'язане з автоматизацією обробки та інтерпретації динамічних даних, поданих цифровими зображеннями. У статті обґрунтовується та на прикладах демонструється застосування інформаційних характеристик розбіжностей у зображеннях геополів: імовірнісна міра, ентропії Гіббса–Шеннона, Реньї. Наведено приклади обчислення різних функціоналів розподілу фізичних величин, поданих цифровими зображеннями, та характеристики ступеня їх близькості: розбіжність Кулбека, альфа-дивергенція (або дивергенція Реньї). Розглянутий підхід до аналізу геопроектів проілюстровано прикладом обчислення інформаційних дивергенцій температурного поля торфовища, отриманого як результат обробки даних з космічного апарата Landsat-8. Результати комп'ютерного моделювання розглянутого прикладу показують значну залежність розглянутих мір розбіжності від просторової розрізненості космічного знімання поля. Для правильного розрахунку інформаційних дивергенцій необхідно використовувати зображення, що отримані з однаковою просторовою розрізненістю. Подальший розвиток методів дивергенцій пов'язаний із впровадженням масштабно-інваріантних мір. Це дасть змогу використовувати сенсори з різною просторовою розрізненістю в системах аерокосмічного моніторингу для визначення динамічних змін геосистем та процесів.

Ключові слова: аерокосмічний моніторинг, зображення динаміки геополів, характеристики відмінності, ймовірність, міра, ентропії, інформаційні дивергенції, температурне поле торфовища

© М. В. Артюшенко, А. Ю. Порушкевич. 2023

Вступ

Багато задач аерокосмічного моніторингу полягають у визначенні змін, що відбуваються за певний період часу на великих площах у просторових і спектральних структурах природних об'єктів. Об'єктами досліджень є частини геосистем, елементи яких об'єднані спільними потоками енергії і речовини. Для характеристики динаміки структурних змін об'єктів у статті розглянуто і на прикладі показано як застосовувати кількісні статистичні та інформаційні характеристики відмінностей для оцінювання ступеня змін і передбачити необхідність проведення заходів для запобігання негативному розвитку процесів. Визначення за космічними знімками кількісних змін у геосистемах, дасть змогу автоматизувати процес спостереження і оповіщення про небезпечні явища. Вирішення цих завдань пов'язано з визначенням інформаційних розбіжностей (дивергенцій) розподілів контрольованих величин за космічними знімками. Для визначення розбіжностей розподілів дискретних величин використовуються різні інформаційні міри. Деякі перспективні статистичні

методи для використання інформаційних мір у задачах природокористування розглядаються в цій статті.

У результаті зондування створюються й аналізуються на цифрових знімках у різних спектральних діапазонах образи природних або штучних фізичних полів: спектральних яскравостей, радіолокаційних, лідарних вимірів та ін. Зокрема, в інфрачервоній мікрохвильовій області спектра формується поле радіаційної температури, структура температурного поля ландшафту з покладом торфу детально аналізується далі при обґрунтуванні інформаційних характеристик визначення розвитку торфової пожежі. Дослідження просторових розподілів фізичних величин на геометричних носіях зручно формулювати в термінах поля. Фізичне поле називається негладким, якщо функція поля є недиференційованою. В процесі комп'ютерної обробки даних, поданих космічними знімками, виявляється залежність структури поля, що спостерігається, від просторової розрізненості. Зі збільшенням розрізненості мінливість геофізичних полів зростає; обмеження на гладкість і диференційованість функції поля, на яких побудована прикладна теорія поля, стають слабко обґрунтованими. Залежності характеристик

*<https://orcid.org/0000-0002-7899-4450>

E-mail: mart47@i.ua

геосистем від масштабу вимірювань об'єктивно відображають неоднорідність і неізотропність середовища зондування. Такі особливості характерні для температурних полів на зображеннях міст, площах з покладами торфу і на територіях з гетерогенним ландшафтом. Причина виникнення негладких полів полягає в тому, що при підвищенні просторової розрізненості, зокрема інфрачервоного знімання, спостерігаються хаотично розташовані об'єкти, які мають різні характеристики теплових випромінювань. Стосовно площ, які мають гетерогенну ландшафтну структуру із вмістом покладів торфу, негладкі температурні поля поверхні формуються не лише під дією сонячної радіації, але також багато в чому визначається внутрішніми екзотермічними процесами в покладах. Перераховані чинники призводять до того, що зі збільшенням масштабу розгляду спостерігаються стрибки функції поля на множині елементарних поверхонь (пікселів зображення), що приймаються за точки поля. Отже, таке поле має безліч особливих точок, в яких функція поля не є такою, що диференціюється. У таких полях інакше відбуваються процеси дифузії, тепло- і масоперенесення. Необхідність дистанційного контролю за геополями негладкої структури зазвичай потребує високої просторової розрізненості, оскільки значення контрольованих величин істотно залежать від розмірів пікселів цифрових зображень. Очевидно, що опис процесів сучасної обробки сигналів і полів неможливо виконати за допомогою формул класичної регулярної математики, отриманих на основі подання гладких функцій.

