

© Г.Б. Сергій, 2008

УДК 550. 84+519.2

*Український державний геологорозвідувальний інститут,
м. Київ*

ДЕЯКІ ПИТАННЯ СПРОЩЕНОГО ПІДХОДУ ДО ОЦІНКИ ЙМОВІРНОСТЕЙ ДОСТОВІРНОСТІ КОМПЛЕКСНОЇ ІНТЕРПРЕТАЦІЇ

Ускладнення сучасних геологічних задач все частіше висуває такі вимоги до геофізичних, геохімічних та інших прогнозуючих методів, які знаходяться на межі їх можливостей. Одним з очевидних і природних шляхів вирішення таких вимог є комплексування методів. Воно ускладнює і здорожчує геолого-геофізичні роботи, тому потребує обґрунтування не лише на інтуїтивному, а і на теоретичному рівні. Через це час від часу з'являються публікації, в яких на основі простіших посилань до теорії ймовірностей стверджується, що достовірність та ефективність комплексної інтерпретації завжди вище, ніж у складових методів, незалежно від їх індивідуальної точності та надійності [1–3]. У прикладах, що містять ці роботи, апріорні оцінки ймовірностей достовірності кожного методу зазвичай високі, бо беруться некритично у авторів та ентузіастів цих методів. За цими міркуваннями, кінцева достовірність прогнозу при комплексній інтерпретації має наближатися до одиниці при збільшенні кількості складових методів будь-якої достовірності.

На жаль, це розходиться з реальною практикою, тому потребує детальнішого аналізу. Важливість цього питання зростає внаслідок стрімкого поширення формалізованих автоматичних способів комплексування.

Обмежимося рамками того ж спрощеного підходу до оцінки ймовірностей [4]. В ньому кожен з методів характеризується ймовірністю правильного прогнозу p_1 , відповідно, хибного прогнозу $q_1=1-p_1$. Можлива різноманітність невірних рішень в рамках спрощеного підходу не враховується. Прогнози різних методів вважаються незалежними і некорельованими. Таким чином, правильному прогнозу одночасному за двома методами відповідає добуток ймовірностей $p_1 p_2$, за трьома $p_1 p_2 p_3$ і так далі. Зрозуміло, що як правильний, так і хибний прогноз може бути і позитивним, і негативним.

Відомо, хоча іноді забувається, що при комплексуванні змінюється ймовірність помилок двох видів – відмови від вірного розв'язку та прийняття невірного. Ці види помилок різноспрямовані, зменшення помилок одного виду супроводжується збільшенням другого. Співвідношення цих змін

визначається в кожному випадку ймовірностями помилок у методах, що комплексуються.

При обґрунтуванні доцільності використання найбільшої кількості методів будь-якої достовірності спираються, зазвичай, на правило додавання ймовірностей. Дійсно, включення до комплексу навіть дуже слабкого методу дещо збільшить ймовірність того, що хоча б один з методів дасть вірний прогноз. Але нерідко не враховується алгоритм, за яким цей прогноз вибирається і приймається. Найчастіше приймається той прогноз, який дається більшістю методів, тобто треба оцінювати ймовірність не просто появи вірного прогнозу, а появи його одночасно в більшості методів. При такому підході стає очевидним, що надійний метод може загубитися серед багатьох ненадійних.

Розглянемо кілька простих прикладів, щоб відслідкувати основні тенденції.

Спочатку розглянемо простішу ситуацію – комплексування двох незалежних методів. Ми оцінюємо в рамках спрощеного підходу лише збіг або незбіг прогнозу за обома методами, яким відповідає добуток ймовірностей $p_1 p_2$ та $q_1 q_2$. Змішані результати при $p_1 q_2$ та $q_1 p_2$ для комплексу з двох методів відносяться до групи невизначеності.

