

## К ВОПРОСУ О МЕТОДАХ ВЕРИФИКАЦИИ ПРОГНОЗОВ

*Рябинин Г.В.*

*Камчатский филиал Геофизической службы РАН, г. Петропавловск-Камчатский, [gena@emsd.ru](mailto:gena@emsd.ru)*

### Введение

В настоящее время при подготовке экспертных заключений Камчатским филиалом Российского экспертного совета по прогнозу землетрясений, оценке сейсмической опасности и риска (КФ РЭС), согласно работе [5], используются более 20 методов прогноза землетрясений. В этой же работе представлены сводные данные по оценке эффективности технологии прогнозирования землетрясений, выполненной по единой методике с помощью формального критерия. Тем не менее, следует признать, что существующая система оценки сейсмической опасности, при всех ее достоинствах, имеет очевидный недостаток – отсутствие стандартов в методах верификации прогнозов. Тому есть объективные причины, связанные, по-видимому, с большим разнообразием исследуемых величин и индивидуальных особенностей прогностических методик и схем. В связи с этим, стандартизация верификационных процедур оценки качества прогностической информации представляется актуальной задачей.

В решении этой задачи может оказаться весьма полезным опыт, накопленный в гидрометеорологии, где широко используются различные критерии оценки качества прогностической информации [6]. Разнообразие используемых критериев обусловило необходимость введения единых стандартов верификации метеорологических прогнозов. В частности, такие стандарты были введены для оценки качества долгосрочных прогнозов (от 30 суток до 2 лет) аномалий температуры воздуха, осадков и температуры поверхности океана [8]. Одним из принятых в рамках этого документа стандартов верификации является критерий ROC. В данной работе представлен пример использования этого критерия для ретроспективной оценки качества прогнозов сильных землетрясений по данным гидрогеохимических наблюдений на Камчатке.

### Верификационный критерий ROC

Аббревиатура ROC расширяется как Relative (или Receiver) Operating Characteristics. В русскоязычной литературе встречаются разные варианты перевода: относительная оперативная характеристика (ООХ), относительная рабочая характеристика (ОРХ), сравнительная оперативная характеристика (СОХ). Критерий ROC представляет собой график, на котором по оси ординат откладываются доли удачных прогнозов (Hit rate), а по оси абсцисс – доли ложных тревог (False alarm rate), вычисленные для различных значений порогового уровня [6]. ROC может быть использована только в задачах верификации прогнозов для бинарных (дихотомических) предиктантов [7]. К классу дихотомических предиктантов относятся многие явления, которые могут рассматриваться как простые бинарные события. Например, осадки, наводнения, торнадо, землетрясения и т.п. либо происходят, либо не происходят, и соответствующие прогнозы либо будут, либо не будут выданы. Подобного рода оценки, выражающиеся в категориях «есть прогноз / нет прогноза», представляют простейший способ прогнозирования и принятия решения. Основная методологическая трудность при верификации таких прогнозов состоит в разумном сведении задачи к бинарной и в выборе управляющего параметра [3].

Процедура использования критерия ROC начинается с конструирования 2×2 таблиц сопряженности для бинарных событий (таблица 1).

Таблица 1. Общий вид таблицы сопряженности для дихотомических предиктантов

Прогноз события	Наличие события	
	ДА	НЕТ
ДА	a	b
НЕТ	c	d

По данным таблицы рассчитываются показатели  $H = a/(a+c)$  и  $F = b/(b+d)$ . Показатель  $H$  (Hit rate – показатель удачных прогнозов) является выборочной оценкой условной вероятности того, что спрогнозированное событие реализовалось и именуется также вероятностью обнаружения. Показатель  $F$  (False alarm rate – показатель ложных тревог) представляет собой условную вероятность

ложных тревог и именуется также вероятностью ошибочного детектирования [6]. Значения показателей  $H$  и  $F$  зависят от величины порогового уровня, которой определяется возможность принятия решения о прогнозе явления. Если решение о прогнозе явления принимается по переменной, имеющей непрерывное распределение, то использование порогового уровня сводит задачу прогноза к бинарной типа «есть прогноз/нет прогноза». Варьирование величиной порогового уровня внутри диапазона изменения переменной генерирует множество таблиц сопряженности и задает множество значений показателей  $H$  и  $F$ . Изображение зависимости показателей  $H$  и  $F$  на