Стосовно задач аналізу і класифікації цифрових зображень на цей час поширені методи, що ґрунтуються на засадах статистичних та інформаційних ентропій та ентропійних мір (Неро et al., 2001; Ампілова et al., 2014; Ампілова та ін., 2015; Jizba et al., 2004). Для багатьох просторово-розподілених геосистем коректно вивести детермінований закон еволюції спостережуваної системи або її частини не можливо з причини великої кількості непередбачуваних і невідомих чинників, що впливають на неї. Проте на мікрорівні доступними для спостереження є статистичні розподіли елементів системи, що дає змогу передбачати її поведінку і оцінювати чинники, що впливають на систему.

У статті розглянуто обґрунтування методів аналізу геосистем та їх станів за зображеннями з космосу, на засадах статистичних та інформаційних ентропій, інформаційних дивергенцій Кульбака і Реньї. У кінці статті розглянутий підхід до інтерпретації космічних даних ілюструється числовим прикладом статистичного аналізу цифрових зображень, отриманих за обробленими знімками з космосу, – температурного поля земної поверхні при виникненні і в процесі торфової пожежі.

Матеріали і методи дослідження

Аналіз функціонування багатьох природних систем (геосистем) демонструє детермінований

характер процесів (геопроектів), який обумовлений їх стохастичними властивостями. В цих системах регулярні детерміновані процеси формуються численними випадковими міжелементними взаємодіями, які відбуваються на мікрорівні. Характер перетворень випадкових взаємодій між елементами системи в регулярний процес залежить від виду системи і закономірностей її функціонування. Наприклад, в термодинаміці стани газоподібної системи можна визначити координатами і швидкостями дуже великої кількості частинок газу. Коли кількість елементів (частинок) велика, то підхід такого мікрорівня є неконструктивним. У такому випадку, для того, щоб досліджувати систему, не вивчаючи поведінку кожної частинки, в статистичній механіці вводиться концепція ентропії.

Аналіз геосистем та їх зображень на засадах статистичних та інформаційних ентропій.

У рамках системного аналізу для моделювання систем різних класів (фізичних, економічних, біологічних, організаційних і т.п.), що мають стохастичну властивість, були адаптовані і застосовані методи статистичної механіки, засновані на ентропії Гіббса, інформаційної ентропії Шеннона і на розрізняльній інформації (розбіжності, дивергенції) Кульбака, що на статистичному рівні задає міру відмінності станів системи (Вільсон, 1978; Хакен, 1991). Ентропія, як деяка математична функція стану системи, дає змогу пояснювати і передбачати багато макровластивостей системи, не вивчаючи поведінку кожного окремого її елемента, кількість і різноманітність яких в системі може бути дуже великою.

Ймовірність та ентропія Гіббса. Введення статистичної ентропії для дискретного випадку припускає наявність системи елементів, що може приймати дискретну множину станів X_i з імовірністю цих станів p_i :

$$X_i, i = \overline{1, w}, p_i = P\{X_i\}, 0 \leq p_i \leq 1, \sum_{i=1}^w p_i = 1. \quad (1)$$

Ймовірність p_i є мірою визначеності реалізації стану X_i . Якщо $p_i = 1$, то стан X_i є цілком достовірним, у випадку $p_i = 0$ стан X_i має мінімальну можливість реалізації.

За міру невизначеності стану X_i приймають величину $(-\log_b p_i)$. Для цілком достовірного стану, коли $p_i = 1$, невизначеність стану дорівнює нулю: $(-\log_b 1=0)$, а при $p_i = 0$ невизначеність стану нескінченно велика величина. При переході до натуральних логарифмів величина невизначеності стану має вигляд:

$$-k \ln p_i, k = 1/\ln b. \quad (2)$$

Статистична ентропія системи за Гіббсом визначається як математичне сподівання невизначеності, що дорівнює моменту першого порядку невизначеності станів системи:

$$S^G = M[-k \ln p_i] = -k \sum_{i=1}^w p_i \ln p_i, \quad (3)$$

де k – стала Больцмана.

Якщо всі стани системи мають однакову ймовірність $p_1 = p_2 \dots = p_w = 1/w$, відповідно до співвідношення (3) статистична ентропія максимальна і з точністю до сталої визначається тільки числом станів системи w :

$$S^B = k \ln w \propto \ln w. \quad (4)$$

Вираз (4) є вихідним визначенням статистичної ентропії за Больцманом. Запис ентропії (3) у вигляді безрозмірної величини без урахування сталої Больцмана k є

$$S^G = -\sum_{i=1}^w p_i \ln p_i. \quad (5)$$

Співвідношення (5) справедливе не тільки для фізичних систем, а й для значно ширшого класу систем, досліджуваних статистичними методами. Статистична ентропія системи є міра організованості системи. При рівномірному розподілі станів ентропія (5) максимальна, а система максимально неупорядкована. Цей випадок відповідає нижчому ступеню системної організації.

Інформаційна ентропія Больцмана–Гіббса–Шеннона. Інформаційна ентропія системи за Шенноном визначається як

$$H = -\sum_{i=1}^w p_i \log_b p_i, \quad (6)$$

де H – міра інформації, що одержують при повному з'ясуванні того, який стан реалізувався.

Якщо приймають $b = 2$, то обсяг інформації обчислюється в бітах. У випадку, коли $p_j = 1$ всі інші $p_j = 0$, $i \neq j$, величина $H = 0$. Це означає, якщо результат був відомий заздалегідь, то отримана інформація дорівнює нулю.