Розглянемо повну групу подій, сумарна ймовірність якої завжди дорівнює одиниці, тобто $p_1 p_2 + p_1 q_2 + q_1 p_2 + q_1 q_2$. Оцінимо ймовірність правильного результату комплексної інтерпретації $P_{1,2}$ та її очікуваного збільшення порівняно зі складовими методами при різних значеннях p_i та $q_i = 1 - p_i$. Для цього розглянемо табл. 1.

В першому випадку, коли ймовірності правильного прогнозу обох методів дорівнюють 0,5, достовірність результату, тобто ймовірність вибору вірного прогнозу із суми подій із обігу прогнозу за двома методами ($p_1 p_2 + q_1 q_2$), лишилася такою ж, як для окремого методу:

$$P_{1,2} = p_1 p_2 / (p_1 p_2 + q_1 q_2) = 0,25 / (0,25 + 0,25) = 0,50. \quad (1)$$

Ефективність інтерпретації знизилася через велику частку невизначених результатів (50%).

У другому випадку, який відповідає оптимістичному варіанту ($p_1 = 0,8$; $p_2 = 0,8$), невизначених результатів менше (32%), а достовірність результату

Таблиця 1

Події/ $p_1 p_2$	0,5	0,8	0,8	0,8	0,5
	0,5	0,8	0,5	0,2	0,2
$p_1 p_2$	0,25	0,64	0,40	0,16	0,10
$p_1 q_2$	0,25	0,16	0,40	0,64	0,40
$q_1 p_2$	0,25	0,16	0,10	0,04	0,10
$q_1 q_2$	0,25	0,04	0,10	0,16	0,40
	1,00	1,00	1,00	1,00	1,00
$P_{1,2}$	0,50	0,94	0,80	0,50	0,20

збільшується відносно окремих методів ($0,64 / (0,64 + 0,04) = 0,94$). В цьому випадку комплексна інтерпретація виправдана та ефективна, якщо не враховувати надто велику частку невизначених результатів.

В інших випадках маємо методи різної точності та надійності. При $p_1 = 0,8$ та $p_2 = 0,5$ достовірність результату порівняно з кращим методом не змінюється, $0,40 / (0,40 + 0,10) = 0,8$, а витрати зростають.

При $p_1 = 0,8$ та $p_2 = 0,2$ результат ще гірший – ймовірність правильного прогнозу стає $0,5$ замість $0,8$ для першого методу без комплексування його із слабшим другим методом. Зовсім погана ситуація при комплексуванні мало надійних методів ($p_1 = 0,5$ та $p_2 = 0,2$). Замість очікуваного зростання надійності ймовірність правильного прогнозу зменшилася з $0,5$ для кращого методу до $0,2 = 0,10 / (0,10 + 0,40)$.

Щоб позбавитися впливу невизначених результатів (другий та третій рядки табл. 1) та перевірити попередні результати, розглянемо аналогічний приклад з трьома методами (табл. 2).

В цьому прикладі достовірність правильного прогнозу визначається сумою ймовірностей перших чотирьох рядків $P_{1,2,3}$, в яких вірний прогноз дають принаймні два методи. Вона наведена в останньому рядку табл. 2.

Як і в попередньому прикладі, при $p_1 = p_2 = p_3 = 0,5$ ми не маємо ні виграшу, ні програшу. Можна припустити, що апріорна ймовірність $p = 0,5$ є граничною, при якій метод і не погіршує, і не покращує комплексний результат.

У другому випадку ($p_1 = p_2 = p_3 = 0,8$) справді досягається ефективність комплексної інтерпретації. Ймовірність вірного прогнозу зростає з $0,8$ до $0,896$. Ще більша достовірність прогнозу при збігу всіх методів ($0,512 / 0,520 = 0,98$).

