

УДК 550.34.013

ДЕКЛАСТЕРИЗАЦИЯ ПОТОКА СЕЙСМИЧЕСКИХ СОБЫТИЙ, СТАТИСТИЧЕСКИЙ АНАЛИЗ

© 2019 г. В. Ф. Писаренко¹, М. В. Родкин^{1,2*}

¹Институт теории прогноза землетрясений и математической геофизики РАН, г. Москва, Россия

²Институт морской геологии и геофизики ДВО РАН, г. Южно-Сахалинск, Россия

*E-mail: rodkin@mitp.ru

Поступила в редакцию 25.12.2018 г.

Принята в печать 25.03.2019 г.

Известные методы выделения кластеров сейсмических событий, которые в основном образуются афтершоками, часто включают в себя дискуссионные исходно задаваемые положения или сложную систему последовательных приближений. Одним из последних по времени и наиболее логически последовательным методом выделения кластеров афтершоков является метод ближайшего соседа (*NND*), но этот метод и наиболее трудно реализуем алгоритмически. В работе предлагается новый метод декластеризации, основанный на обобщенном расстоянии (*OP*), использующий некоторые положения метода *NND*, но столь же простой в реализации, как и ранее предложенные для этой цели оконные методы. При анализе и обосновании этого нового подхода применяется процедура случайного тасования сейсмических событий по времени, призванная задать модель реального каталога, но лишенного генетических связей между разными событиями. Для 17 регионов проводится сравнение эффективности известных оконных методов, *OP*-метода и метода *NND* по ряду критериев. Показано, что *OP*-метод, в среднем, работает заметно эффективнее стандартных оконных методов и практически не уступает методу *NND*. В заключении обсуждается определенная условность разделения событий на основные и зависящие толчки.

Ключевые слова: алгоритмы выделения афтершоков, случайное перемешивание, окно обобщенного расстояния.

DOI: <https://doi.org/10.31857/S0002-33372019538-52>

ВВЕДЕНИЕ

Поток сейсмических событий включает сгущения по времени и по пространству — кластеры событий. Понятие «кластер» подразумевает какое-то обобщенное расстояние между событиями, которое для пар событий из одного кластера меньше (или статистически меньше), чем между событиями из данного кластера и другими событиями. Часто встречаются и одиночные события, достаточно далеко отстоящие от других. Обычно в кластере имеется землетрясение с заметно большей магнитудой — главное событие. События, предшествующие главному, называют форшоками, а следующие за главным — афтершоками (возможны также кластеры, состоящие из близких по величине толчков — сейсмические рои, но такие явления относительно редки). Афтершоки, как указывает само название, это сейсмические толчки, следующие после основного

толчка, и индуцированные этим толчком. Это определение не является, однако, точным, количественным (что значит индуцированные?).

Наиболее распространена модель сейсмического процесса ETAS (Epidemic Type Aftershock Sequence) [Ogata, Zhuang, 2006; Console et al., 2010; Zhuang et al., 2011; и др.], в которой афтершоки и основные толчки принципиально не различаются: каждый толчок является в той или иной степени афтершоком **всех** предыдущих толчков. Из статьи [Ogata, Zhang, 2006]: «Indeed, the recognition of seismic anomalies (*форшоков, афтершоков, затихший и пр.*) appears to be subjective and still seems under development, and even controversial. To overcome these difficulties we first need to use a practical statistical space–time model that represents the ordinary seismic activity, rather than carrying out a declustering algorithm that removes aftershock events from a catalog.»

ETAS-модель условной (при заданном прошлом) интенсивности точечного сейсмического процесса имеет вид:

$$\lambda(t) = \mu + K \cdot \sum_{t_j < t} (t - t_j + c)^{-p} \exp(\alpha(m_j - m_0)), \quad (1)$$

где: μ — постоянная фоновая интенсивность независимых Пуассоновских толчков; K — некоторая константа; сумма включает **все** толчки до момента t ; t_j — моменты времени толчков, произошедших до момента t ; c, p, α — параметры модели; m_j — магнитуды прошлых толчков; m_0 — магнитуда нижнего порога каталога.

