

УДК 528.931:(553.98:519)

Валідація результатів картування контурів покладів вуглеводнів за матеріалами наземного спектрометрування, дистанційними та геолого-геофізичними даними

О. В. Титаренко*

ДУ “Науковий центр аерокосмічних досліджень Землі ІГН НАН України”, Київ, Україна

Оцінка вірогідності нафтогазоперспективних рішень вкрай необхідна при прийнятті рішень про промислову експлуатацію нафтогазоперспективних об'єктів. Вимога кількісної оцінки ефективності автоматизованої технології картування контуру покладу вуглеводнів за матеріалами наземного спектрометрування, дистанційними і геолого-геофізичними даними потребує спеціальної процедури валідації. Метою процедури валідації є встановлення відповідності досягнутих показників точності і достовірності заданим (прийнятим) вимогам Вхідними даними для проведення валідації одержаного за матеріалами наземного спектрометрування контуру вуглеводнів є місцеположення його точок відносно контурів, виділених іншими незалежними методами — дистанційними та геолого-геофізичними. Оскільки наземне спектрометрування виконується вздовж просторових профілів, то величиною, що оцінюється, є геометрична відстань точок контурів, одержаних різними методами, вздовж профілю. У статті запропоновано алгоритм валідації результатів картування контурів покладів вуглеводнів за матеріалами наземного спектрометрування, дистанційними та геолого-геофізичними даними. Апробація алгоритму проведена за результатами картування контурів покладу вуглеводнів на прикладі Новотроїцького та Східнорогінцівського нафтового родовища (Україна). Для розрахунку валідації результатів уточнення контуру покладу вуглеводнів за даними наземного спектрометрування рослинності на території Східнорогінцівського нафтового родовища використовувалися результати вимірювань вздовж 14 просторових профілів, Новотроїцького — 28 профілів. Одержано середню імовірність помилки 0.28, що є задовільним для нафтогазоперспективних робіт на суходолі. Отримані результати свідчать про задовільну достовірність картування контурів покладів вуглеводнів за даними польового спектрометрування. Застосування результатів проміжної валідації безпосередньо в ході картування контурів покладів вуглеводнів дозволить об'єктивно оцінювати статистичну узгодженість одержуваних кількісних вимірів і здійснювати їх обґрунтований відбір перед одержанням прикінцевих інформаційних продуктів, що збільшить їхню достовірність.

Ключові слова: валідація, просторові дані, нафтогазоносність, дані дистанційного зондування, спектрометрування
© О. В. Титаренко. 2019

DOI: <https://doi.org/10.36023/ujrs.2019.21.151>

Використання технологій ДЗЗ для пошуків нафти і газу вимагає підтвердження геологічної інформацією, оскільки, в основному, поклади вуглеводнів розташовані на значній глибині від поверхні. Крім того, оцінка вірогідності нафтогазоперспективних рішень вкрай необхідна при прийнятті рішень про промислову експлуатацію нафтогазоперспективних об'єктів.

Вимога кількісної оцінки ефективності автоматизованої технології картування контуру покладу вуглеводнів за матеріалами наземного спектрометрування, дистанційними і геолого-геофізичними даними потребує спеціальної процедури валідації. Метою процедури валідації є встановлення відповідності досягнутих показників точності і достовірності заданим (прийнятим) вимогам [7].

Основним видом помилок при вирішенні задачі картування контуру покладу вуглеводнів (ВВ) є по-

зиційний. Позиційні помилки відображають, наскільки точно вдається встановити відповідність між координатами точок контуру за різними даними (дистанційними, геологічними, геофізичними). Оцінки позиційної точності мають задовольняти вимогам статистичної вірогідності і бути відповідними в будь-якій точці контуру.

В загальному випадку дійсні значення позиції об'єкта отримують на основі інформації про координати наземних контрольних точок [13]. Оскільки при проведенні геологорозвідувальних робіт надійні контрольні точки, як правило, відсутні, залишається оцінювати статистичні показники точності позиціонування за різними незалежними методами.

Коректність валідації забезпечується польовими вимірюваннями на спеціальних тестових ділянках території, які найбільш повно відображають умови застосування технології, що перевіряється, та відповідають всім її вимогам. Крім того, повинна бути передбачена можливість одержання, при не-

* E-mail: <olgatitarenko66@ukr.net>

обхідності, додаткової калібрувально-завіркової інформації.

