

УДК 519.87:303.732

*М.В. Артюшенко, А.В. Хижняк*

## ЭНТРОПИЙНЫЕ МЕТОДЫ МОДЕЛИРОВАНИЯ САМООРГАНИЗУЮЩИХСЯ СИСТЕМ ПО ДАННЫМ ГИПЕРСПЕКТРАЛЬНОГО ЗОНДИРОВАНИЯ

**Ключевые слова:** информационные энтропии, статистический анализ систем, самоорганизация, спектральные данные, численное моделирование, поиск углеводородов.

### Введение

Анализ функционирования многих природных систем (геосистем) демонстрирует детерминированный характер процессов (геопроцессов), который обусловлен их стохастическими свойствами. В рассматриваемых системах регулярные детерминированные процессы формируются в результате многочисленных случайных межэлементных взаимодействий, которые происходят на микроуровне. Характер преобразований случайных взаимодействий между элементами системы в регулярный процесс зависит от вида системы и закономерностей ее функционирования. Так, например, состояния газообразной системы можно определить, используя координаты и скорости очень большого количества частиц газа. Когда количество элементов (частиц) большое, то такой микроуровневый подход является неконструктивным. В этом случае, для того чтобы исследовать систему, не изучая поведения каждой частицы, в статистической механике вводится концепция энтропии. В рамках системного анализа для моделирования систем различных классов (физических, экономических, биологических, организационных и т.п.), обладающих таким свойством, были адаптированы и применены методы статистической механики, основанные на энтропии Гиббса, информационной энтропии Шеннона и на различающей информации (расхождении) Кульбака, которая на статистическом уровне задает меру различия состояний системы [1, 2]. Энтропия системы как некоторая математическая функция состояния позволяет объяснять и предсказывать многие макросвойства системы, не изучая поведения каждого отдельного ее элемента, количество и разнообразие которых в системе может быть очень большим. Энтропийные методы математического и последующего компьютерного моделирования являются эффективным средством исследования широкого класса сложных систем, обладающих способностью к самоорганизации. В данной работе авторы придерживаются определения самоорганизации, которое сформулировал Г. Хакен [2]. Согласно этому определению, самоорганизующаяся система обладает способностью обретать пространственную, временную или функциональную структуру без какого-либо специфического воздействия на систему извне, которое навязывает ей эту структуру. В статье данное определение распространяется на исследование спектральных структур геосистем.

Для многих пространственно-распределенных геосистем корректно вывести детерминированный закон эволюции наблюдаемой системы или ее части не представлено.  
© М.В. АРТЮШЕНКО, А.В. ХИЖНЯК, 2021  
*Международный научно-технический журнал  
«Проблемы управления и информатики», 2021, № 2*

ляется возможным ввиду большого числа непредсказуемых и неизвестных факторов, оказывающих на нее влияние. Однако на микроуровне доступными для наблюдения являются статистические распределения элементов системы, что позволяет предсказывать ее поведение и оценивать факторы, действующие на систему. Последующее развитие энтропийных методов моделирования для описания и исследования сложных самоорганизующихся систем показало, что перечисленные выше энтропии не обеспечивают согласия с явлениями, для которых имеют место степенные распределения [3, 4]. Степенные распределения характерны для различных классов сложных самоорганизующихся систем как органического, так и неорганического происхождения [5, 6], систем с фрактальной структурой [7], исследование которых можно проводить дистанционными аэрокосмическими методами [8].

В работе [9] для идентификации степенного распределения по данным гиперспектрального зондирования самоорганизующейся природной системы был предложен метод и проведена его валидация. Для этой демонстрации был выбран спектральный диапазон с хорошим приближением к степенному виду распределения. Следовательно, исследования таких режимов работы самоорганизующихся систем должны базироваться на другом виде энтропии, отличающемся от информационной энтропии Гиббса–Шеннона [3, 4]. Особенностью степенного закона распределения, называемого также законом Ципфа, распределением Парето, или распределением с «тяжелым хвостом», является медленное убывание вероятностей больших значений случайной величины. В случае нестепенной статистики область больших значений величины характеризуется малой вероятностью. Так, например, при распределении Гаусса, которому подчиняется большое количество случайных процессов, имеет место очень быстрое убывание плотности вероятности по мере отклонения величины от математического ожидания, что позволяет не учитывать в расчетах вероятность крупных событий. В системах со степенным законом распределения крупные события происходят недостаточно редко по сравнению с экспоненциальной формой распределения Гиббса, на котором основана классическая статистическая термодинамика [4], в силу чего графики степенных распределений имеют «тяжелые хвосты», что свидетельствует о возможности эмерджентного развития катастрофических явлений. К распределениям с «тяжелыми хвостами» также относятся распределения: Парето, логнормальное, Леви, Вейбула и др. Перечисленные выше особенности класса сложных самоорганизующихся систем привели к смене парадигмы использования методов классической статистической термодинамики. Так, в частности, для анализа природных систем по данным дистанционного зондирования используются методы интерпретации данных, основанные на обобщенной энтропии Реньи [8].