У випадку рівномірного розподілу, якщо $p_i = 1/w$, кількість інформації при повному з'ясуванні стану максимальна і дорівнює

$$H = \log_b w. \quad (7)$$

Аналіз отриманих результатів показує, що інформаційна ентропія, яка обумовлена виразами (6), (7), і статистична ентропія Гіббса (3), Больцмана (4) з точністю до сталої рівні. Це дає змогу прийняти вираз (6) за інформаційну ентропію Гіббса–Шеннона і використовувати її для статистичного аналізу просторово розподілених систем.

Важливою властивістю інформаційної ентропії (6), що дає змогу використовувати її в додатках як міру, є властивість адитивності. Властивість адитивності означає, що при об'єднанні двох незалежних систем їх ентропії додаються:

$$S = S_1 + S_2.$$

Дуже важливе обмеження ентропії Больцмана $S^B = \ln w$ складається з умови рівних ймовірностей усіх w станів системи. У статистичній механіці такі системи описуються мікроканонічним ансамблем (Хакен, 1991; Башкиров, 2006), що дає змогу записати формулу Больцмана у вигляді

$$S_i^B = -\ln p_i \text{ для усіх } i = \overline{1, w}. \quad (8)$$

Для випадку, якщо p_i різні, вводиться ансамбль мікроканонічних підсистем, таких, що для кожної i -ї підсистеми всі w_i її станів мають рівні ймовірності p_i ,

і ентропія Больцмана цієї підсистеми визначається виразом (8). Далі усереднення (обчислення моменту першого порядку) ентропії Больцмана ансамблю підсистем приводить до ентропії Гіббса–Шеннона:

$$S^G = \langle S_i^B \rangle = -\sum_i^w p_i \ln p_i. \quad (9)$$

З огляду на різні міркування, отримуємо одну формулу запису інформаційної ентропії Гіббса–Шеннона (5), (9), що, як уже відзначалося, справедливо не тільки для фізичних систем, а й для значно ширшого класу систем, досліджуваних статистичними методами (Башкиров та ін., 2007).

Однак, подальший розвиток ентропійних методів для опису і дослідження складних систем, показав, що перераховані вище ентропії не забезпечують відповідності із спостережуваними явищами, для яких визначені степеневі розподіли (Башкиров, 2006; Башкиров та ін., 2007). Степеневі розподіли притаманні різним складним системам як штучного, органічного, так і неорганічного походження (Вак, 1996; Newman, 2005), системам з фрактальною структурою (Федер, 1991; Schroeder, 1991), негладким полям, досліджувати які можна дистанційними аерокосмічними методами (Artiushenko, 2018a; Artiushenko, 2018b).

Функціонал ентропії Реньї. Цей вид ентропії був виведений Альфредом Реньї (Renyi, 1961; Башкиров та ін., 2007) на основі аксіоматичних вимог. Ентропія Реньї (10) так само, як і інформаційна ентропія Гіббса–Шеннона (9), задовольняє вимогу адитивності для статистично незалежних систем:

$$H_q^{(R)} = \frac{1}{1-q} \ln \sum_{i=1}^w p_i^q, \quad q \geq 0, \quad \sum_{i=1}^w p_i = 1. \quad (10)$$

При значенні параметра $q = 1$, з умови нормування ймовірностей p_i , логарифм статистичної суми у формулі ентропії Реньї дорівнює 0. У цьому випадку виникає невизначеність у записі (10), розкриття якого, за умови $q \rightarrow 1$, приводить до формули інформаційної ентропії Гіббса–Шеннона (9). Таким чином, було введено нове сімейство ентропії (для кожного значення q), що містить у собі ентропію Гіббса–Шеннона.

Методи ентропії, математичного і подальшого комп'ютерного моделювання є ефективним засобом дослідження широкого класу систем.

Обчислення інформаційних розбіжностей геосистем і геополів за зображеннями з космосу.

Дослідження просторових розподілів фізичних величин по поверхні Землі може бути подано у вигляді цифрового зображення. Нехай при проведенні геомоніторингу з космосу в кожен дискретний момент часу область земної поверхні \tilde{V} відображається в кадрі V цифрового зображення, образом якого є стаціонарне поле $U = f(x, y)$ деякої фізичної величини. Цифрове зображення поля розміром $m \times n$, що побудоване на растрі пікселів, задається матрицею \mathbf{K} . Значення матриці

$K(i, j)$, $i = \overline{1, m}$, $j = \overline{1, n}$ дорівнюють квантованим яскравостям пікселів у “точках” $(i, j) \in V$ і відповідають значенням поля з площі що є проєкцією піксела на область \tilde{V} . Цифрова функція $K(i, j)$ – це і є цифрове зображення просторового розподілу фізичної величини.

Розглянемо розподіл значень U двовимірного поля області \tilde{V} і характеристики цього розподілу, що визначаються за розподілом яскравостей у кадрі V цифрового зображення. Термін “розподіл” слід розуміти у ймовірнісному змісті як співвідношення, що встановлює зв’язок між можливими значеннями випадкової величини і відповідними їм ймовірностями. Для опису властивостей розподілів використовуються числові характеристики, такі як моменти розподілу випадкової величини. Початковий момент M_q порядку q випадкової дискретної величини X з n значеннями x_i і ймовірностями цих значень v_i визначається відповідно до формули

$$M_q = \sum_{i=1}^n x_i^q v_i, \quad (11)$$

де q належить множені натуральних чисел.