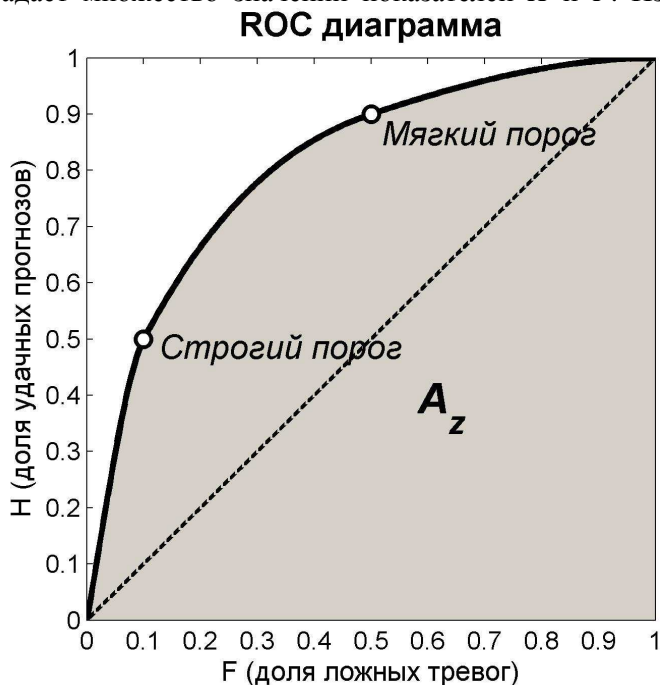


Рис. 1. Идеализированный пример ROC диаграммы. Сплошной линией обозначена ROC кривая. Круглыми маркерами показаны примеры положения на ROC кривой "строгих" и "мягких" пороговых уровней. Пунктирной линией обозначена главная диагональ ( $H = F$ ). Заливкой серого цвета обозначена площадь ( $A_z$ ) под ROC кривой.

диагональю. Такое положение ROC кривой означает, что система прогнозирует не явление, а его отсутствие. Если поменять местами в таблицах сопряженности, по которым рассчитывалась исходная ROC кривая, подсчеты в ячейках  $a - b$  и  $c - d$ , то получившаяся кривая будет лежать в левой треугольной области диаграммы и представлять собой зеркальное отражение исходной кривой относительно главной диагонали. Величина критерия  $A_z$  после такой трансформации будет превышать значение 0.5 и характеризовать эффективность прогноза отсутствия явления. Критерий  $A_z$  дает интегральную характеристику эффективности прогностической системы (схемы, метода) и позволяет сравнивать различные прогностические методики между собой. В практике метеорологических прогнозов принято, что прогностическую систему, характеризующуюся величиной  $A_z \approx 0.8$ , можно считать хорошей, а систему, характеризующуюся величиной  $A_z \approx 0.7$  – полезной [1].

Самостоятельной задачей при использовании критерия ROC является задача выбора порогового уровня, который определяет решение о прогнозе явления или события. Каждая точка на ROC кривой соответствует некоторому порогу. Строгие пороги, расположенные в левом нижнем углу диаграммы, минимизируют количество ложных тревог за счет малого количества удачных прогнозов. Мягкие пороги в правом верхнем углу диаграммы максимизируют количество удачных прогнозов за счет большого количества ложных тревог [3]. Выбор определенного порогового уровня принятия решения должен зависеть от социальной значимости прогнозируемого явления и возможных экономических и социальных последствий от ложного прогноза [3].

### Пример использования верификационного критерия ROC

Рассмотрим процедуру (алгоритм) использования критерия ROC для ретроспективной оценки качества прогнозов сильных землетрясений по данным гидрогеохимических наблюдений на

координатной плоскости носит название ROC диаграммы. Пример такой диаграммы представлен на рисунке 1.