В статье рассмотрены энтропийные методы моделирования самоорганизующихся систем растительных покровов и на примерах демонстрируется применение этих методов для определения на месторождениях границ залежей углеводородов. Актуальность решения этой задачи обусловлена тем, что во многих регионах ресурс легкооткрываемых залежей углеводородов исчерпан, и тенденцией развития геологоразведочных работ является выявление глубокозалегающих, слабопроявленных залежей на уже промышленно освоенных месторождениях. Однако значительная недостоверность существующих индикаторов скопления углеводородов (залежей) и методов их определения, включая и основные геофизические методы сейсморазведки, далеко не всегда приводят к ожидаемым результатам, что является неприемлемым для чрезвычайно дорогого поискового нефтегазового бурения, особенно глубинного. В связи с этим развиваются недорогие методы определения дополнительных индикаторных признаков залежей углеводородов, к которым относятся и методы фи-

тоиндикации. В результате многочисленных исследований установлено, что пространственно-спектральные характеристики растительного покрова существенно изменяются в аномальных районах, связанных с месторождениями нефти и газа. Формирование биохимических аномалий растительного покрова над залежью углеводородов связано с вертикальной миграцией газоподобных углеводородов в корневой слой грунта. Биохимические аномалии и изменения отражательной способности покровов хорошо определяются по изменениям спектральных характеристик (спектрограммам) растений [9]. Эти изменения происходят в результате процесса самоорганизации оптической системы листьев, способной дозировать необходимое для нормального развития растения количество поглощаемой и отраженной световой энергии в зависимости от абиотических факторов среды.

Статья завершается иллюстрацией применения рассмотренных энтропийных моделей в численных экспериментах с реальными данными, полученными с месторождения природного газа. На основе данных гиперспектральных излучений гомогенного растительного покрова [9] проводится верификация применения методов: расхождения Кульбака, дивергенции Реньи ( $\alpha$ -дивергенция, расхождение Реньи) и статистических фракталов (мультифракталов) для определения границ залежей углеводородов.

### **Цели, объект и предмет исследований**

Ниже перечислим цели исследований.

1. Разработка энтропийных методов моделирования изменений в состояниях природных систем на основе пространственно-распределенных спектральных данных дистанционного зондирования.

2. Верификация энтропийных методов интерпретации данных дистанционного зондирования — для определения на месторождениях границ залежей углеводородов.

Объект исследований — процесс изменений состояния растительного покрова, произрастающего над залежами углеводородов.

Предмет исследований — энтропийные методы анализа изменений в состояниях сложных самоорганизующихся систем на основе спектральных данных.

### **Методы моделирования**

**Способ получения исходных данных.** Используется гиперспектральное зондирование, которое основано на разложении светового потока от объекта зондирования с помощью дифракционной решетки или оптической призмы, установленной в оптико-электронном приборе, на сотни узких смежных спектральных поддиапазонов длин волн. В статье термины «гиперспектральные» и «спектральные» данные используются как эквивалентные. Для того чтобы сакцентировать внимание на применении высокого спектрального разрешения, гиперспектральным анализом принято называть анализ спектральный, в котором использовано от ста до тысячи спектральных поддиапазонов. Спектральный анализ с числом поддиапазонов свыше тысячи принято называть ультраспектральным.

Полученный в результате гиперспектрального зондирования большой объем данных позволяет анализировать пространственно-спектральную структуру геофизического поля геосистемы:  $U = f(x, y, \lambda)$ , где  $\lambda$  — длина волны,  $(x, y)$  — координаты точки поверхности,  $U$  — функция поля, выраженная как спектральная энергетическая яркость или коэффициент спектральной яркости [9]. Результатом гиперспектральной съемки, производимой с аэро- или космических аппаратов, является трехмерный массив данных  $B$ , получивший название гиперкуб

$$B \in \{B_{ijs}, s = \overline{1, N}; i = \overline{1, m}; j = \overline{1, n}, \quad (1)$$

где  $s$  — номер спектрального поддиапазона,  $\varepsilon(i, j)$  — координаты центра пиксела гиперспектрального снимка,  $B_{ijs}$  — энергетическая яркость пиксела в спектральном поддиапазоне  $s$ ,  $N$  — суммарное количество анализируемых спектральных поддиапазонов. Гиперспектральный снимок образован  $N$  равными по размерам изображениями, каждое из которых характеризует энергию излучения в разных спектральных поддиапазонах. Каждый гиперспектральный пиксел  $\varepsilon(i, j)$  характеризуется спектральной характеристикой  $\{B_s\}_{s=1}^N; i, j = \text{const}$ , которая определяет значение спектрального поля  $U = f(\lambda)$  с элементарной поверхности, равной площади проекции пиксела изображения на земную поверхность, и рассматривается как геометрическая точка поля. Энергетическая яркость в спектральных характеристиках первоначально задается как коды выходных сигналов сенсора и далее может быть пересчитана в спектральную плотность энергетической яркости или коэффициент спектральной яркости (далее спектральная яркость)  $k(s) = \rho(s) / \rho_0(s)$ , который равен отношению яркости данной поверхности в заданном направлении к яркости идеальной рассеивающей поверхности при одинаковом освещении. Каждый гиперспектральный пиксел  $\varepsilon$  представлен своей спектральной характеристикой (спектрограммой)

$$k_s = k(\lambda_s), s = \overline{1, N}, \quad (2)$$

где  $\lambda_s$  — спектральные поддиапазоны (каналы сенсора), выраженные в длинах электромагнитных волн;  $N$  — количество спектральных каналов, используемых для анализа;  $k_s$  — спектральная яркость канала.