Ситуація різко погіршується, якщо один з надійних методів замінити на слабкий ($p_1 = p_2 = 0,8; p_3 = 0,2$). Ймовірність вірного прогнозу для комп-

Таблиця 2

Події/ $p_1 p_2 p_3$	0,5	0,8	0,8	0,8
	0,5	0,8	0,8	0,2
	0,5	0,8	0,2	0,2
$p_1 p_2 p_3$	0,125	0,512	0,128	0,032
$p_1 p_2 q_3$	0,125	0,128	0,512	0,128
$p_1 q_2 p_3$	0,125	0,128	0,032	0,128
$q_1 p_2 p_3$	0,125	0,128	0,032	0,008
$p_1 q_2 q_3$	0,125	0,032	0,128	0,512
$q_1 p_2 q_3$	0,125	0,032	0,128	0,032
$q_1 q_2 p_3$	0,125	0,032	0,008	0,032
$q_1 q_2 q_3$	0,125	0,008	0,032	0,128
	1,000	1,000	1,000	1,000
$P_{1,2,3}$	0,500	0,896	0,704	0,296

лексу трьох методів отримуємо лише 0,704 замість 0,800. Ймовірність 0,800 зберігається тільки при збігу всіх методів.

Зовсім погано, коли з надійним методом комплексуються два значно слабших ($p_1 = 0,8; p_2 = p_3 = 0,2$). В цьому випадку замість ймовірності 0,8 для кращого методу отримуємо ймовірність правильного прогнозу комплексу трьох методів лише 0,296. При збігу всіх трьох методів вона дорівнює 0,200, тобто ще погіршується, а не збільшується.

У випадку комплексування трьох методів можна для оцінки очікуваної ймовірності вірного прогнозу користуватися формулою, яка легко отримується з лівої графи табл. 2 після множення та розкривання дужок:

$$P_{123} = p_1 p_2 + p_1 p_3 + p_2 p_3 - 2 p_1 p_2 p_3. \quad (2)$$

Подібну формулу можна отримати для будь-якої кількості методів.

Такі приклади можна наводити ще довго, але для основних висновків достатньо і цих, тим більше, що розглядається тільки спрощений підхід до оцінки достовірності результатів формалізованого прогнозування комплексом методів.

Очевидно, що для оцінок достовірності комплексної інтерпретації, навіть спрощених і приблизних, треба мати апіорні оцінки надійності для кожного методу. На жаль, треба констатувати, що таких оцінок, яким можна було б довіряти, ми практично не маємо. Справа не тільки в тому, що оцінки авторів або пропагандистів методу не завжди об'єктивні. Для теоретичних оцінок точності треба для кожної геологічної ситуації чи задачі провести великі дослідження, для яких ще треба розробляти теорію та методику. Достовірної статистики теж не вистачає, в кращому випадку є узагальнення якихось даних майже без їх аналізу. При цьому також неминучі проблеми. Наведемо один приклад.

Свого часу на Кадницькій площі ДДЗ та на Лудинській площі Волино-Поділля були проведені роботи комплексами геохімічних та геофізичних методів. На обох площах виникла однакова ситуація – різні методи дали прогнози, які не співпадали між собою, а свердловини були пробурені без врахування цих прогнозів, на “чистому місці”. Такий збіг свідчить про те, що подібні ситуації досить поширені. В обох випадках свердловини виявилися непродуктивними. Формально прогнози всіх методів підтвердилися, всім їм з повним правом треба ставити “плюс” для статистики підтвердження. Але ж прогнози різних методів не збіглися, тобто принаймні більшість з них на цей “плюс” не заслуговують. Як бути зі статистикою в подібних ситуаціях, неясно.

Комплексування методів на якісному рівні в ході геологічної інтерпретації, безумовно, є необхідним. Комплекс методів, що застосовуються, та

їхня вага змінюються залежно від геологічних задач та умов. Очевидно, вони будуть різними для задач пошуку і деталізації пасток вуглеводнів, оцінки нафтогазонасиченості, побудови геологічної моделі зонального та регіонального рівнів тощо.

При переході до кількісної комплексної інтерпретації необхідно оцінити очікувану достовірність результату.

Не маючи надійних оцінок апріорної достовірності методів, варто користуватися хоча б приблизними, навіть інтуїтивними оцінками для визначення набору методів, що комплексуються, і методики власне комплексування. Без таких оцінок до формалізованого комплексування треба ставитися дуже обережно, можливо, зовсім від нього утримуватися або розглядати як допоміжний засіб, перевіряючи результати геологічними міркуваннями.