В основе ROC критерия лежит математическая теория детектирования сигналов, которая была разработана при решении задач анализа способности электронных систем выделять сигналы из шума [7]. Все эмпирические ROC кривые обязательно проходят через точки (0,0) и (1,1), а также другие точки, расположенные внутри единичного квадрата. Наиболее широко используемой мерой эффективности прогностической системы (схемы) является мера  $A_z$ , численно равная площади под ROC кривой. Диапазон возможных значений меры  $A_z$  изменяется от 0 до 1. Идеальная прогностическая эффективность достигается тогда, когда величина  $A_z$  равна 1. Отсутствию эффективности соответствует величина  $A_z$  равная 0.5, когда кривая ROC совпадает с главной диагональю ( $H = F$ ), что равноценно ситуации случайного угадывания. При значениях  $A_z < 0.5$  кривая ROC проходит под главной

Камчатке. Результаты этих наблюдений показывают, что наиболее достоверным прогнозным признаком подготовки землетрясения является бухтообразное уменьшение концентрации отдельных показателей ионного состава воды самоизливающихся скважин [4]. В качестве рабочего примера возьмем временной ряд изменения концентрации хлор-иона (рис. 2а), где отмеченный прогнозный признак имеет наилучшую воспроизводимость. Для данного примера задача может быть сформулирована следующим образом: с использованием критерия ROC оценить эффективность ретроспективных прогнозов сильных землетрясений по прогнозному признаку, представляющему собой уменьшение концентрации хлор-иона в воде самоизливающейся скважины ГК-1. С целью выделения прогнозных признаков на фоне неинформативных факторов проведем предварительную обработку временного ряда  $Cl$ , в результате которой получим временной ряд  $Cl_{prg}$  (рис. 2б). Предварительная обработка, в данном случае, включала в себя удаление тренда и сглаживание остатков скользящим средним во временном окне 40 суток. Выборка сейсмических событий осуществлялась по каталогу КФ ГС РАН и включала в себя землетрясения с локальной магнитудой  $M_l \geq 6.0$ , произошедшие на глубине менее 100 км, эпицентры которых попадали внутрь окружности радиусом 350 км с центром в г. Петропавловске-Камчатском.