**Модель состояния геосистемы на основе спектральных данных.** В статистике переменными являются вероятности, а состояние системы определяется распределением вероятностей. Статистический анализ спектральной характеристики точки поля (пиксела) предусматривает нормирование спектральных яркостей на единицу, что позволяет ввести величины  $p_s$ , которые можно интерпретировать как вероятности. Эти вероятности равны относительным значениям спектральных яркостей в  $N$  спектральных поддиапазонах

$$p_s = \frac{k_s}{K}, \sum_{s=1}^N k_s = K, p_s > 0, \sum_{s=1}^N p_s = 1; \quad (3)$$

$$P = \{p_s, s = \overline{1, N}\}, \quad (4)$$

где  $K$  — суммарная спектральная яркость. Величины  $P$ , заданные выражением (4), можно рассматривать как распределение вероятностей по спектральным каналам (поддиапазнам), которое характеризует состояние системы в точке поля с координатами  $(x, y)$ .

**Энтропия Гиббса–Шеннона для распределений спектральных данных.** Обратной величиной к вероятностной мере  $p_s$  является величина  $(-\ln p_s)$  [2], которая для рассматриваемой задачи трактуется как мера неопределенности спектрального поддиапазона, или как информация, полученная от отдельного спектрального канала. Для редуцирования описания состояния системы, заданного распределением (4), вводится энтропия данного распределения. Информационная энтропия Гиббса–Шеннона [3, 4] для распределения гиперспектральных дан-

ных (4) определяется как математическое ожидание неопределенности состояния системы, которая равна моменту первого порядка неопределенности состояния поддиапазонов, т.е. среднему значению неопределенностей

$$H = M[-\ln P] = -\sum_{s=1}^N p_s \ln p_s. \quad (5)$$

Энтропия (5) рассматривается так же, как полная средняя информация, получаемая от всех  $N$  задействованных спектральных поддиапазонов (каналов прибора). При равномерном распределении вероятностей, когда все  $p_s$  равны  $p_s = 1/N$ , энтропия  $H$  максимальна. Система в данной точке поля максимально неупорядочена — этот случай соответствует самой низкой степени организованности системы.

**Метод 1. Определение изменений в состоянии систем на основе энтропии Гиббса–Шеннона для распределений спектральных данных, метод расхождения Кульбака.** Предположим, что спектральные измерения в одной точке поля привели к распределению

$$G = \{g_s, s = \overline{1, N}\}, \quad (6)$$

а в другой — к распределению  $P = \{p_s, s = \overline{1, N}\}$ , тогда расхождение в распределениях может быть вычислено как несимметричная информационная мера  $\mu_1$  расхождения Кульбака [2, 10], равная среднему значению изменения неопределенности системы

$$\Delta_s = -\ln g_s - (-\ln p_s), \quad (7)$$

$$\mu_1(P \| G) = \sum_{i=1}^N p_s \Delta_s = \sum_{i=1}^N \left[ p_s \ln \frac{p_s}{g_s} \right]. \quad (8)$$

Мера  $\mu_1(P \| G)$  интерпретируется как удаленность распределения  $G$  от распределения  $P$  (см. пример 1). Мера несимметричная,  $\mu_1(P \| G) \neq \mu_1(G \| P)$ , всегда больше нуля, кроме случая равных распределений  $\mu_1(P \| P) = 0$ .

**Энтропия Реньи для распределений спектральных данных.** Для анализа систем, спектральные данные которых распределены в соответствии со степенным законом и характеризуются как распределения с «тяжелыми хвостами», целесообразно использовать энтропию Реньи [11] в следующей модификации

$$H_\alpha = \frac{1}{1-\alpha} \ln \sum_{s=1}^N p_s^\alpha, \quad \alpha \geq 0, \quad \sum_{s=1}^N p_s = 1. \quad (9)$$

При значении параметра  $\alpha = 1$ , в силу условия нормирования вероятностей  $p_s$ , согласно выражению (3), логарифм суммы в формуле (9) равен 0. В этом случае возникает неопределенность в записи (9), раскрытие которой, при условии  $\alpha \rightarrow 1$ , приводит к формуле информационной энтропии Гиббса–Шеннона (5). Таким образом, с помощью выражения (9) вводится семейство функций, характеризующих состояние системы (для каждого значения  $\alpha$ ), которое включает в себя энтропию Гиббса–Шеннона (5).

**Метод 2. Определение изменений в состоянии систем на основе энтропии Реньи, метод  $\alpha$ -дивергенции спектральных данных.** Вычисление изменений в состоянии системы выполняется по спектральным распределениям

методом расхождения Реньи [11] (дивергенция Реньи,  $\alpha$ -дивергенция). Расхождение распределения  $P$  от распределения  $G$  порядка  $\alpha$  определяется выражением

$$\mu_\alpha(P\|G) = \frac{1}{\alpha-1} \ln \left( \sum_{s=1}^N \frac{p_s^\alpha}{g_s^{\alpha-1}} \right) = \frac{1}{\alpha-1} \ln \left( \sum_{s=1}^N p_s^\alpha g_s^{1-\alpha} \right), \quad \alpha > 0, \alpha \neq 1, \quad (10)$$

где  $P$  и  $G$  — распределения, заданные соотношениями (4) и (6) (см. пример 2). Меры  $\alpha$ -дивергенции симметричны для одинаковых значений  $\alpha$ ,  $\mu_\alpha(P\|G) = \mu_\alpha(G\|P)$ . Меры больше нуля и равны нулю только для равных распределений  $\mu_\alpha(P\|P) = 0$ . При  $\alpha \rightarrow 1$  дивергенция Реньи вырождается в расхождение Кульбака (8).