Как мы видим, модель ETAS не выделяет одного конкретного предка для любого толчка: все предшествующие толчки вносят свой вклад в интенсивность, порождающую любой толчок. Такой подход представляется обоснованным, и он позволяет получить некоторое вероятностное разделение толчков на фоновые и зависимые события, но он неудобен для исследования статистики и физики афтершокового процесса, так как не дает разбиения на отдельные афтершоковые последовательности. В работе [Baiesi, Paczuski, 2004] предлагается считать «предком» данного афтершока единственный предшествующий толчок, вклад которого можно полагать максимальным согласно некоторому правилу. Такой подход позволяет рассматривать в рамках модели ETAS также и отдельные афтершоковые последовательности. Он, однако, противоречит исходной предпосылке модели, ввиду чего логичность модели ETAS несколько нарушается.

МЕТОДЫ ВЫДЕЛЕНИЯ ПОСЛЕДОВАТЕЛЬНОСТЕЙ АФТЕРШОКОВ

Оконные методы

Разнообразие подходов к выделению афтершоков весьма велико, ниже мы кратко упоминаем только те из них, что наиболее тесно соотносятся с нашим подходом. Более полный и объемный обзор возможных способов выделения афтершоков можно найти в работах [Stiphout et al., 2012; Дещеревский и др., 2016]. Исторически первыми методами выделения афтершоковых последовательностей были оконные методы (см., например, [Gardner, Knopoff, 1974; Uhrhammer, 1986]). Эти методы, вследствие простоты реализации, часто и теперь оказываются наиболее используемыми. Оконный метод состоит в том, что к выделенному определенным способом главному событию $t_k, \lambda_k, \varphi_k, h_k, m_k$ приставляют пространственно-временное окно

$$0 < t - t_k < W_1(m_k), \quad (2)$$

$$0 < r(\lambda_k, \varphi_k; \lambda, \varphi) < W_2(m_k), \quad (3)$$

где: $r(\lambda_k, \varphi_k; \lambda, \varphi)$ — расстояние между точками на поверхности Земли с указанными широтами и долготами, а $W_1(m_k), W_2(m_k)$ — пороги, зависящие от магнитуды главного толчка (чем больше магнитуда, тем больше порог). Такой вид имеют, например, окна Гарднера–Кнопова, Урхаммера и др., (см. [Gardner, Knopoff, 1974; Uhrhammer, 1986; Stiphout et al., 2012]). События, попавшие в окно (2)–(3), считаются афтершоками главного события $t_k, \lambda_k, \varphi_k, h_k, m_k$. Несколько более сложный алгоритм используется в методе [Reasenber, 1985], где допускаются вторичные афтершоковые последовательности (афтершоки от афтершоков), при этом эти вторичные афтершоки включаются в первичную последовательность афтершоков.

Метод ближайшего соседа

Более сложный вариант выделения зависимых и независимых событий связан с заданием некоторого обобщенного расстояния между толчками [Baiesi, Paczuski, 2004; Zalyapin et al., 2008; Zalyapin, Ben-Zion, 2013; 2013a; и др.]. Субъективный момент в эти построения вносит выбор формулы, используемой для расчета расстояния, и выбор значения порогового расстояния для различения независимых и зависимых событий. Обычно применяется субъективное решение, в той или иной степени подкрепляемое некоторыми качественными соображениями. Такая субъективность, несомненно, уменьшает объективность получаемых выводов. Мы попытаемся, насколько это возможно, вывести далее некоторые статистические свойства афтершоков, не прибегая к субъективному заданию процедуры их выделения.