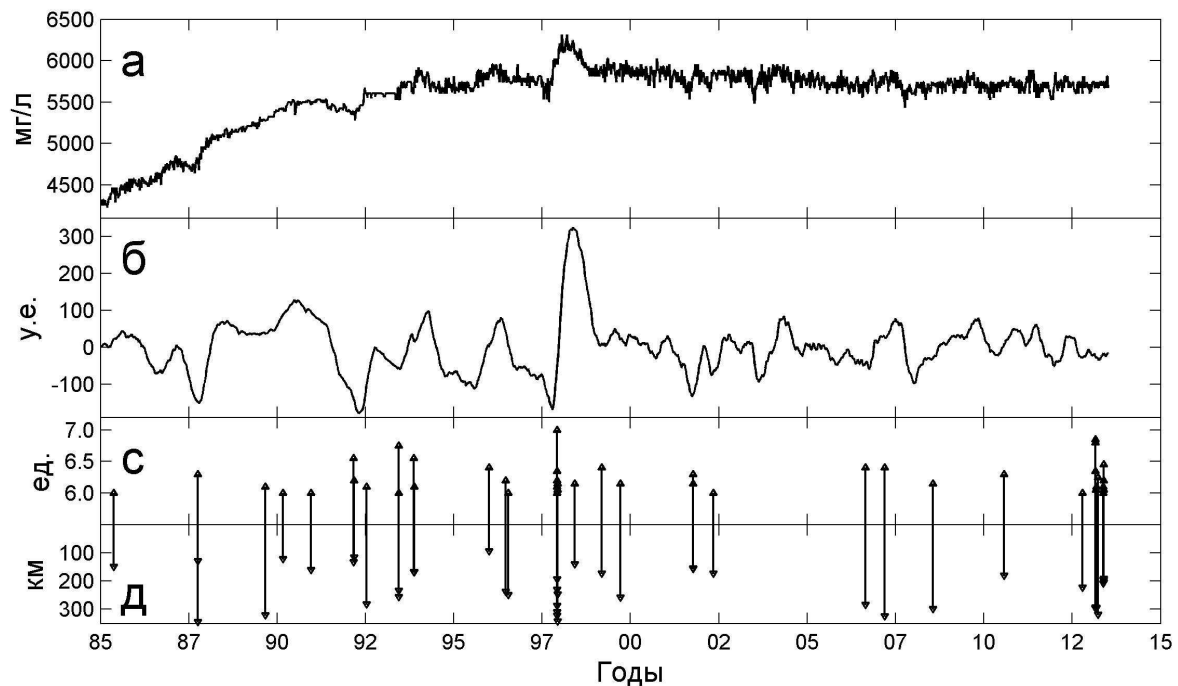


Рис. 2. Вариации концентрации хлор-иона в воде скважины ГК-1 (станция Пиначево) в сопоставлении с сильными ( $M_l > 6.0$ ) сейсмическими событиями: а – фактический временной ряд вариаций хлор-иона ( $Cl$ ); б – временной ряд вариаций хлор-иона после процедуры удаления тренда и сглаживания ( $Cl_{prg}$ ); в и д – временные последовательности сейсмических событий, выраженные в величинах локальных магнитуд ( $M_l$ ) и эпицентральных расстояний ( $\Delta$ , км) соответственно.

Наиболее сложным этапом верификационной процедуры является конструирование таблиц сопряженности – сведение задачи к бинарной и выбор управляющих параметров. Для последовательности сейсмических событий дихотомическая ситуация типа "есть землетрясение – нет землетрясения" определялась одним параметром – магнитудой землетрясения. Относительно магнитуды рассматривались два варианта. В первом варианте событие считалось реализовавшимся, если оно имело магнитуду  $M_l \geq 6.0$ , во втором варианте –  $M_l \geq 6.5$ . Для вариаций ряда  $Cl_{prg}$  дихотомическая ситуация "есть прогноз – нет прогноза" определялась величиной порогового уровня. В рассматриваемом случае альтернатива «есть прогноз» считалась реализовавшейся, если концентрация хлор-иона оказывалась меньше, чем пороговая концентрация. Связь между прогнозами и сейсмическими событиями оценивалась на скользящем временном интервале  $T$ . Для этого интервал  $T$  разбивался на две равные части. В первой (левой) части интервала устанавливался факт наличия или отсутствия прогноза. Во второй (правой) части интервала – наличие или отсутствие сейсмического события. Фактически, здесь использовался тот же подход, что и в работе [2], предполагающий определение типа интервала (есть событие – нет события) по поведению анализируемого параметра на предыдущем интервале (есть прогноз – нет прогноза). Четыре

комбинации возможных вариантов, при последовательном смещении скользящего временного интервала  $T$  на величину  $dT$  вдоль оси времени, определяли подсчеты в четырех ячейках  $2 \times 2$  таблицы сопряженности.

Фиксация величин всех рассмотренных параметров и величины пороговой концентрации даст одну таблицу сопряженности и одну точку на ROC диаграмме. Для получения полной ROC кривой (при фиксированных параметрах  $M_i$ ,  $T$  и  $dT$ ) необходимо повторить описанную выше процедуру для всех возможных значений пороговой концентрации, от ее минимума до ее максимума. На практике, для построения ROC кривой разумно будет ограничиться последовательностью из нескольких значений пороговой концентрации. При равенстве пороговой концентрации ее максимальному значению, для рассматриваемого примера, все «левые» части скользящих интервалов  $T$  будут содержать метку «есть прогноз», что приведет к обнулению ячеек  $c$  и  $d$  таблицы сопряженности. Соответственно, значения показателей  $F$  и  $H$  будут равняться единице, а точка на ROC кривой попадет в верхний правый угол ROC диаграммы. При равенстве пороговой концентрации ее минимальному значению все «левые» части интервалов  $T$  будут содержать метку «нет прогноза», что приведет к обнулению ячеек  $a$  и  $b$ . Величины показателей  $F$  и  $H$  при этом будут равняться нулю, и точка на ROC кривой попадет в нижний левый угол ROC диаграммы. Промежуточные значения пороговой концентрации зададут конфигурацию кривой на ROC диаграмме. Величина площади под полученной кривой ( $A_z$ ) и будет являться ретроспективной оценкой эффективности рассматриваемой прогностической схемы.