Висновки та результати валідації мають базуватися переважно на даних кількісних вимірів і розрахунків. Разом з тим допускаються, при необхідності, експертні оцінки. Кількість та розбіжність кількісних даних вимірів мають задовольняти вимогам статистичної репрезентативності. Розрахунок обсягу репрезентативної вибірки полягає в тому, що потужність вихідної множини вимірів є досить великою [10]. Проте на практиці це припущення не завжди виконується.

Процедуру валідації призначено, у першу чергу, для здійснення підсумкової перевірки ефективності запропонованої технології. Разом з тим, її застосування може виявитися корисним для оцінки попередніх результатів картування контуру покладу ВВ безпосередньо в ході нафтогазопошукових робіт.

Отже, вхідними даними для проведення валідації одержаного за матеріалами наземного спектрометрування контуру ВВ є місцеположення його точок відносно контурів, виділених іншими незалежними методами — дистанційними та геолого-геофізичними. Оскільки наземне спектрометрування виконується вздовж просторових профілів, то величиною, що оцінюється, є геометрична відстань точок контурів, одержаних різними методами, вздовж профілю.

Розглянемо просторовий профіль напряму x , вздовж якого n різними незалежними методами визначено точки перетину контуру покладу x_p , $i = 1 \dots n$. Припустимо, що в якості x_n використано вимірювання місцеположення контуру методом, який валідується, у даному випадку — наземним спектрометруванням, а $x_1 \dots x_{n-1}$ — опорні незалежні вимірювання. Оцінка імовірності помилки місцеположення ε визначається густиною розподілу параметра x , що спостерігається, та апіорними ймовірностями відповідних статистичних гіпотез. Оптимальне рішення визначається величиною відношення правдоподібності — точки рівності густин імовірності правильного рішення до помилки, а імовірність помилки — інтегралами “хвостів” відповідних густин імовірності [4].

Але на практиці ці густини розподілів, як правило, невідомі, їх оцінюють за навчальною вибіркою та потім інтегрують числовими методами [2]. При отриманні ж аналітичних рішень для багатьох видів розподілів виникають значні, а інколи й нездоланні математичні труднощі, тому часто вважають імовірнісні розподіли образів нормальними. В цьому випадку імовірність помилки класифікатора максимальної правдоподібності складе [8].

$$\varepsilon = P_1 \cdot \Phi\left(\frac{\ln x^* - 0.5d}{\sqrt{d}}\right) + P_2 \Phi\left[1 - \left(\frac{\ln x^* + 0.5d}{\sqrt{d}}\right)\right], \quad (1)$$

де x^* — поріг прийняття рішення класифікатора максимальної правдоподібності,

$$d = x^T \text{cov}^{-1}(x) x, \quad (2)$$

де відстань Махаланобіса між нормальними розподілами x , $\text{cov}(x)$ — коваріаційна матриця x , для одновимірних вимірювань, якими є місцеположення контурів $\text{cov}(x) = \sigma_x^2$, і відстань Махаланобіса $d = \frac{\Delta x^2}{\sigma_x^2}$.

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{\xi^2}{2}} d\xi - \quad (3)$$

табульований інтеграл імовірностей [6].

У випадку рівних апіорних імовірностей $P_1 = P_2$ співвідношення для імовірності помилки (1) спрощується до вигляду [5]

$$\varepsilon = 2\Phi\left(\frac{\sqrt{d}}{2}\right) - 1. \quad (4)$$

Основна складність застосування (4) для валідації розташування контурів покладів вуглеводнів полягає в малому розмірі статистичних вибірок: типові значення $n = 4 \dots 5$. В цьому випадку нормальність вхідного розподілу обов'язково потребує перевірки [9].