Одно из наиболее эффективных обобщенных расстояний используется в так называемом методе «ближайшего соседа» (Nearest Neighbor Distance — NND) [Baiesi, Paczuski, 2004; Zalyapin, Ben-Zion, 2013; 2013a]. Каждое событие каталога характеризуется 5 числами: $t, \lambda, \varphi, h, m$ (время, широта, долгота, глубина, магнитуда). **Несимметричное** расстояние $D_k(i)$ от события $t_k, \lambda_k, \varphi_k, h_k, m_k$ до другого события $t_i, \lambda_i, \varphi_i, h_i, m_i$ задается формулой:

$$D_k(i) = \begin{cases} (t_k - t_i) r^d \cdot 10^{-bm_i}; & (t_k - t_i) > 0 \\ \infty; & (t_k - t_i) \leq 0. \end{cases} \quad (4)$$

Здесь: r — расстояние между эпицентрами событий (глубины очагов не учитываются, хотя в принципе их несложно учесть); d — некоторый показатель, обычно полагаемый равным

среднему значению фрактальной размерности поля эпицентров $d = 1.6$. Параметр b — показатель степени в законе Гутенберга–Рихтера ради простоты часто полагают равным единице $b = 1$, что обычно выполняется с удовлетворительной точностью. У каждого события $t_k, \lambda_k, \varphi_k, m_k$ (кроме первого события в каталоге) есть один ближайший «предок» — это то из предшествующих событий, которое имеет наименьшее расстояние до него. Это минимальное расстояние и есть NND_k^B :

$$NND_k^B = \min(D_k(i)), \quad (5)$$

где минимум берется по i . Заметим, что при этом допускается, что магнитуда основного толчка меньше чем афтершока; такие случаи естественно рассматривать как форшоки.

В работе [Zalyapin, Ben-Zion, 2013; 2013a] показано, что NND_k^B является весьма эффективным средством для разделения каталога на кластеры. Распределение величин оказывается двугорбым (см. рис. 1 или, например, рис. 4 в работе [Zalyapin, Ben-Zion, 2013]), где горб с меньшими расстояниями естественно принять как формируемый зависимыми событиями, а горб с большими расстояниями — как независимыми. Если ввести некоторый порог, то можно выделить кластеры (цепочки) событий, в которых каждый ближайший предок отстоит на расстоянии меньше этого порога. Вырожденные цепочки состоят из одного события. Но при выборе порога также неизбежен некоторый произвол, хотя обычно приводят некоторые соображения в пользу того или иного выбора.

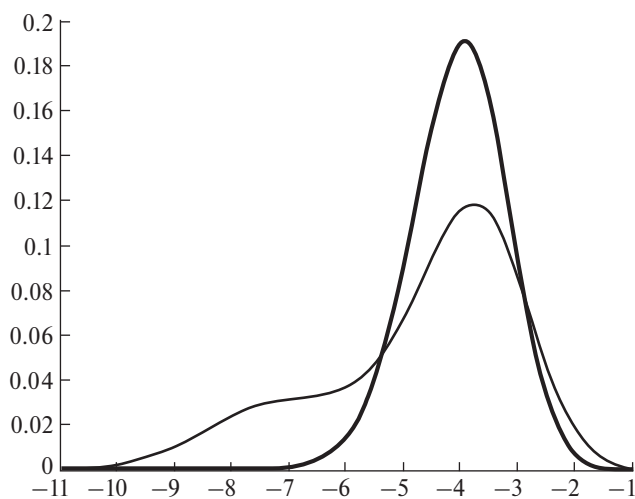


Рис. 1. Эмпирическая плотность вероятности $\lg(NND_k^B)$ для исходного каталога (двугорбая кривая) и для каталога со случайно переставленными моментами времени (одногогорбая кривая); $d = 1.6$.

Метод случайного перемешивания

Основным средством нашего статистического анализа будет служить метод случайного перемешивания (случайного тасования) событий по времени. Этот метод является разновидностью классического бутстреп-метода [Efron, 1979]. Он неоднократно применялся ранее для статистического анализа каталогов землетрясений [Писаренко, Родкин, 2007; Pisarenko, Rodkin, 2010; 2013]. Применительно к проблеме анализа афтершоков этот метод упоминается в работе [Дещеревский и др., 2016]. В этой работе проведен тщательный анализ группируемости землетрясений в смысле изменений вероятности возникновения ближайшего (по времени) землетрясения; такой подход, однако, как представляется, существенно ограничивает разнообразие проявлений эффекта группируемости. По-видимому, именно в результате такого ограничения авторами был получен несогласующийся с другими данными вывод о практической независимости параметров группирования от магнитуды землетрясений. Мы будем пользоваться случайной перестановкой моментов времени событий для разрушения причинных («генетических») связей событий в кластерах. Мы пробовали и вариант перемешивания со случайным заданием момента событий на заданном интервале времени; результаты такого перемешивания близки к варианту случайной перестановки моментов времени событий. Далее мы будем обсуждать вариант случайной перестановки моментов времени как меньше искажающий структуру реального каталога землетрясений. На рис. 1 показаны полученные эмпирические плотности вероятности $\lg(NND_k^B)$ при $d = 1.6$ (детали вычисления этой плотности в Приложении) для исходного каталога и для каталога со случайной перестановкой моментов времени. Был использован глобальный Гарвардский каталог GCMT 1976–2016 гг., $M_w \geq 5.3$, глубина очага меньше 70 км.