Нехай деяка фізична величина розподілена по об’єкту на земній поверхні і подана зображенням. Розіб’ємо кадр V на N рівних квадратних клітин зі стороною r . Нехай поле об’єкта подано на зображенні в градаціях яскравості. Для характеристики розподілу яскравості по зображенню значення яскравості пікселів у кожній i -й клітині додаються, і яскравість клітини нормується на одиницю. Нормовані яскравості клітин розглядаються як ймовірності p_i :

$$\sum_{i=1}^N k_i = K, \quad p_i = \frac{k_i}{K}, \quad \sum_{i=1}^N p_i = 1, \quad (12)$$

де p_i – нормовані значення яскравості в i -й клітині зображення; k_i – квантовані значення яскравості клітини з номером i ; K – загальна яскравість квантованих значень зображення поля, що не залежить від розміру клітин розподілу кадру; N – сумарна кількість зайнятих зображенням клітин.

Визначивши значення яскравості k_i і ймовірності цих значень p_i , за формулою (12) обчислюємо розподіл яскравості і відповідних їм фізичних величин (наприклад, температури), по області земної поверхні. Повний набір

$$\{p_i\}_{i=1}^N \quad (13)$$

при мінімальному значенні r (розмір піксела), дає табличний опис розподілу яскравості зображення. Однак кількість клітин, що дорівнює N , надто велика і залежить від розміру клітин. Така характеристика не дасть змоги порівнювати поля за їхніми розподілами – необхідно одержати числові характеристики розбіжностей у розподілах фізичних величин, поданих ймовірностями розподілів яскравостей двох зображень.

Інформаційні міри розбіжностей за С. Кульбаком. Для визначення розбіжностей розподілів дискретних величин використовуються

різні інформаційні міри, що ґрунтуються на ентропійних характеристиках. Найбільш відомими є інформаційні міри розбіжностей Кульбака (Кульбак С., 1967), що добре адаптуються для визначення просторових розподілів за їх зображеннями, та ґрунтуються на формулі ентропії Гіббса–Шеннона (9). Для двох величин X та Y зі значеннями x , y , визначених в областях, що аналізуються, розподіл імовірності знаходиться зокрема так. Области покриваються рівними клітинами з номерами $i = 1, 2, \dots, N$, у межах яких визначаються ймовірності (12) знаходження випадкових величин та їх розподіли згідно із співвідношеннями (13), (14):

$$p_i = p(x_i), P = \{p_i\}, \quad g_i = g(y_i), G = \{g_i\}. \quad (14)$$

За цих умов обчислюються такі інформаційні розбіжності:

асиметрична міра Кульбака

$$B_K(P\|G) = \sum_{i=1}^N [p_i \ln \frac{p_i}{g_i}] \quad (15)$$

та симетрична міра Кульбака

$$B_{K_S}(P\|G) = \sum_{i=1}^N [(p_i - g_i) \ln \frac{p_i}{g_i}]. \quad (16)$$

Визначення мір розбіжностей може провадитись як для фізичних величин одного типу, так і для різних. Ймовірності визначаються як відносні значення величин, що розташовані в клітинах, сума яких (12) по всій області дорівнює одиниці. У зв’язку з тим, що неперервні форми величин при використанні даних ДЗЗ і польових вимірювань не мають застосування, неперервні форми розбіжностей Кульбака не розглядаються. Недоліком форми (15) є те, що вона не є симетричною: $B_K(P\|G) \neq B_K(G\|P)$, і не може розглядатися як метрика. Наприклад, якщо G є еталонний розподіл, а P розподіл на досліджуваній ділянці, то інформаційна відстань від розподілу G до P не дорівнює відстані від P до G , що не дає змоги використовувати форму (15) для визначення розбіжностей на маршрутах обстежень, або застосувати її, якщо еталонів декілька. Форма (16) дивергенції Кульбака не має такої обмеженості.

Загальна обмеженість використання форм розбіжності (15) та дивергенції Кульбака (16), що побудовані на формулі ентропії Гіббса–Шеннона, як критеріїв схожості, стає зрозумілим, якщо деяка фізична величина (наприклад температура або спектральна інтенсивність) подані зображеннями. Визначити схожість зображень лише за одним критерієм недостатньо.

Розбіжність Реньї або alpha-дивергенція та інші розбіжності. Перспективним слід вважати використання функціонала дивергенції Альфреда Реньї (Renyi, 1960; Hero et al., 2001; Ampilova et al., 2014), що як і ентропія Реньї (10) побудована на моментах розподілу ймовірностей за клітинами $i = 1, 2, \dots, N$, на які розбито область, що аналізується. Розбіжність Реньї порядку q із розподілу $P = \{p_i\}$ у розподіл $G = \{g_i\}$ визначається для різних значень параметра q із співвідношення:

$$A_q(P\|G) = \frac{1}{q-1} \ln \left(\sum_{i=1}^N \frac{p_i^q}{g_i^{q-1}} \right) = \frac{1}{q-1} \ln \left(\sum_{i=1}^N p_i^q g_i^{1-q} \right), \quad q > 0, q \in \mathbf{R}, \quad (17)$$

що дає змогу оцінити розбіжності за багатьма критеріями. Для двох розподілів, що порівнюються, обчислюється вектор розбіжностей A_q , який розглядається як характеристика схожості.