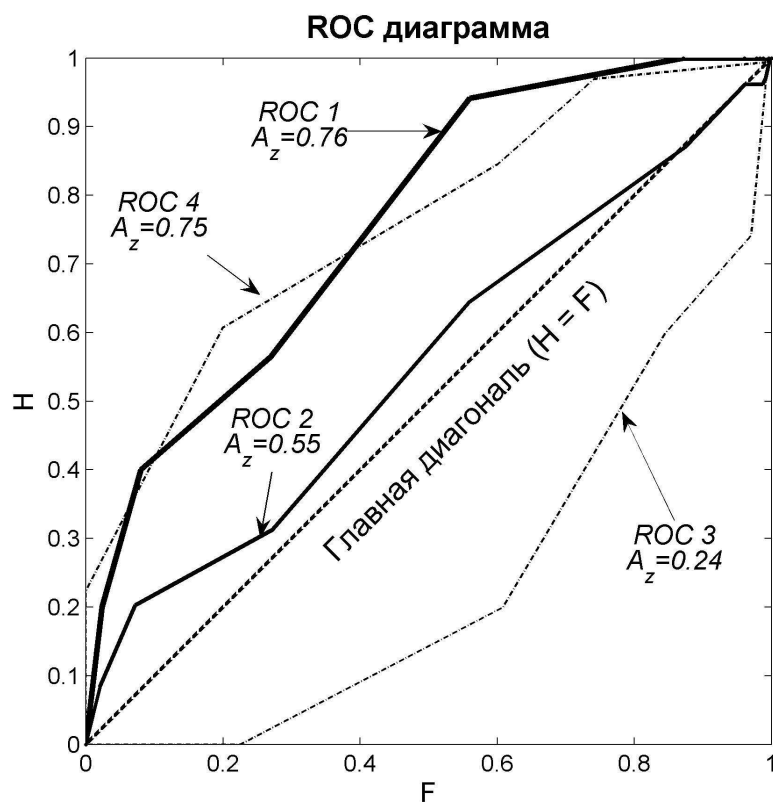


Рис. 3. ROC диаграмма, иллюстрирующая качество ретроспективных прогнозов землетрясений по вариациям временного ряда Clrg. Пояснения в тексте.

На рисунке 3 приведено несколько примеров ROC кривых. Для всех ROC кривых величины параметров  $T$  и  $dT$  равнялись 100 и 3 суткам соответственно. Пороговые значения концентрации варьировались от -200 до 350 условных единиц с шагом в 50 условных единиц (-200, -150, -100, ..., 0, 50, 100, ..., 350). Кривые ROC1 и ROC2 представляют собой оценку эффективности прогноза сейсмических событий с магнитудами ( $M_i$ ) больше 6.5 и 6.0 соответственно. И в первом и во втором случае альтернатива «есть прогноз» реализовалась, если концентрация хлор-иона оказывалась ниже пороговой. Площади под ROC кривыми оценивались методом трапеций [7]. На рисунке видно, что обе кривые лежат над главной диагональю. Формально из этого следует, что предложенная прогностическая схема эффективно прогнозирует и события с  $M_i \geq 6.0$  и события с  $M_i \geq 6.5$ . Однако в случае кривой ROC2 эффективность лишь незначительно отличается от «случайного угадывания». Практически это означает, что использование рассматриваемого метода для прогноза землетрясений

с  $M_l \geq 6.0$  малоэффективно, поскольку количество правильных прогнозов будет соизмеримо с количеством ложных тревог при любом значении пороговой концентрации. В случае с кривой ROC2 величина параметра  $A_z$  равняется 0.76. Такую прогностическую эффективность можно считать вполне удовлетворительной. Из этого можно сделать вывод о том, что уменьшение концентрации хлор-иона в воде скважины ГК-1 является достаточно надежным прогностическим признаком землетрясений с  $M_l \geq 6.5$  и  $\Delta \leq 350$  км.