Вычисление NND_k^B для реальных каталогов приводит к двумодальным распределениям, состоящим из двух четко выраженных горбов (см. рис. 1 или, например, рис. 4 в работе [Zalyapin, Ben-Zion, 2013]). Такое двугорбое распределение указывает на наличие кластеров. Левый горб (меньшие минимальные расстояния) относится, в основном, к парам землетрясений из одного кластера, а правый (большие минимальные расстояния) к парам из разных кластеров. Следует сразу отметить один недостаток при интерпретации гистограмм NND_k^B . Распределение значения NND_k^B зависит от номера « k »

в каталоге, точнее, от того, сколько событий в каталоге предшествует событию « k ». Ведь минимум в (5) берется по всем предшествующим событиям, а их ровно $(k-1)$. Таким образом, гистограммы состоят из неоднородных наблюдений. Для малых k берется минимум из меньшего числа наблюдений, чем для больших k . По этой причине наблюдаемая величина NND_k^B систематически смещается в меньшую сторону с увеличением k . Этот эффект особенно ярко проявляется на каталогах со случайно переставленными моментами времени. В реальных каталогах он также присутствует, но маскируется сильной нестационарностью реального сейсмического потока. Следует учесть также и нестационарность каталога, связанную с ростом со временем доли более слабых событий за счет совершенствования системы регистрации. Однако эта тенденция достаточно слаба и практически не влияет на результаты анализа, если ограничиться событиями выше порога достаточно надежной регистрации. Для используемого далее GCMT-каталога этот порог соответствует магнитудам $M_w \geq 5.3$.

Для иллюстрации эффекта систематического смещения мы использовали глобальный GCMT-каталог, $n = 17\,826$ событий; $5.3 \leq m \leq 9.1$; $h \leq 70$ км. Моменты времени каталога на интервале [01.01.1976–12.01.2016] случайно переставлялись, а координаты не менялись. Таким способом было сгенерировано 10 рандомизированных реализаций каталога. Измерялись средние значения $\lg(NND_k^B)$ в скользящих интервалах

$(k-50; k+50)$, $k = 51, \dots, 17\,776$. Разброс этих средних оценивался стандартным отклонением от выборочного среднего (по 10 реализациям).

На рис. 2а для каталогов со случайными перестановками видно, что значения $\lg(NND_k^B)$ быстро уменьшаются с увеличением k примерно до значений $k \approx 6000$, а затем уменьшение практически прекращается. На рис. 2б аналогичный график приведен для реального каталога. Уменьшение здесь менее четко выражено, но несмотря на значительный разброс видно, что средние значения для $k < 6000$, в основном, больше значений для $k > 6000$.

Чтобы избежать влияния этой нестационарности мы будем в дальнейшем использовать минимальные расстояния NND_k^B только для $k > 6000$. Естественно, что для других каталогов значение k изменится. Таким образом, в дальнейшем (если это не оговорено специально) мы будем рассматривать для метода *NND укороченный* каталог для значений $6001 \leq k \leq 17\,826$ с магнитудами $m \geq 5.3$, что соответствует временному отрезку 19.11.1992–12.01.2016 (11826 событий).