Для перевірки нормальності розподілу малих вибірок найбільше підходить критерій Колмогорова-Смирнова з поправкою Ліл'єфорса z_{\max} [12]:

$$Z_{\max} = \max_i \left[\frac{i}{n} - \Phi\left(\frac{x_i - \bar{x}}{\sigma_x}\right), \Phi\left(\frac{x_i - \bar{x}}{\sigma_x}\right) - \frac{i-1}{n} \right]. \quad (5)$$

Близькість величини (5) до порогового значення z_{\max}^* дозволяє визнати чи не визнати емпіричний імовірнісний розподіл близьким до нормального з наперед заданим рівнем значущості α . Нажаль, порогові значення z_{\max}^* критерію (5) не виражаються аналітично, тому зазвичай рішення приймаються за таблицями, одержаними методом Монте-Карло [9].

Проте, для оцінки рівня значущості α нормальності емпіричного імовірнісного розподілу може застосовуватися аналітична апроксимація розподілу Ліл'єфорса [11]:

$$a = \exp(-7.01256 \cdot z_{\max}^2 \cdot (n+2.78019) + 2.99587 \cdot z_{\max} \cdot (n+2.78019)1/2 - 0.122119 + 0.974598/n^{1/2} + 1.67997/n). \quad (6)$$

Отже, складається наступний алгоритм валідації результатів картування контуру покладу вуглеводнів за матеріалами наземного спектрометрування:

Крок 1. Введення даних про місцезнаходження контуру x за n різними незалежними дистанційними та геолого-геофізичними методами, причому x_n — місцезнаходження контуру за матеріалами наземного спектрометрування.

Крок 2. Розрахунок вибіркового середнього x та дисперсії σ_x^2 імовірнісного розподілу x : $x =$

$$\frac{1}{n} \sum_{i=1}^n x_i, \sigma_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Крок 3. Розрахунок критеріального значення зпах нормальності імовірнісного розподілу x за формулою (5).

Крок 4. Оцінка рівня значущості α нормальності емпіричного імовірнісного розподілу x за формулою (6). Якщо рівень значущості незадовільний, то поточне вимірювання відкидається, інакше — перехід до кроку 5.

Крок 5. Розрахунок відстані Махаланобіса d за формулою (2), причому $\Delta x = x_{n-1} - x_n$, де $x_{n-1} = \frac{1}{n-1} \sum_{i=1}^{n-1} x_i$.

Крок 6. Розрахунок імовірності помилки ϵ за формулою (4).

Алгоритм виконується для всіх наявних профілів, за якими визначалися точки контуру покладу вуглеводнів, одержана імовірність помилки усереднюється.

Для розрахунку валідації результатів уточнення контуру покладу вуглеводнів за даними наземного спектрометрування рослинності на території Східнорогінцівського нафтового родовища використовувалися результати вимірювань вздовж 14 просторових профілів, які містяться в табл. 1. Для кожного профілю за допомогою викладеного алгоритму оцінювалися рівень значущості нормальності одержаного статистичного розподілу та імовірність помилки визначення місцеположення точки контуру за даними назем-

ного спектрометрування (середнє значення імовірності помилки 0.215). Ці результати також містяться в табл. 1.

Для розрахунку валідації результатів методики уточнення контуру покладу вуглеводнів за даними наземного спектрометрування рослинності на території Новотроїцького нафтового родовища використовувалися результати вимірювань вздовж 28 просторових профілів, які містяться в табл. 2. Для кожного профілю за допомогою викладеного алгоритму за вибіркою $n = 5$ оцінювалися рівень значущості α нормальності одержаного статистичного розподілу та імовірність помилки ϵ визначення місцеположення точки контуру за даними наземного спектрометрування (середнє значення імовірності помилки 0.282). Ці результати також містяться в табл. 2.

Отримані результати свідчать про задовільну достовірність картування контурів покладів вуглеводнів за даними польового спектрометрування. Головною проблемою при цьому залишається гостра нестача достовірних даних про реальне розташування контурів нафтогазоносності на територіях дослідження. Перспективним напрямком подальших досліджень може стати залучення до валідації даних просторового моделювання статистичних розподілів параметрів нафтогазоносних геосистем [1]. Застосування результатів проміжної валідації безпосередньо в ході картування контурів покладів вуглеводнів дозволить об'єктивно оцінювати статистичну узгодженість одержуваних кількісних вимірів і здійснювати їх обґрунтований відбір перед одержанням прикінцевих інформаційних продуктів, що збільшить їхню достовірність.

Таблиця 1.