**Метод 3. Определение изменений в состоянии систем мультифрактальной структуры на основе обобщенной энтропии Реньи, метод варибельности спектра размерностей Реньи.** Этот энтропийный метод положен в основу мультифрактального анализа [7] спектральных характеристик геосистем [8, 9]. Метод предусматривает обобщение функционала энтропии (9) на отрицательные значения степенных показателей  $q$ :

$$H_q = \frac{1}{1-q} \ln \sum_{s=1}^{N(r)} p_s^q(r), \quad -\infty < q < \infty, \quad \sum_{s=1}^{N(r)} p_s(r) = 1, \quad (11)$$

где  $r$  — переменный размер спектральных поддиапазонов, выраженный в относительных единицах к общему спектральному диапазону, принятому для исследования геосистемы;  $N(r)$  — суммарное количество спектральных поддиапазонов, которое зависит от их размера  $r$ ;  $p_s(r)$  — вероятность, с которой встречаются спектральные яркости в спектральных поддиапазонах  $s$ .

Для каждого размера  $r$  спектральные яркости  $k_s(r)$  нормируются на единицу

$$p_s(r) = \frac{k_s(r)}{K}, \quad \sum_{s=1}^{N(r)} k_s(r) = K, \quad \sum_{s=1}^{N(r)} p_s(r) = 1. \quad (12)$$

Основной характеристикой мультифрактальной структуры исследуемой системы является спектр (множество) размерностей Реньи  $D_q$ , которые выражаются через энтропию (11)

$$D_q = - \lim_{r \rightarrow 0} \left[ \frac{H_q}{\ln r} \right]. \quad (13)$$

На практике величины размерностей (13) можно оценить, применяя следующую последовательность вычислений (см. пример 3). Выделим статистическую сумму в выражении (11) как функцию двух переменных

$$Z(q, r) = \sum_{s=1}^{N(r)} p_s^q(r). \quad (14)$$

Выражение (13) запишем в виде

$$D_q = \tau(q) / (1-q), \quad (15)$$

где функция  $\tau(q)$  определяется как предел

$$\tau(q) = - \lim_{r \rightarrow 0} \frac{\ln Z(q, r)}{\ln r}. \quad (16)$$

Оценим поведение статистической суммы при разных значениях  $r$  приближением к степенной зависимости. В билогарифмических координатах это семейство прямых

$$Z(q, r) = c(q)r^{-\tau(q)}, \quad \ln Z(q, r) = \ln c(q) - \tau(q) \ln r. \quad (17)$$

При фиксированном значении  $q$  уравнение (17) содержит пару неизвестных  $\ln c(q)$ ,  $\tau(q)$ , для нахождения которых достаточно вычислить значения  $(\ln Z(q, r), \ln r)$  для двух размеров  $r$ . Вычислив эти значения для большего количества размеров  $r$  спектральных поддиапазонов, можно оценить  $\tau(q)$  более точно. Для этого применяется метод наименьших квадратов или определяется  $\tau(q)$  как угол наклона прямой регрессии (17) для каждого значения  $q$ . По значениям функции  $\tau(q)$  и соотношению (15) вычисляются размерности Реньи  $D_q$  для дискретных значений показателя степени  $q$  в интервале  $q_{\min} \leq q \leq q_{\max}$ .

Введение меры, которая позволяет оценивать состояние системы мультифрактальной структуры, основано на перечисленных ниже свойствах функциональной зависимости размерностей  $D_q$ .

*Свойство 1* размерностей  $D_q$ . Функция размерностей Реньи  $D_q$ , вычисляемая по соотношениям (14)–(17), есть невозрастающая функция (числовая последовательность), значения которой монотонно убывают или остаются постоянными с увеличением показателя степени  $q$ ,  $D_q \geq D_{q'}$ ,  $q' > q$  (график этой функции приведен в примере 3). Неоднородному распределению спектральных яркостей  $k_s$  по поддиапазонам соответствует строгое неравенство  $D_q > D_{q'}$ .

*Свойство 2* размерностей  $D_q$ . Максимального значения функция  $D_q$  достигает при  $q \rightarrow -\infty$ ,  $\lim_{q \rightarrow -\infty} (D_q) = \max(D_q) = D_{-\infty}$ , а минимального значения — при  $q \rightarrow \infty$ ,  $\min(D_q) = D_{\infty}$ .

*Свойство 3* размерностей  $D_q$ . Знак равенства  $D_q = D_{q'} = D_{\text{support}}$ ,  $q \neq q'$  имеет место только при однородном распределении. В случае однородного распределения величины по поддиапазонам все размерности Реньи  $D_q$  равны между собой и равны одному значению — размерности Хаусдорфа носителя (support) мультифрактального множества [8]. В этом случае во всех спектральных поддиапазонах содержатся равные значения  $k_s(r)$  и, соответственно, все вероятности  $p_s(r)$  равны. Единица вероятности равномерно распределена по всем поддиапазонам. График функции размерностей Реньи для  $q = 0 : D_q = D_{\text{support}} = D_0$  вырождается в прямую линию, параллельную оси  $q$  (см. пример 3). Отметим, что энтропия Гиббса–Шеннона (5) такого распределения максимальна.