Як і розбіжність Кульбака (6), розбіжність Реньї не є від'ємною $A_q(P\|G) > 0$ при $q > 0$; при рівних розподілах $A_q(G\|G) = 0$; також міра A_q розбіжностей є симетричною $A_q(P\|G) = A_q(G\|P)$.

Розбіжність Реньї узагальнює низку відомих інформаційних розбіжностей, наприклад:

- подвійний від'ємний логарифм від коефіцієнта Бахаттачарія, $q = 1/2$:

$$A_{q=1/2}(P\|G) = -2 \ln \sum_{i=1}^N \sqrt{p_i g_i}, \quad (18)$$

- розбіжність Кульбака, див. (16), $q = 1$:

$$A_{q=1}(P\|G) = \sum_{i=1}^N [p_i \ln \frac{p_i}{g_i}], \quad (19)$$

- логарифм очікуваного відношення ймовірностей, $q = 2$:

$$A_{q=2}(P\|G) = \ln \left\langle \frac{p_i}{g_i} \right\rangle. \quad (20)$$

Апробація методу узагальнювальної alpha-дивергенції для визначення кількісних змін у негладких температурних полях, за даними з TIRS Landsat-8. Враховуючи те, що alpha-дивергенція узагальнює, зокрема, і розбіжність Кульбака (при $q = 1$), а також інші інформаційні міри, обмежимося при моделюванні розбіжностей температурних полів за зображеннями демонстрацією застосування alpha-дивергенції.

Вибір об'єкта досліджень. Для апробації методів інформаційних мір розбіжностей (дивергенцій) геополів необхідно було вибрати територію з різними характеристиками теплових відбиттів і випромінювання. Наприклад, територію, що містить поклад торфу та має виражений гетерогенний ландшафт: вода, трава, відкритий ґрунт. Негладке температурне поле формується не тільки під дією сонячної радіації, а й визначається внутрішніми екзотермічними процесами в покладі торфу.

Початковими даними дослідження методів визначення розбіжностей слугують значення температури земної поверхні, отримані в роботі (Lischenko et al., 2016) як результат обробки з космічного апарата Landsat-8 даних довгохвильового інфрачервоного випромінювання території. В експерименті розбіжності аналізуються для області розміром 7680×7680 кв. м, значення температури подані матрицею розміром 256×256 чисел, (див. Рис. 1, виконано Криловою Г. Б.). Дані, отримані з приладу TIRS із просторовим розрізненням 100 м, інтерпольовані постачальником даних до 30 м. Для подальшого аналізу використано три знімки за

2015 рік: 12 серпня, 28 серпня, 15 жовтня. Окрім покладу торфу ландшафтний склад обраної області містить водну поверхню, ділянки відкритого ґрунту і рослинність: лісову, лугову, болотяну. За наземними спостереженнями на 15 жовтня 2015 року виявлено осередки прихованого горіння торфу. 12 серпня, 28 серпня пожежа не спостерігалась.

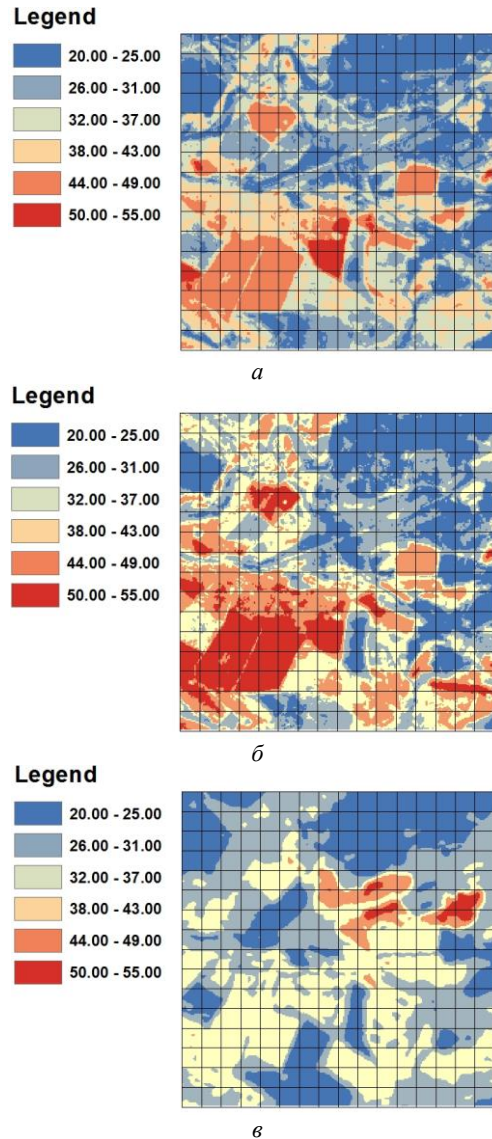


Рис. 1. Зображення поля температури поверхні під час розвитку торфової пожежі у 2015 році на дату: $a - (t_1 - 12$ серпня), $b - (t_2 - 28$ серпня), $v - (t_3 - 15$ жовтня, пожежа)

На Рис. 1 a, b, v зображені поля температури на три дати знімання 2015 року та встановлені такі розподіли: ($t_1 - 12$ серпня, розподіл температур G), ($t_2 - 28$ серпня, розподіл P_2) і ($t_3 - 15$ жовтня, розподіл P_3). Окрім покладу торфу ландшафтний склад обраної області містить водну поверхню, ділянки відкритого ґрунту і рослинність: лісову, лугову, болотяну.