Результати вимірювань місцеположення контуру покладу вуглеводнів на території Східнорогінцівського родовища

Номер профілю (дата) тип рослинності	ВНК за структурною картою горизонту В-19	ВНК за структурною картою горизонту В-17	ВНК за структурною картою горизонту В-26	Контур за даними дистанційного зондування	Контур за даними наземного спектрометрування	Рівень значущості нормальності імовірнісного розподілу	Імовірність помилки визначення місцеположення контуру
№1-4 (1995) береза	714	868	555	464	1085	1.662562	0.618832
№1-1 (1995) клен	53	13	34	217	105	0.512911	0.125378
№1-1 (1995) клен	886	763	971	732	816	1.598597	0.090763
№1-1 (1995) копитняк	886	763	971	732	816	1.598597	0.090763
№1-1 (1995) ліщина	53	13	34	217	360	0.136670	0.652909
№1-1 (1995) ліщина	886	763	971	732	816	1.598597	0.090763
№2-6 (1997) дуб	856	1004	690	623	818	1.695999	0.066309
№3-4 (1999) дуб	1217	1353	1125	1115	965	1.311227	0.594123
№3-4 (1999) дуб	1217	1353	1125	1115	1264	0.902595	0.242868
№3-5 (1999) дуб	1766	1648	1864	1623	1729	1.303349	0.015482
№3-6 (1999) дуб	2546	2489	2636	2452	2536	0.872550	0.030163
№3-6 (1999) клен	2546	2489	2636	2452	2589	1.731287	0.305678
№4-1 (2009) береза	757	703	884	644	741	0.513680	0.026993
№4-1 (2009) липа	757	703	884	644	735	0.566667	0.053881

Таблиця 2.

Результати вимірювань місцеположення контуру покладу вуглеводнів на території Новотроїцького нафтового родовища

Номер профілю (дата) тип рослинності	ВНК за структурною картою горизонту В-23	ВНК за структурною картою горизонту Т-1 за Ф. Д. Савченком	ВНК за структурною картою горизонту Т-1 за Б. М. Угаровим	Контур за даними дистанційного зондування	Контур за даними наземного спектрометрування	Рівень значущості імовірнісного розподілу	Імовірність помилки визначення місцеположення контуру
№ 1–1 (1995) сосна	1128	901	1061	681	643	0.826	0.507845
№ 2–1 (1995) сосна	1225	444	732	285	656	0.737	0.017293
№ 3–1 (1995) сосна	1029	476	638	496	257	0.878	0.518649
№ 3–2 (1995) сосна	1161	478	652	491	712	0.270	0.023721
№ 4–1 (1995) сосна	813	498	634	226	280	1.071	0.409106
№ 4–2 (1995) сосна	1309	752	962	699	1173	1.059	0.355174
№ 4–3 (1995) сосна	768	483	634	154	734	0.797	0.346444
№ 1–1 (1996) береза	1007	363	521	268	475	0.142	0.090137
№ 1–2 (1996) береза	1478	1161	1345	1339	662	0.331	0.704561
№ 3–1 (1996) береза	1457	830	1009	904	456	0.934	0.590216
№ 4–1 (1996) береза	1899	1387	1574	1290	1075	1.572	0.542769
№ 4–2 (1996) береза	820	501	649	162	250	1.206	0.395492
№ 5–1 (1996) береза	49	248	401	749	246	0.663	0.175332
№ 1–3 (1996) дуб	1177	419	590	425	274	0.245	0.407771
№ 1–4 (1996) дуб	2111	1485	1667	1541	1933	0.879	0.336367
№ 2–1 (1996) дуб	1817	1556	1345	1374	1279	0.291	0.425336
№ 3–2 (1996) дуб	1900	1381	1592	1306	1548	0.861	0.005628
№ 3–3 (1996) дуб	829	521	619	154	732	0.808	0.300847
№ 4–3 (1996) дуб	1702	1872	2043	680	302	0.228	0.586617
№ 4–4 (1996) дуб	1702	1872	2043	680	1642	0.059	0.050862
№ 1–5 (1996) сосна	1005	392	548	449	633	0.468	0.056889
№ 2–2 (1996) сосна	2053	1167	1343	1345	1680	0.230	0.226788
№ 3–4 (1996) сосна	1667	1164	1371	1072	1323	0.989	0.007856