Рассмотренные свойства множества размерностей Реньи позволяют ввести меру  $\delta$  для оценки степени неоднородности спектральной характеристики гиперспектрального пиксела, получившую название «вариабельность спектра размерностей Реньи» [8]

$$\delta = (D_{-\infty} - D_{\infty}) \approx (D_{q_{\min}} - D_{q_{\max}}), \quad (18)$$

Мера  $\delta$  определяется по степени изменения функции размерностей  $D_q$  и позволяет различать состояние систем по спектрограммам или гиперспектральным снимкам.

## Численные эксперименты

**Описание эксперимента.** Демонстрация результатов применения энтропийных методов для определения границы залежи природного газа по постоянно гомогенного растительного покрова проводится на основе данных, полученных с Глебовского месторождения газа (Крым) [12]. Дистанционное спектрометрирование было выполнено полевым спектрорадиометром FieldSpec® 3 FR. В момент спектрометрирования на месторождении произрастала пшеница, которая находилась в стадии вегетации. Съемка велась дистанционно с расстояния 5 м. Маршрут исследования имел протяженность около 2 км и пересекал залежь газа с известными координатами границы залежи. Зона видимости местности прибором составляет 18 см по малой оси и 20 см по большой оси эллипса. На маршруте было выделено 16 контрольных точек (участков) спектрометрирования, в каждой из которых выполнено 4 измерения растительности со спектральным разрешением 1 нм в диапазоне длин электромагнитных волн 350÷2500 нм. В процессе обработки данных спектральный диапазон был ограничен интервалом 350÷749 нм. Последовательность нумераций спектрограмм соответствует направлению движения по маршруту. Всего поступило в обработку 64 спектрограммы. Цель вычислительных экспериментов состояла в автоматизированном определении компьютерными программами, составленными для каждого из трех энтропийных методов, границы залежи углеводородов по спектральным данным и верификации методов и компьютерных программ.

**Пример 1. Обработка спектральных данных методом расхождения Кульбака (метод 1).** На рис. 1, *a* приведен график спектрограммы № 1,  $\lambda$  — длины волн спектральных каналов в нанометрах,  $k$  — коэффициент спектральной яркости. Спектрограмма № 1 первой точки маршрута расположена вне залежи газа, и она принимается за базовую, с ней сравниваются все остальные спектрограммы маршрута № 2÷№ 64. Вычисление распределения спектральных яркостей выполняется по соотношениям (2) и (6) для спектрограммы № 1,  $G = \{g_s, s = \overline{1, N}\}$ , и для всех остальных спектрограмм — по соотношениям (3), (4),  $P = \{p_s, s = \overline{1, N}\}$ . На рис. 1, *б* приведен график мер расхождений Кульбака  $\mu_1$  как функция номера спектрограммы (обозначение  $x$ ), вычисленной по соотношению (8).

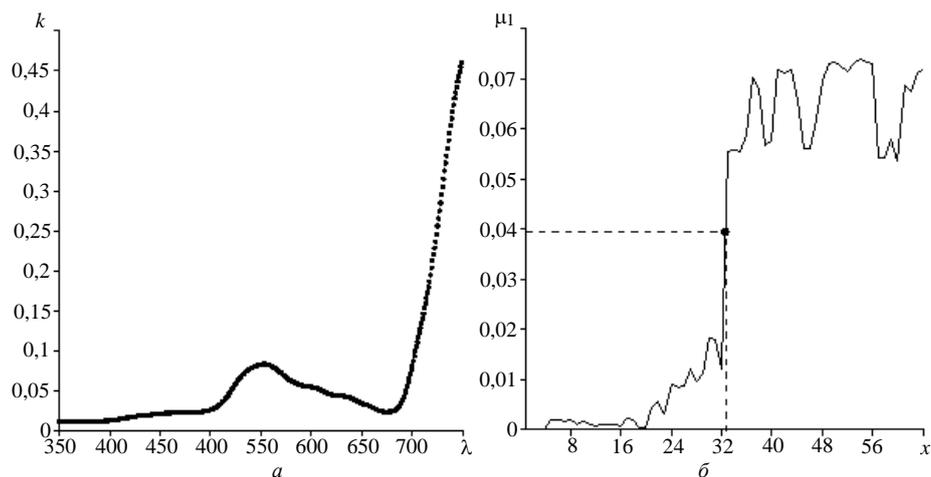


Рис. 1

Для определения границы залежи газа применяется пороговая бинарная сегментация значений меры  $\mu_1$  с помощью построения гистограммы. Численное значение

порога бимодального распределения меры равно  $\mu_1 = 0,039$ . Этому значению соответствует спектрограмма № 32 (см. пунктирные линии на рис. 1, б), которая получена с участка № 8. Граница залежи проходит по участкам № 8÷№ 9. Следовательно, участки № 1÷№ 8 (спектрограммы № 1÷№ 32) находятся вне залежи газа, остальные участки № 9÷№ 64 (спектрограммы № 33÷№ 64) располагаются над залежью.