Окно Обобщенного расстояния (ОР-окно)

В ранее предложенных окнах (2)–(3) ограничения по времени и по пространству действуют независимо друг от друга. А в методе *NND* эти ограничения используются более гибко — в виде произведения функций от расстояния и времени. Однако метод *NND* довольно сложен в применении, он требует прослеживания связей,

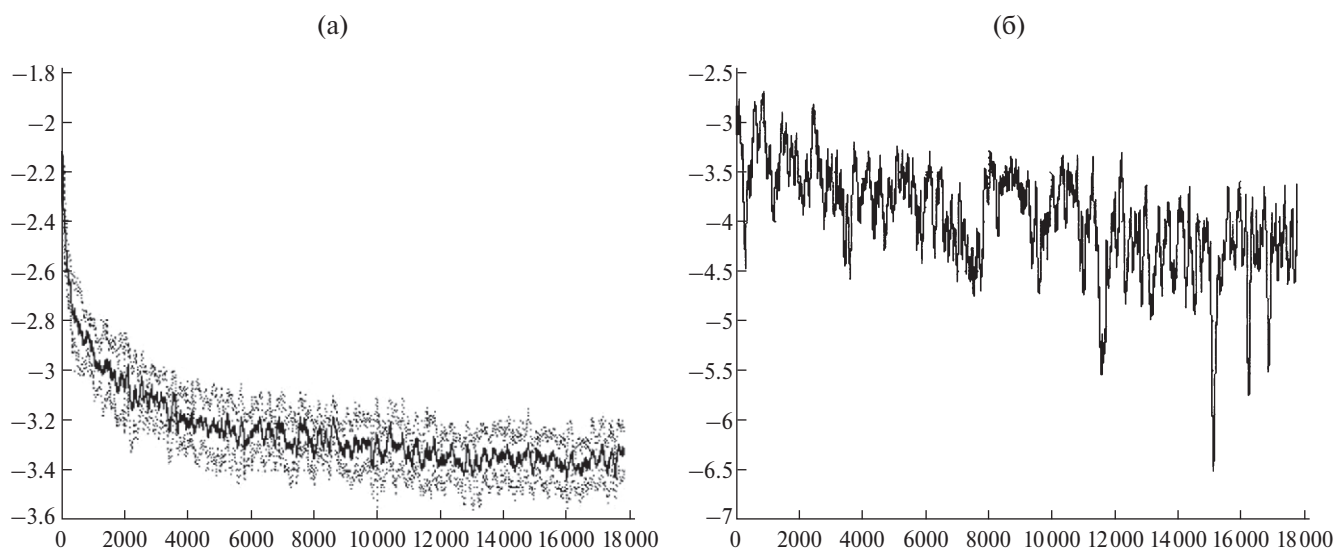


Рис. 2. Средние значения $\lg(NND_k^B)$ в скользящем окне от $(k-50)$ события до $(k+50)$ события и разброс этих средних (стандартное отклонение по 10 реализациям) для каталога с перестановками по времени (а) и для реального каталога (б).

не превосходящих заданный порог, выделения связанных графов и т. д. Возникает мысль воспользоваться расстоянием NND в простой оконной форме. Пусть имеется главное событие t_k , λ_k , φ_k , h_k , m_k . Окно, определяемое порогом W , задается ограничением:

$$(t_k - t) \cdot r^d < W \cdot 10^{bm}_k, (t_k - t) > 0. \quad (6)$$

Здесь: r — расстояние между эпицентрами событий; d — некоторый показатель. Окно (6) будем называть окном Обобщенного Расстояния (ОР-окно). Оно использовалось в работах [Писаренко, Родкин, 2018; Pisarenko, Rodkin, 2018]. Здесь мы обсудим его более подробно и сравним эффективность этого метода выделения афтершоков с другими подходами.

Заметим, что отбор в качестве афтершоков событий с помощью порогового соотношения (6) означает, что в начале афтершоковой серии $((t_k - t)$ мало) в качестве афтершоков могут отбираться и достаточно удаленные от главного толчка события. В последнее время все больше свидетельств приводится в пользу того, что афтершоки, в особенности, удаленные, инициируются проходящей сейсмической волной, дающей начало процессу генерации афтершоков. При этом вызванные события могут происходить со значительной задержкой [Freed, 2005; Felzer, Brodsky, 2007; Соболев, Закржевская, 2013; Lippiello et al., 2015; Кочарян, 2016]. Окно (6) согласуется с таким механизмом генерации афтершоков, предполагая возможность отбора ранних, удаленных событий в качестве афтершоков.