Література

- Бурштейн Л. М. Статистические оценки параметров распределения скоплений нефти по величине в слабоизученных седиментационных бассейнах. *Геология и геофизика*. 2006. Т. 47(9) С. 1013–1023.
- Вентцель Е. С., Овчаров Л. А. Теория вероятностей и её инженерные приложения. М.: Высшая школа, 2000. 480 с.
- ДСТУ ISO 5479:2009. Видання. Статистичне опрацювання даних. Критерії відхилення від нормального розподілу. Вид. офіц. Київ: Держспоживстандарт України, 2009. 32 с.
- ДСТУ ISO 10576-2006. Видання. Статистичні методи. Настанови щодо оцінювання відповідності заданим вимогам. Вид. офіц. Київ: Держспоживстандарт України, 2006. 19 с.
- Орлов В. В. Эффективность распознавания образов при оценивании эталонов по классифицированным выборкам. *Труды Одесского политехнического университета*. Одесса, 2004. Вып.1(21). С. 121–124.
- Поллард Дж. Справочник по вычислительным методам статистики. Финансы и статистика. М.: 1982. 344 с.
- Попов М. О., Цимбал Т. В. Правила та процедура сертифікації методик використання даних дистанційного зондування Землі при вирішенні тематичних задач. *Космічна наука і технологія*. 2004. Т. 10. № 4. С. 104–112.
- Ту Дж., Гонсалес Р. Принципы распознавания образов. М.: Мир, 1978. 414 с.
- Abdi H., P. Molin Lilliefors test of normality. *Encyclopedia of Measurement and Statistics / Ed. by N.J. Salkind*. Thousand Oaks: Sage Publications, 2007. P. 540–544.
- Cochran W. G. *Sampling Techniques*. N. Y.: John Wiley, 1977. 428 p.
- Dallal G. E., L. Wilkinson. An analytic approximation to the distribution of Lilliefors's test statistic for normality. *The American Statistician*. 1986. Vol. 40. No. 4. P. 294–296.
- Lilliefors H. W. On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*. 1967. Vol. 62. No. 318. P. 399–402.
- McCoy R.M. *Fields Methods in Remote Sensing*. N.Y.: Guilford Press, 2005. 160 p.

References

1. Burshtein, L.M. (2006). Statistical estimation of parameters of size distribution of oil accumulations in poorly studied sedimentary basins. *Geologiya i geofizika*, 47(9), 1013–1023. (Russian).
2. Venttsel, Ye.S., Ovcharov, L.A. (2000). *Probability theory and its engineering applications*. Moscow: Vysshaya Shkola. (Russian).
3. DSTU ISO 5479:2009. Edition. Statistical data processing. Criteria for deviation from normal distribution, Kyiv: Derzhspozhyvstandart Ukrainy, 2009. (Ukrainian).
4. DSTU ISO 10576-2006. Edition. Statistical methods. Guidelines for compliance assessment with the requirements, Kyiv: Derzhspozhyvstandart Ukrainy, 2009. (Ukrainian).
5. Orlov, V. V. (2004). Pattern recognition efficiency in the assessment of standards on classified samples. *Trudy Odeskogo politekhnicheskogo universiteta*, iss 1(21), 121–124. (Russian).
6. Pollard, J. (1982). *Handbook of Numerical and Statistical Techniques*. Moscow: Finansy i statistika.
7. Popov, M. O., Tsybal, T. V. (2004). Rules and policies for methods certification for using remote sensing data from the Earth in order to support topic-specific objectives. *Kosmichna nauka i tekhnologija*, 10(4), 104–112. (Ukrainian).
8. Tu, J., Gonzalez, R. (1978). *Pattern recognition principles*. Moscow: Mir. (Russian).
9. Abdi, H., Molin, P. (2007). *Lilliefors test of normality. Encyclopedia of Measurement and Statistics*. N.J. Salkind (Ed.) (pp. 540–544). Thousand Oaks: Sage Publications.
10. Cochran, W. G. (1977). *Sampling Techniques*. N. Y.: John Wiley.
11. Dallal, G. E., Wilkinson, L. (1986). An analytic approximation to the distribution of Lilliefors's test statistic for normality. *The American Statistician*, 40(4), 294–296.
12. Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62(318), 399–402.
13. McCoy, R. M. (2005). *Fields Methods in Remote Sensing*. N.Y.: Guilford Press.