**Пример 2. Обработка спектральных данных методом  $\alpha$ -дивергенции (метод 2).** Распределение спектральных яркостей в спектрограмме № 1 (обозначение  $G$ ), вычисленное по соотношению (6), сравнивается по формуле  $\alpha$ -дивергенции (10) с распределениями в спектрограммах № 2÷№ 64 (обозначение распределения для каждой спектрограммы  $P$ ). Численные результаты сравнения представлены мерами  $\mu_\alpha$  на графиках рис. 2, а. При каждом показателе степени  $\alpha$  значение функции  $\mu_\alpha$  может быть использовано как мера расхождения распределений  $G$  и  $P$ . Графики построены по вычисленным, согласно выражению (10), значениям  $\mu_\alpha$  для  $\alpha = 1, 1, 1; 1, 2; \dots; 10$ . На рис. 2, а приведены результаты сравнений спектрограммы № 1 со спектрограммами №  $x = \{5, 17, 21, 25, 29, 33, 37, 41, 47, 57, 61\}$ , нумерация спектрограмм на графиках снизу–вверх. На рис. 2, б приведен график  $\alpha$ -дивергенций  $\mu_{\alpha=10}$  как функция номера спектрограммы (обозначение  $x$ ), вычисленной по соотношению (10) при значении степенного показателя  $\alpha = 10$ .

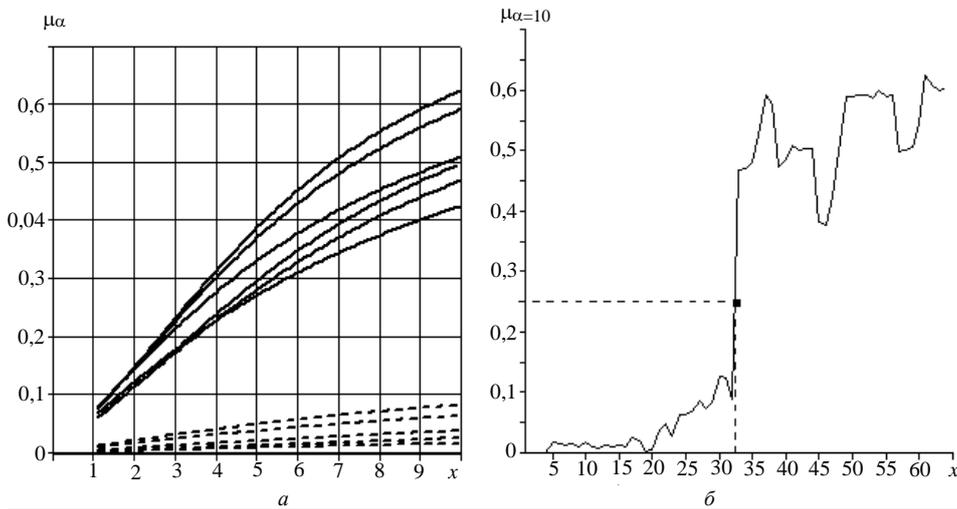


Рис. 2

Для определения границы залежи применена пороговая бинарная сегментация значений меры  $\mu_{\alpha=10}$  с помощью построения гистограммы. Численное значение порога бимодального распределения меры равно  $\mu_{\alpha=10} = 0,25$ . Этому значению соответствует спектрограмма № 32 (см. пунктирные линии на рис. 2, б), которая получена с участка № 8. Граница залежи проходит по участку № 8. Следовательно, как и в примере 1, участки № 1÷№ 8 (спектрограммы № 1÷№ 32) находятся вне залежи газа, остальные участки № 9÷№ 64 (спектрограммы № 33÷№ 64) располагаются над залежью.

**Пример 3. Обработка данных методом варибельности спектра размерностей Ренья (метод 3).** На рис. 3, а показаны графики фрагментов степенной зависимости (17) статистической суммы от размеров спектральных поддиапазонов  $r$  в билогарифмической шкале, выполненные для первой спектрограммы растительного покрова на маршруте исследования. На рис. 3, б приведены графики

размерностей Реньи  $D_q$ , полученные с шести контрольных участков маршрута. На графиках демонстрируются монотонно убывающие функции  $D_q$  для степенных показателей в интервале значений  $q: -90 \leq q \leq 90$ . Графики на рис. 3 наглядно подтверждают наличие масштабно-инвариантной мультифрактальной структуры в гиперспектральных отражениях растительного покрова (см. свойства размерностей  $D_q$ ).