Результати дослідження

Індикатор розбіжностей температурного поля. Розрахунок alpha-дивергенцій температурних полів за зображеннями на Рис. 1 виконаний згідно з

виразами (12), (17). Для обчислення дивергенцій розмір клітин покриття зображень r визначається у відносних одиницях до розміру зображення і становить $r = 1/16$ розміру квадратного кадру зображення. На Рис. 2 наведено розбіжності температурного поля торфовища, заданого трьома зображеннями (Рис. 1 а, б, в), за інформаційною мірою Реньї $A_q(P|G)$. Розбіжності визначено відносно зображення на Рис. 1 а, що має розподіл температур G .

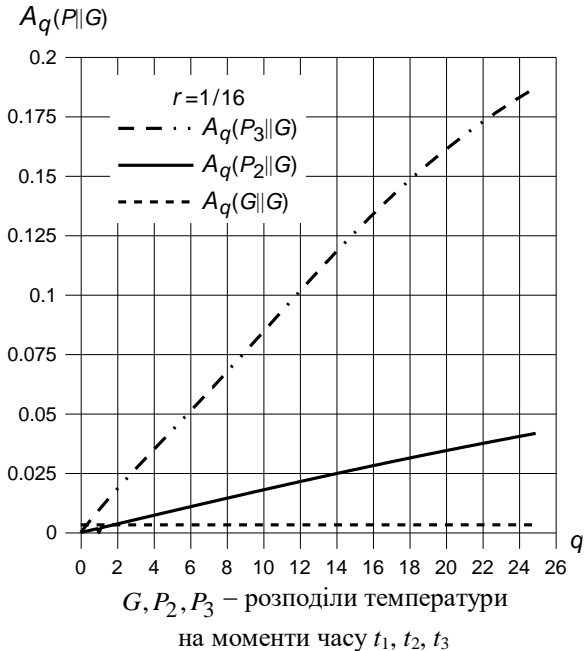


Рис. 2. Alpha-дивергенція температурного поля торфовища, або сімейство A_q розбіжностей Реньї

Розбіжності на Рис. 2 обчислено на моменти часу t_2 , t_3 відносно моменту t_1 . Розглянуті графіки демонструють монотонну залежність функціональної міри розбіжностей від параметра q . Швидкість зростання міри низька для малих відмінностей в розподілах і стає великою при значних інформаційних розбіжностях $A_q(P_3|G)$ між двома зображеннями (Рис. 1 а, Рис. 1 в), що визначає пожежу на торфовищі. За даними авторів роботи (Lisichenko, Razynych, 2016), проводилися польові дослідження торфовища та виявлено осередки прихованого горіння торфу на 15.10.2015 р. На зображенні з космосу (Рис. 1 в) неможливо за температурою визначити горіння торфу. За даними знімків з космосу, максимальна, мінімальна і середня температури поверхні в градусах за Цельсієм становили відповідно: Рис. 1 а – (50,88; 27,99; 36,14), Рис. 1 б – (45,0; 22,35; 31,29), Рис. 1 в – (51,42; 12,48; 22,53).

Таблиця. Інформаційні розбіжності температурного поля

q	1/2	2	6	8	10
$A_q(G G), t_1$	0.0	0.0	0.0	0.0	0.0
$A_q(P_2 G), t_2$	0.0009055	0.003626	0.01086	0.01444	0.01796
$A_q(P_3 G), t_3$	0.004871	0.01836	0.05107	0.06757	0.08451

Обчислення виконано при $r = 1/16$.

У таблиці надано значення розрахованих альфа-дивергенцій (величина безрозмірна) температурного поля, що використовуються як індикатори розбіжностей та стан температурного поля на дати: (1) – 12 серпня, 224 доба), б – (2) – 28 серпня, 248 доба), в – (3) – 15 жовтня, 288 доба).

Цей експеримент дає змогу зробити висновок про те, що безпосередньо за значеннями температури поверхні торфовища, що досліджується з космічного апарата, визначити стан підземного екзотермічного процесу та момент займання поклада торфу складно (див. Рис. 1). Однак, використовуючи методи альфа-дивергенцій розподілу температур, можна впевнено характеризувати стан процесу статистичними методами і створеними індикаторами. Для цього можна використовувати як числові дані, так і графічні, що забезпечуються комп'ютерною програмою обчислення дивергенцій.

З графіка (Рис. 2) і таблиці випливає, що методи розбіжностей (18)–(20) як індикаторні міри при: $q = 1/2$, $q = 1$, $q = 2$ не забезпечать прийнятної точності обчислень. Доцільно використовувати інформаційні розбіжності як індикаторні міри при значеннях $q = \{6, 8, 10\}$.

Дослідження залежності індикатора альфа-дивергенції від просторової розрізненості.

Залежність альфа-дивергенції від просторової розрізненості космічного знімання можна відстежити як залежність розбіжностей від розміру r клітин покриття цифрових зображень, що аналізуються. На Рис. 3 надано графіки, що демонструють залежність від розміру r при обчисленні розбіжностей зображення температурного поля торфовища (Рис. 1) на момент часу пожежі t_3 (розподіл температури P_3), відносно зображення на Рис. 1 а (розподілом температури G). Обчислення проведено для розмірів клітин покриття $r = \{1/2, 1/4, 1/8, 1/16, 1/64\}$.