Сравнение эффективности выделения кластеров различными методами

Сравним ОР-окно с другими традиционными окнами и покажем его относительную эффективность. Для выделения афтершоков с помощью ОР-окна будем начинать с максимального события в каталоге и присоединять к нему в качестве его кластера все события с расстоянием (6) меньше выбранного порога W . Затем будем выбрасывать этот кластер и переходить к следующему максимальному событию. Для изучения характеристик ОР-окна мы также будем использовать укороченный каталог, но только укорачивать его нужно с конца, т. е. отбрасывать какое-то количество последних наблюдений. Как показали предварительные расчеты, для ОР-окна в качестве подходящей верхней границы можно взять момент времени 05.08.2005 г. Поэтому в дальнейшем, если не будет специальных

оговорок, мы будем использовать для ОР-окна и других окон укороченный каталог на временном отрезке 01.01.1976–05.08.2005 гг. (12000 событий $M \geq 5.3$). Завершая описание нового ОР-окна, заметим, что любой оконный метод не может быть абсолютно безошибочен. Действительно, для любого пространственно-временного окна и любого каталога существует не равная нулю вероятность попадания независимого события в это окно.

Для сравнения эффективности выделения кластеров различными методами мы снова будем использовать метод случайной перестановки моментов наступления событий, разрушающий причинные («генетические») связи, существующие в отдельных кластерах. Из реального каталога событий мы будем формировать случайные каталоги с помощью *случайной перестановки моментов наступления событий*. Получив таким способом N различных случайных каталогов, можно будет усреднить оценки различных параметров с целью уменьшения их разброса. Такую процедуру можно рассматривать как разновидность известного бутстреп-метода [Efron, 1979].

Начнем с ОР-метода. Сформируем из данного реального каталога N случайных каталогов. Рассмотрим нормированные расстояния между всевозможными парами событий:

$$(t_k - t_i) \cdot r^d \cdot 10^{-m}_k, (t_k - t) > 0. \quad (7)$$

Отметим, что в плане настоящего рассмотрения нас фактически интересуют не все возможные пары событий, а только события с расстояниями, приведенными на рис. 1 или близкие к таковым. Действительно, при рассмотрении задачи кластеризации землетрясений нам не важны пары точек с большими расстояниями, сравнимыми с пространственно-временными размерами рассматриваемой области. При этом конкретный выбор этого максимально принимаемого в расчет расстояния не существен, можно задать любое, несколько отступая вправо от правых горбов на рис. 1. При расчетах использованы ограничения по времени — не более 1 года. По расстоянию — не более 100 км. Функцию распределения этих расстояний для пар из реального каталога обозначим $F_{real}(x)$, а в случайном каталоге — $F_{rand}(x)$.

Мы не сравниваем эти расстояния с каким-либо порогом W и не принимаем решений о принадлежности каких-либо пар к разным кластерам (или к одному и тому же кластеру),

а просто отмечаем, что для случайных каталогов с перемешиванием нормированные расстояния (7) будут **в среднем** больше, чем для реального каталога (так как в реальном каталоге значительная часть событий входит в кластеры с малыми взаимными расстояниями). Определим в качестве оптимального такой порог W , при котором это отличие максимально, и сравним, насколько оно выражено в разных случаях. Степень отличия мы будем характеризовать минимальной вероятностью:

$$\min[F_{rand}(W) + (1 - F_{real}(W))], \quad (8)$$

где минимум берется по всем значениям W . При заданном W величина $F_{rand}(W)$ равна вероятности того, что пара из случайного каталога имеет расстояние меньше W , а величина $(1 - F_{real}(W))$ равна вероятности того, что пара из реального каталога имеет расстояние больше W . Таким образом, их сумму $F_{rand}(W) + (1 - F_{real}(W))$ можно рассматривать как суммарную вероятность ошибки при использовании порога W для классификации пар. Эту величину мы и принимаем в качестве меры отличия перекрывающихся распределений $F_{real}(W)$ и $F_{rand}(W)$. Заметим, что можно было бы минимизировать и любую взвешенную сумму вероятностей ошибок. В качестве

статистических оценок функций $F_{real}(W)$ и $F_{rand}(W)$ берутся стандартные эмпирические функции распределения расстояний пар или их сглаженные аналоги. Мы сглаживали эти функции ядерным методом (Приложение 1). При этом для оценки среднего и разброса $F_{rand}(W)$ мы брали $N = 25$ случайных каталогов.