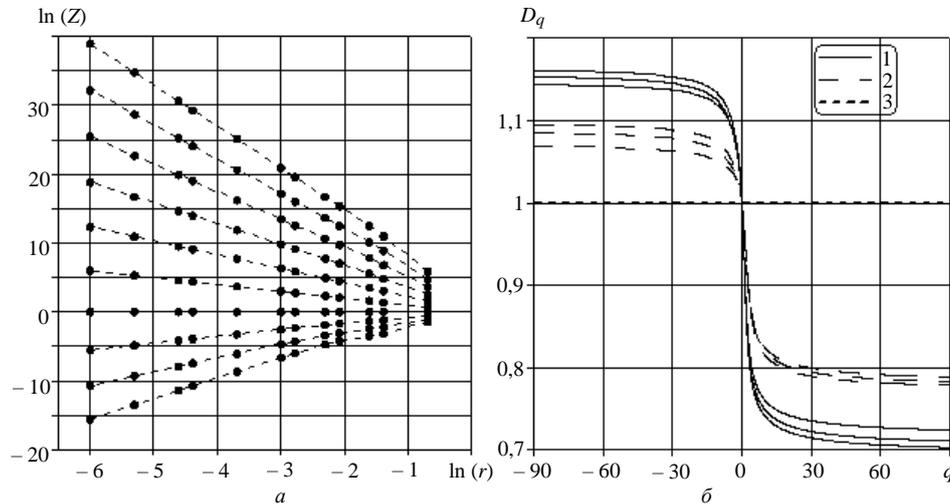


Рис. 3

Графики на рис. 3, *a* построены для значений  $q = \{-5, -4, \dots, 0, 1, \dots, 5\}$ , нумерация графиков сверху–вниз. В примере 3 каждый график построен по вычисленным 12 точкам для значений  $r = \{1/400, 1/200, 1/100, 1/80, 1/40, 1/20, 1/16, 1/10, 1/8, 1/5, 1/4, 1/2\}$ , через которые с помощью метода наименьших квадратов проводились прямые линии. По вычисленным углам наклона прямых (17) определены значения  $\tau(q)$ , и согласно выражению (15), вычислены функции  $D_q$  для каждой спектрограммы на маршруте. На рис. 3, *б* приведены графики функций  $D_q$ , для которых приняты следующие обозначения: 1 — для трех спектрограмм с участков, находящихся вне залежи газа; 2 — для трех спектрограмм с участков, расположенных над залежью; 3 — линия уровня однородного распределения  $D_{\text{support}} = 1$  (см. свойство 3 размерностей  $D_q$ ).

На рис. 4, *a* приведен график вариабельности размерностей Реньи (мера  $\delta$ ) как функция номера спектрограммы (обозначение  $x$ ). Мера  $\delta$  вычислена по соотношению (18) при значениях показателей  $q_{\min} = -90, q_{\max} = 90$ . На графике рис. 4, *a* виден скачок численного значения меры на 32 спектрограмме. Так же как и в двух предыдущих примерах, участки № 1–№ 8 (спектрограммы № 1–№ 32) находятся вне залежи газа, остальные участки № 9–№ 64 (спектрограммы № 33–№ 64) располагаются над залежью. Однако теперь большим значениям меры соответствуют спектрограммы, полученные с участков, расположенных вне залежи газа.

Если результат обработки данных готовится для визуального восприятия, то сегментация меры на маршруте может быть выполнена графически, как это показано на рис. 4, *б*. Для этого яркость дисплея в оттенках серого задается пропорционально значению меры  $\delta$ . На этом рисунке  $i$  — номер участка,  $j$  — номер спектрограммы на участке. Начиная с номера участка  $i = 9$ , располагается залежь газа. Используя метод вариабельности спектра размерностей Реньи, можно ввести меру отклонения системы от состояния равновесия.

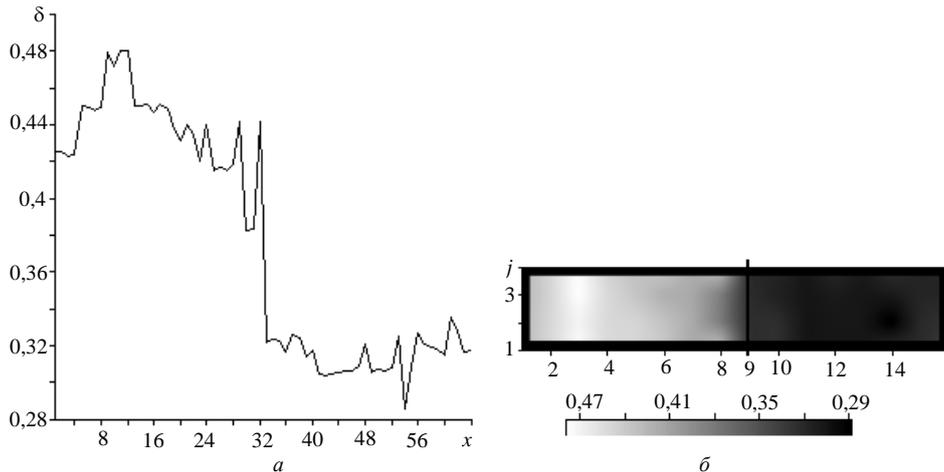


Рис. 4

### Заключение

В статье рассмотрены три метода моделирования изменений в состояниях спектральных структур сложных природных систем, обладающих свойством самоорганизации. Методы основаны на различных математических формах информационных энтропий, однако в численных экспериментах с реальными данными они привели к одинаковым результатам. Все три метода продемонстрировали правильное определение границы залежи природного газа на основе спектральных данных растительного покрова. Дополнительными исследованиями установлено, что частотное распределение [13] спектральных яркостей растительного покрова в диапазоне 350÷749 нм не отвечало степенному виду, столь характерному для критических состояний систем [5]. Самоорганизация оптической системы растительного покрова над залежью природного газа происходит в не критическом режиме, что позволяет применять для определения границ залежей любой из трех рассмотренных энтропийных методов. Обобщая полученные результаты с помощью модели систем самоорганизованной критичности (Self-Organized criticality), заметим, что более универсальными методами определения состояния систем следует признать методы  $\alpha$ -дивергенции и вариабельности спектра размерностей Реньи, которые позволяют анализировать различные типы самоорганизации.