Методи, які для певної задачі не мають високої оцінки достовірності, також не повинні ігноруватися. Їх треба використовувати з меншою вагою або лише “з правом дорадчого голосу”, але ніяк не виключати з розгляду.

Треба пам'ятати, що всі наведені вище міркування відповідають дуже спрощеному підходу до оцінки інформативності і достовірності окремих методів та їх комплексу. Справжня ситуація має бути значно складнішою.

Так, навряд чи можна з повним правом вважати різні методи незалежними. Хоча вони і незалежні за методикою, технологією та теоретичними засадами, але їх результати висвітлюють одне і те саме геологічне середовище. Можливо, при оцінці ймовірностей треба враховувати умову збігу певних аномалій над одним і тим же геологічним об'єктом, використовувати і параметри з однаковою просторовою частотою тощо.

Перспективним, хоча і досить складним, виглядає підхід, в якому визнається залежність методів, а достовірність результату визначається умовними ймовірностями. Наведемо два приклади обґрунтованості такого підходу.

В Англії було виконано аналіз результатів буріння в акваторії Атлантичного океану за 10 років [5]. До буріння там було проведено прогноз наявності покладів методом AVO. В середньому цей прогноз підтвердився на рівні 50%, але виявилось, що за відсутності структурних пасток підтвердження прогнозу нафтогазонасиченості майже нульове, але там, де були виявлені пастки, прогноз AVO підтвердився повністю. Очевидно, за подібних умов доцільно визначати умовну ймовірність прогнозу залежно від надійності виявлення пастки.

Другий приклад стосується методів ОНМ, яка призначена для прямого прогнозування нафтогазонасиченості на основі електромагнітних зондувань [6]. Її автори чітко вказують, що висока надійність методу (до 80%) досягається за умови правильності даних структурної сейсмозвідки, які використовуються при прогнозуванні методом ОНМ.

Такі підходи до оцінки ймовірностей достовірності комплексної інтерпретації, як і багато інших можливих підходів, вимагають подальших досліджень з урахуванням специфіки конкретних задач і умов. Єдиний узагальнений розв'язок може виявитися неможливим.

Головний висновок, який можна зробити в рамках використаного в цій роботі спрощеного підходу, – формалізоване комплексування методів може не тільки не покращити результат, а навпаки, може суттєво знизити його достовірність порівняно з використанням одного, найбільш надійного методу, якщо в комплексуванні задіяні навіть поодинокі методи із значно нижчою надійністю.

Не варто використовувати для формалізованого комплексування методи з різною апріорною надійністю, якщо не застосовувати відповідних вагових функцій. Не можна брати до такого комплексу методи, у яких ймовірність вірного прогнозу менша 0,5. Ці висновки повністю узгоджуються з відомим положенням, за яким надійність системи визначається її слабкою ланкою.

1. Бенько В.М., Бабаєв В.В. та ін. До питання про комплексування методів пошуку нафти і газу // Нафт. і газова пром-ть. – 2004. – №4. – С. 3–4.
2. Saunders D.F., Burson K.R. et al. Combined geological and surface geochemical methods discovered Agaritta and Brady Greek fields // AAPG Bull. – 1993. – №7. – P. 1219–1240.
3. Sorokin A.P., Slobodyanyuk S.O., Sirchenko V.V. Authenticity of hydrocarbons deposit prediction rise on a base the integrated model of field // Geosciences to discover and develop. Intern. Conf. – Saint Petersburg. – 2006. – A030.
4. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и ее инженерные приложения. – М.: Наука, 1988. – 480 с.
5. Loizon N. Post well analysis of exploration drilling on the UK Atlantic Margin provides clues to success // First break. – 2003. – №4. – P. 45–50.
6. Mc Barnet A. How OHM is getting the case for EM sounding methods heard by the offshore industry // First break. – 2003. – №11. – P. 65–68.