Мы применили описанную процедуру к 17 каталогам регионального масштаба. Каталоги выбирались как отвечающие определенной сеймотектонической структуре, скажем Курильской или Алеутской зоне субдукции. Рассмотрение отдельных региональных каталогов и указанных выше ограничений по времени и пространству удобно в данном случае и в том плане, что снимается времяемкая и физически неоправданная процедура поиска возможных фор- и афтершоков в сильно отдаленных сеймотектонических зонах. Схема с расположением рассматриваемых регионов с их номерами дана на рис. 3.

Для сравнения эффективности ОР-метода мы выбрали метод NND [Baiesi, Paczuski, 2004; Zalyapin, Ben-Zion, 2013] и два часто используемых окна: окна Гарднера–Кнопова и Урхаммера. Наиболее часто используемое окно Гарднера–Кнопова [Gardner, Knopoff, 1974] имеет вид:

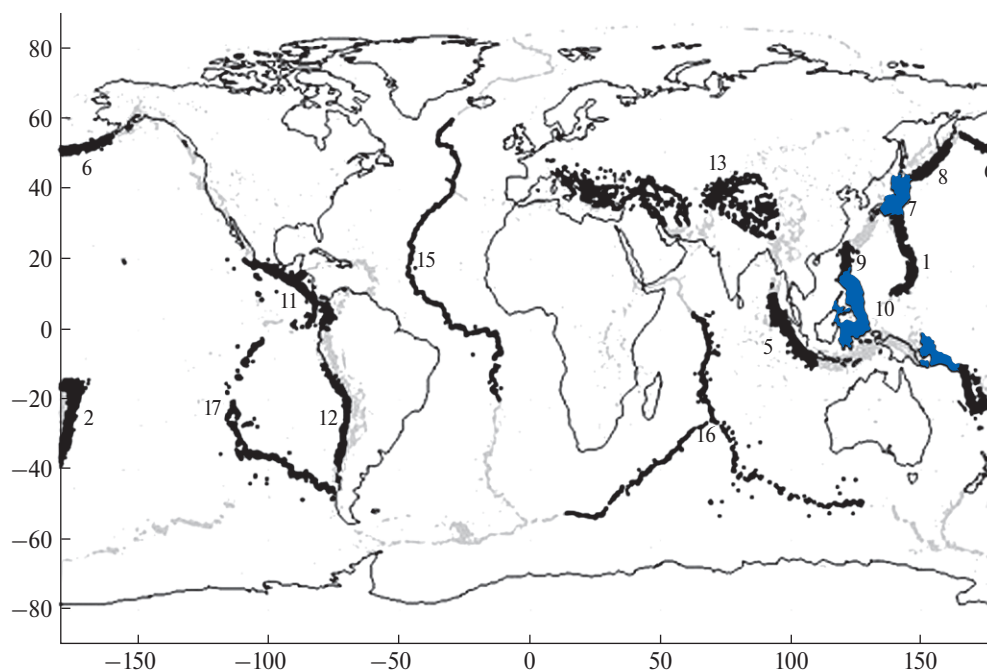


Рис. 3. Схема расположения рассматриваемых регионов и их нумерация. События в выбранных регионах даны черными или синими точками (номера регионов – цифрами), соприкасающиеся регионы даны разным тоном серого.

$$T_{norm} = \begin{cases} (t_k - t_i) \cdot 10^{-2.739 - 0.32M}; & M \geq 6.5; \\ (t_k - t_i) \cdot 10^{-0.547 - 0.54M}; & M \leq 6.5; \end{cases} \quad (9)$$

$$D_{norm} = 10^{-0.983 - 0.124M}.$$

Окно Гарднера–Кнопова $T_{norm} \leq 10^W$; $D_{norm} \leq 10^W$ имеет стандартный размер при $W=0$.