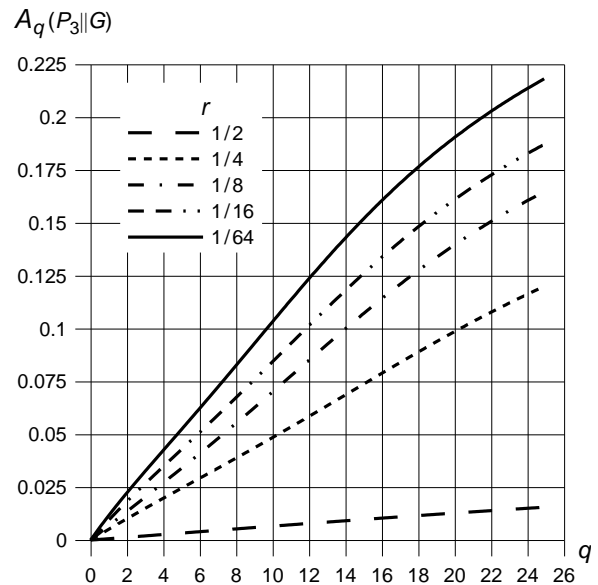


Рис. 3. Графіки залежностей значень альфа-дивергенції зображення на момент t_3 (розподіл P_3 температури) відносно моменту часу t_1 (розподіл G) від розміру клітин покриття r ; моделює залежність розбіжностей A_q від просторової розрізненості

Висновки

Для тематичних завдань космічного геомоніторингу, що полягають у визначенні змін, які відбуваються за певний період часу в просторових і спектральних структурах природних об'єктів, розроблено, науково обґрунтовано та на прикладах продемонстровано, як застосовувати інформаційні характеристики відмінності, які на кількісному рівні дадуть змогу автоматизувати процес пошуку за космічними знімками таких змін, оцінити ступінь змін і обґрунтувати необхідність проведення заходів щодо запобігання негативного розвитку процесів. Адаптовано низку перспективних статистичних методів для розробки інформаційних мір у задачах природокористування: ентропії Гіббса–Шеннона, Реньї, міри розбіжностей Кульбака, альфа-дивергенція. Геопросторовим модулюванням та порівняльним аналізом застосування різних інформаційних мір як індикаторів стану процесів у геосистемах визначено доцільність введення мір на засадах методу альфа-дивергенції. Недоліком індикаторів стану процесів при використанні мір методу альфа-дивергенції є їх залежність від просторової розрізненості космічного знімання. Міри не є масштабно-інваріантними, що потребує проведення аерокосмічного спостереження за процесом, використовуючи сталу просторову розрізненість. Подальший розвиток цього напрямку аналізу геопроектів пов'язаний з уведенням масштабно-інваріантних мір, що дасть змогу використовувати їх як індикатори, які мало залежать від просторової розрізненості знімання.

Література

- Ампілова, Н. Б., Сергеев, В. Д., Соловьев, И. П. (2015). *Применение расхождения Реньи к анализу и классификации изображений*. https://lib.herzen.spb.ru/media/magazines/contents/1/176/ampilova_176_35_44.pdf
- Башкиров, А. Г. (2006). Энтропия Реньи как статистическая энтропия для сложных систем. *Теоретическая и матем. физика*. 149, № 2. 299–317.
- Башкиров, А. Г., Витязев, А. В. (2007). Энтропия Реньи и степенные распределения в естественных и гуманитарных науках. *ДАН РФ*. 412, № 4. 476–479.
- Вильсон, А. Дж. (1978). *Энтропийные методы моделирования сложных систем. Теория и методы системного анализа*. Москва: Наука. Глав. ред. физ.-мат. лит.
- Кульбак, С. (1967). *Теория информации и статистика* / под ред. и с предисл. Колмогорова, А. Н. Москва: Наука.
- Федер, Е. (1991). *Фракталы*. Москва: Мир.
- Хакен, Г. (1991). *Информация и самоорганизация. Макроскопический подход к сложным системам*. Москва: Мир.
- Ampilova, N., Soloviev, I. (2014). On Application of Entropy Characteristics to Texture Analysis. *WSEAS Transactions on Biology and Biomedicine*. Vol. 11, Art. 25, 194–202.
- Artiushenko, M. (2018a). Statistical analysis of the unsmooth geophysical fields by remote sensing data. *J. of Automation and Information Sciences*. 50(6), 14–27. DOI: 10.1615/JAutomatInfScien.v50.i6.20
- Artiushenko, M. (2018b). Identification and Interpretation of Power-Law Distributions by Spectral Data of Remote Sensing. *J. of Automation and Information Sciences*. 50(12), 17–33. DOI: 10.1615/JAutomatInfScien.v50.i12.20