ВАЛИДАЦИЯ РЕЗУЛЬТАТОВ КАРТИРОВАНИЯ КОНТУРОВ ЗАЛЕЖЕЙ УГЛЕВОДОРОДОВ ПО МАТЕРИАЛАМ НАЗЕМНОГО СПЕКТРОМЕТРИРОВАНИЯ, ДИСТАНЦИОННЫМ И ГЕОЛОГО-ГЕОФИЗИЧЕСКИМ ДАННЫМ

О. В. Титаренко.

Оценка вероятности нефтегазоносности территорий крайне необходима при принятии решений о промышленной эксплуатации нефтегазоперспективных объектов. Количественная оценка эффективности технологии картирования контура залежи углеводородов по материалам наземного спектрометрирования, дистанционным и геолого-геофизическим данным требует специальной процедуры валидации. Целью ее является установление соответствия достигнутых показателей точности и достоверности заданным (принятым) требованиям. Исходными данными для проведения валидации полученного по материалам наземного спектрометрирования контура углеводородов является местоположение его точек относительно контуров, выделенных другими независимыми методами — дистанционными и геолого-геофизическими. Поскольку наземное спектрометрирование выполняется вдоль пространственных профилей, то величиной, которая оценивается, является геометрическое расстояние точек контуров, полученных разными методами, вдоль профиля. В статье предложен алгоритм валидации результатов картирования контуров залежей углеводородов по материалам наземного спектрометрирования, дистанционными и геолого-геофизическими данными. Апробация алгоритма проведена на примере Новотроицкого и Восточно-Рогинцевского месторождений (Украина). Для расчета валидации на территории Восточно-Рогинцевского месторождения использовались результаты измерений вдоль 14 пространственных профилей, Новотроицкого — 28 профилей. Получена средняя вероятность ошибки 0,28, что свидетельствует об удовлетворительной достоверности картирования контуров залежей углеводородов по данным полевого спектрометрирования. Применение результатов промежуточной валидации непосредственно в ходе картирования контуров залежей углеводородов позволит объективно оценивать статистическую согласованность получаемых количественных измерений. Кроме того, появляется возможность осуществлять их обоснованный отбор перед получением заключительных информационных продуктов, что увеличит их достоверность.

Ключевые слова: валидация, пространственные данные, нефтегазоносность, данные дистанционного зондирования, спектрометрирование

HYDROCARBON DEPOSIT MAPPING VALIDATION BY THE MEANS OF GROUND-BASED SPECTROMETRY, REMOTE SENSING AND GEOPHYSICAL DATA

O. V. Titarenko.

The probability estimation of oil and gas inside certain area is essential for decision making on the industrial exploitation of oil and gas bearing features. A quantitative assessment of the hydrocarbon contour mapping accuracy using ground-based spectrometric measurements, remote, geological and geophysical data requires a special validation procedure. Its purpose is to evaluate achieved accuracy and reliability as well as the conformance to specified requirements. The input data for validation of the hydrocarbon deposit contour by field spectrometry are the one points' locations relative to the other contours detected by independent methods, such as remote, geological and geophysical. As the field spectrometry performed along spatial trace,

the geometric drifts of other methods' cross-points are estimated. The algorithm for the validation of hydrocarbon deposit contour mapping by field spectrometry, remote, geological and geophysical data is proposed in this paper. The algorithm was tested on over the Novotroitsky and East Roginty hydrocarbon deposits (Ukraine). Measurements along 14 spatial traces over the Novotroitsky's deposit and 28 traces over the East Roginty's one was carried out to perform validation. The average error probability was 0.28, which demonstrates an admissible reliability of hydrocarbon deposits contours' mapping by field spectrometry data. The preliminary validation estimates engagement during the hydrocarbon deposits mapping provides the fact-based statistical consistency of the quantitative measurements received. In addition, it is possible to filter the outliers reasonable before final information product release, which will enhance the overall reliability.

Keywords: validation, spatial data, oil and gas potential, remote sensing data, spectrometry