*М.В. Артюшенко, А.В. Хижняк*

### ЭНТРОПІЙНІ МЕТОДИ МОДЕЛЮВАННЯ СИСТЕМ, ЩО САМООРГАНІЗУЮТЬСЯ ЗА ДАНИМИ ГІПЕРСПЕКТРАЛЬНОГО ЗОНДУВАННЯ

Для дослідження складних систем, що самоорганізуються, створюються різні математичні моделі. У геосистемах детермінований характер процесів обумовлений їх стохастичними властивостями. У розглянутих системах регулярні детерміновані процеси формуються численними випадковими міжелементними взаємодіями, які відбуваються на мікрорівні. У багатьох випадках коректно вивести детермінований закон еволюції системи, що досліджується, або її частини неможливо в зв'язку з великою кількістю непередбачуваних і невідомих чинників, що на неї впливають. Однак на мікрорівні доступними для спостереження є статистичні розподіли елементів системи, що дозволяє прогнозувати її поведінку і оцінювати фактори, які діють на систему. Найбільш універсальні методи моделювання систем зі стохастичними властивостями засновано на фундаментальних поняттях статистичної механіки — інформаційних ентропіях Гіббса–Шеннона і Реньї. У статті досліджуються ентропійні методи обчислень кількісних оцінок стану просторово-розподілених геосистем і їх

дивергенції в процесі самоорганізації:  $\alpha$ -розбіжностей, розбіжностей Кульбака, варіабельності спектра розмірностей Реньї. Розглянуто особливості дослідження систем з мультифрактальними структурами за даними гіперспектральних вимірювань. На прикладах ілюструється застосування ентропійних моделей в чисельних експериментах з реальними даними, отриманими з родовища природного газу. Ґрунтуючись на даних гіперспектральних випромінювань гомогенного рослинного покриву, проводиться верифікація застосування ентропійних методів для визначення меж покладів вуглеводнів.

**Ключові слова:** інформаційні ентропії, статистичний аналіз систем, самоорганізація, спектральні дані, чисельне моделювання, пошук покладів вуглеводнів.

*M.V. Artiushenko, A.V. Khyzhniak*

## ENTROPY METHODS OF SELF-ORGANIZING SYSTEMS MODELING USING HYPERSPECTRAL REMOTE SENSING DATA

Various mathematical models are created for exploring complex self-organizing systems. In geosystems, the deterministic nature of processes is due to their stochastic properties. In such systems, regular deterministic processes are formed by numerous random interelement interactions that occur at the micro level. In many cases, it is not possible to define correctly the deterministic law of evolution of the observable system or its part because of the large number of unpredictable and unknown factors that influence it. However, at the micro level, statistical distributions of system elements are available for observation, which allows to predict its behavior and evaluate the factors influencing the system. The most universal methods for modeling systems with stochastic properties are based on the fundamental concepts of statistical mechanics — the Gibbs-Shannon and Renyi information entropies. The article studies the entropy methods for calculating quantitative estimations of the states of spatially distributed geosystems and their divergences in the process of self-organization: alpha-divergence, Kullback divergence, variability of the spectrum of Renyi dimensions. The features of the above systems with multifractal structures according to hyperspectral measurements are observed. The examples illustrate the use of entropy models in numerical experiments with real data obtained from a natural gas field. Verification of the entropy methods for determining the boundaries of hydrocarbon deposits based on the hyperspectral data of emission of a homogeneous vegetation cover was carried out.

**Keywords:** information entropies, statistical analysis of system, self-organizing, spectral data, numerical simulation, search for hydrocarbons.

1. Вильсон А.Дж. Энтропийные методы моделирования сложных систем. Теория и методы системного анализа. М. : Физматлит, 1978. 248 с.
2. Хакен Г. Информация и самоорганизация: Макроскопический подход к сложным системам. М. : Мир, 1991. 240 с.
3. Башкиров А.Г. Энтропия Реньи как статистическая энтропия для сложных систем. *Теоретическая и матем. физика*. 2006. **149**, № 2. С. 299–317.
4. Башкиров А.Г., Витязев А.В. Энтропия Реньи и степенные распределения в естественных и гуманитарных науках. *Доклады академии наук РФ*. 2007. **412**, № 4. С. 476–479.
5. Вак Р. How nature works: the science of self-organized criticality. New York: Copernicus, 1996. 207 p.
6. Newman M.E.J. Power laws, Pareto distributions and Zipf's law. *Journal Contemporary Physics*. 2005. **46**. P. 323–351.
7. Федер Е. Фракталы. М. : Мир, 1991. 254 с.
8. Artiushenko M.V. Statistical analysis of the unsmooth geophysical fields by remote sensing data. *Journal of Automation and Information Sciences*. 2018. **50**, N 6. P. 14–27.
9. Artiushenko M.V. Identification and interpretation of power-law distributions by spectral data of remote sensing. *Journal of Automation and Information Sciences*. 2018. **50**, N 12. P. 17–33.
10. Кульбак С. Теория информации и статистика. М. : Наука, 1967. 408 с.
11. Renyi A. On measures of entropy and information. *Proc. Fourth Berkeley Symp. on Math. Stat. and Prob.* Univ. of Calif. Press. 1961. **1**. P. 547–561.
12. Лялько В.І., Шпортюк З.М., Сибірцева О.М., Дугін С.С., Воробйов А.І. Дослідження варіацій індексів червоного краю спектрів відбиття пшениці над газовим родовищем. *Космічна наука і технологія*. 2010. **16**. № 6. С. 5–10.
13. Кендалл М.Дж., Стьюарт А. Теория распределений. М. : Наука, 1966. 588 с.

Получено 14.05.2020

ISSN 0572-2691