- Bak, P. (1996). *How Nature Works: The Science of Self-Organized Criticality*. New York: Copernicus.
- Hero, A. O, Ma Bing, Michel, O., Gorman, J. (2001). Alpha-Divergence for Classification, Indexing and Retrieval. *Communications and Signal Processing Laboratory Technical Report CSPL*, 3–16.
- Jizba, P., Arimitsu, T. (2004). The world according to Rényi: thermodynamics of multifractal systems. *Annals of Physics (Elsevier)*. 312, 17–59. DOI:10.1016/j.aop.2004.01.002.
- Lisichenko, L., Pazynych, N., Fedorov O. (Ed.) (2016). Remote study of peatlands in connection with their flammability risk. *Space research in Ukraine 2014–2016* / Kyiv: Akadem periodyka. 54–59. ISBN 978-966-360-317-9.
- Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *J. Contemporary Physics*, 46, 323–351.
- Renyi A. (1960, June 20 – July 30) On Measures of Entropy and Information. *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, 547–561.
- Schroeder, M. (1991). *Fractals, chaos, power laws*. New York: W.H. Freeman and Co. ISBN 0-7167-2136-8.

References

- Ampilova, N., Sergeev, V., Soloviev, I. (2015). An application of Renyi divergence to image analysis and classification. (in Russian). https://lib.herzen.spb.ru/media/magazines/contents/1/176/ampilova_176_35_44.pdf.
- Ampilova, N., Soloviev, I. (2014). On Application of Entropy Characteristics to Texture Analysis. *WSEAS Transactions on Biology and Biomedicine*. Vol. 11, Art. 25, 194–202.
- Artiushenko, M. (2018a). Statistical analysis of the unsmooth geophysical fields by remote sensing data. *J. of Automation and Information Sciences*. 50(6), 14–27. DOI: 10.1615/JAutomatInfScien.v50.i6.20.
- Artiushenko, M. (2018b). Identification and Interpretation of Power-Law Distributions by Spectral Data of Remote Sensing. *J. of Automation and Information Sciences*. 50(12), 17–33. DOI: 10.1615/JAutomatInfScien.v50.i12.20.
- Bak, P. (1996). *How Nature Works: The Science of Self-Organized Criticality*. New York: Copernicus.
- Bashkirov, A. (2006). Renyi entropy as statistical entropy for complex systems. *TMF*, 149(2), 299–317. (in Russian).
- Bashkirov, A., Vityazev, A. (2007). Renyi entropy and power-law distributions in natural sciences and humanities. *DAN RF*, 412(4), 476–479. (in Russian).
- Feder, J. (1991). *Fractals*. Moscow: Mir. (in Russian).
- Hero, A. O, Ma Bing, Michel, O., Gorman, J. (2001). Alpha-Divergence for Classification, Indexing and Retrieval. *Communications and Signal Processing Laboratory Technical Report CSPL*, 3–16.
- Jizba, P., Arimitsu, T. (2004). The world according to Rényi: thermodynamics of multifractal systems. *Annals of Physics (Elsevier)*. 312, 17–59. DOI:10.1016/j.aop.2004.01.002.
- Kullback, S. (1967). *Information Theory and Statistics*. Moscow: Nauka. (in Russian).
- Lisichenko, L., Pazynych, N., Fedorov, O. (Ed.) (2016). Remote study of peatlands in connection with their flammability risk. *Space research in Ukraine 2014–2016*. Kyiv: Akadem periodyka. 54–59. ISBN 978-966-360-317-9.
- Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *J. Contemporary Physics*, 46, 323–351.
- Renyi, A. (1960, June 20 – July 30) On Measures of Entropy and Information. *Proceedings of the 4th Berkeley Symposium on Mathematics, Statistics and Probability*, 547–561.
- Schroeder, M. (1991). *Fractals, chaos, power laws*. New York: W.H. Freeman and Co. ISBN 0-7167-2136-8.

APPLICATION OF INFORMATION DIVERGENCES TO THE ANALYSIS OF GEOSYSTEMS AND PROCESSES USING REMOTE SENSING DATA

M. V. Artiushenko, A. Yu. Porushkevich

Scientific Centre for Aerospace Research of the Earth of the Institute of Geological Sciences of the National Academy of Sciences of Ukraine, Olesia Honchara Str., 55-b, Kyiv, 01054, Ukraine

Visual observation of changes occurring in the states of objects and processes on the Earth's surface is successfully solved through aerospace monitoring. Further improvement of information technologies of monitoring is associated with automating processing and interpretation of dynamic data represented by digital images. The article substantiates and demonstrates using of information characteristics of differences in images of geofield: a probabilistic measure, Gibbs-Shannon and Renyi entropy. Examples of calculation of various functionals of distributions of physical quantities represented by digital images and characterizing the degree of their proximity are given: Kullback divergence, alpha-divergence or Renyi divergence. The considered approach is illustrated by the example of calculating the informational divergences of the peatland temperature field obtained as a result of data processing using images taken by Landsat-8 satellite. The results of computer simulation of the above example show a significant dependence of the considered divergence measures on the spatial resolution of the field images. For the correct calculation of informational divergences, it is necessary to use images obtained with the same spatial resolution. Further development of information divergence methods is associated with introducing of scale-invariant measures. This will make it possible to use sensors with different spatial resolutions in the aerospace monitoring system to determine dynamic geosystem changes and processes.

Keywords: aerospace monitoring, images of geofield dynamics, difference characteristics, probability measures, entropy, informational divergences, analysis of peatland temperature field

Рукопис статті отримано 06.03.2023

Надходження остаточної версії: 27.05.2023

Публікація статті: 29.06.2023