Интересный результат получается, если в качестве прогностического признака рассматривать не уменьшение, а увеличение концентрации хлор-иона. Этот результат иллюстрирует кривая ROC3, построенная для тех же значений параметров, что и кривая ROC1. Как видно на рисунке, кривая ROC3 расположена ниже главной диагонали. Если трансформировать таблицу сопряженности, заменив подсчеты в ячейке  $a$  на подсчеты в ячейке  $b$ , а подсчеты в ячейке  $c$  на подсчеты в ячейке  $d$ , то получившаяся кривая ROC4 будет представлять собой зеркальное отражение кривой ROC3 относительно главной диагонали. Это означает, что кривая ROC3 (ROC4) представляет собой оценку эффективности прогноза отсутствия сейсмических событий с  $M_l \geq 6.5$  и  $\Delta \leq 350$  км. Другими словами, можно сказать, что, чем выше концентрация хлор-иона в воде скважины ГК-1, тем выше вероятность того, что землетрясение с  $M_l \geq 6.5$  и  $\Delta \leq 350$  км не произойдет.

### Обсуждение результатов

Представленная работа, в большой мере, имеет методический характер. Автор далек от мысли, что изложенный в работе алгоритм построения верификационных оценок на основе ROC критерия является оптимальным и может быть без изменений использован для верификации методик, основанных на прогностических признаках иного рода, отличных от тех, которые рассмотрены в данной работе. Основная трудность, и об этом уже упоминалось выше, заключается в сведении задачи к бинарной. Для разных прогностических методик алгоритм конструирования таблиц сопряженности может несколько отличаться в зависимости от индивидуальных особенностей используемых прогностических признаков землетрясений и набора управляющих параметров.

Характеризуя критерий ROC как метод верификации необходимо отметить следующие его достоинства. Во-первых, критерий ROC дает интегральную характеристику эффективности прогностической методики при разных пороговых уровнях принятия решения, в отличие от множества других критериев, основанных на единственном значении пороговой величины [6]. Во-вторых, критерий ROC дает формальные основания для выбора величины порогового уровня принятия решения в зависимости от социальной и экономической значимости прогнозируемого явления. Очевидно, что для задачи прогноза сильных социально значимых сейсмических событий пороговый уровень принятия решения о прогнозе землетрясения должен находиться в области "строгих порогов" для того, чтобы свести к минимуму вероятность ложных тревог.

Относительно рассмотренного в данной работе примера можно заключить, что результаты формальной оценки качества ретроспективных прогнозов сильных землетрясений по вариациям хлор-иона в воде скважины ГК-1 демонстрируют вполне приемлемую эффективность используемой прогностической схемы применительно к событиям с  $M_l \geq 6.5$  и  $\Delta \leq 350$  км.

### Список литературы

1. Бундель А.Ю. и др. Верификация кратко- и среднесрочных прогнозов осадков на основе системы ансамблевого моделирования Гидрометцентра России // Метеорология и гидрология. 2011. № 10. С. 20-34.
2. Гусев А.А. Прогноз землетрясений по статистике сейсмичности // Сейсмичность и сейсмический прогноз, свойства верхней мантии и их связь с вулканизмом на Камчатке. Новосибирск: Наука, 1974. С. 109-119.
3. Муравьев А.В. Стандартизированная система верификации долгосрочных метеорологических прогнозов (SVS LRF) // 80 лет Гидрометцентру России. [http://method.hydromet.ru/publ/books/80\\_years/murav.pdf](http://method.hydromet.ru/publ/books/80_years/murav.pdf)
4. Рябинин Г. В., Хаткевич Ю. М. Гидрогеохимические эффекты, предшествующие сильным землетрясениям Камчатки. Алгоритм идентификации и морфологический анализ // Вестник КРАУНЦ. Науки о Земле. 2009. Вып. 13. № 1. С. 107-122.
5. Чебров В.Н. и др. Прогнозирование землетрясений на Камчатке. М.: Светоч Плюс, 2011. 304 с.
6. Jolliffe I.T., Stephenson D.B. Forecast Verification: A Practitioner's Guide in Atmospheric Science. John Wiley & Sons, Ltd., West Sussex, England, 2003. 240 p.
7. Mason I.B. A model for assessment of weather forecast // Austral. Met. Mag. 1982. 30. P. 291-303.
8. Standardized Verification System (SVS) for Long-Range Forecasts (LRF) // New Attachment II-9 to the Manual on the GDPS (WMO-No. 485), 2002, vol. I. 23 p.