Окно Урхаммера [Uhrhammer, 1986]:

$$T_{norm} = (t_k - t_i) \cdot 10^{1.2464 - 0.5316M}; \quad M \geq 6.5; \quad (10)$$

$$D_{norm} = 10^{0.4447 - 0.3992M}.$$

Окно Урхаммера $T_{norm} \leq 10^W$; $D_{norm} \leq 10^W$ имеет стандартный размер при $W=0$.

ОР-окно имеет вид:

$$((t_k - t_i)/365) \cdot r^d \cdot 10^{-bm_k} < 10^W, \quad (11)$$

где: b — коэффициент наклона закона Гутенберга–Рихтера для данного региона; d — фрактальная размерность. Стандартный размер окна соответствует значениям $W=-5$, $d=1.6$. Отметим, что ОР-окно зависит от локального (регионального) параметра b , что дает возможность несколько адаптироваться к данному региону, чего нет у других окон.

На рис. 4–рис. 5 приведены графики суммарной вероятности ошибки для двух регионов: Курил ($b=0.930$) и Алеут ($b=0.921$), оцененных с помощью ОР-окна. Показана исходная кривая и сглаженный с помощью ядер-плотности вариант. Оценки параметра b получались стандартным методом правдоподобия с использованием усеченного распределения Гутенберга–Рихтера.

В табл. 1 приведены значения минимальной суммарной вероятности ошибки для всех 17 регионов.

Иногда бывает полезно отдельно рассмотреть характеристики сейсмического режима для различных типов сеймотектонических провинций, а именно, для зон субдукции, внутриконтинентальной коллизии и для срединно-океанических хребтов. Эти обобщенные зоны были сформированы на основе приведенных выше регионов. Средние по зонам оценки приведены в табл. 2. Из табл. 2 видно, что наиболее явное разделение, минимальное значение суммы $F_{rand}(W) + (1 - F_{real}(W))$, получается для внутриконтинентальных зон коллизии. И наихудшее разделение, в среднем, получается для срединно-океанических хребтов. Можно предположить, что такой результат связан с наиболее точной информацией для внутриконтинентальных зон сейсмичности и наименее полной (вследствие удаленности сеймостанций) для зон срединно-океанических хребтов.

Из табл. 1, табл. 2 мы видим, что метод NND и ОР-окно имеют вероятность ошибки в среднем на 0.12 меньше, чем два других окна. Это различие выходит за рамки погрешностей, является значимым. Поэтому можно заключить, что метод NND и ОР-окно предпочтительней упомянутых традиционных окон. Подчеркнем, что этот вывод мы сделали, не используя исходного задания величин порогов.

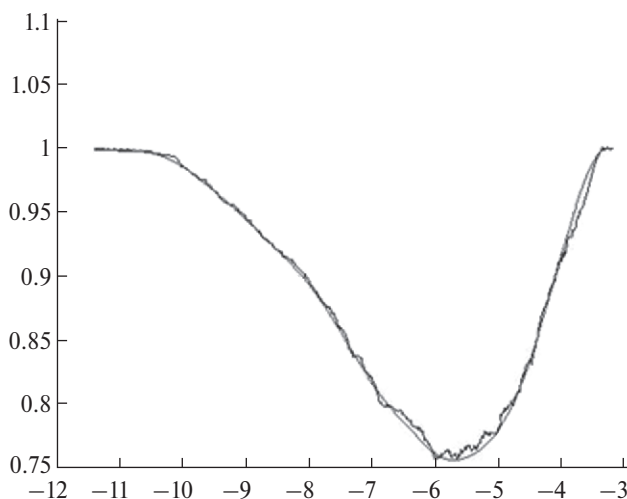


Рис. 4. Алеуты. ОР-окно. Графики суммарной ошибки $F_{rand}(W) + (1 - F_{real}(W))$. Исходная кривая и сглаженный вариант.

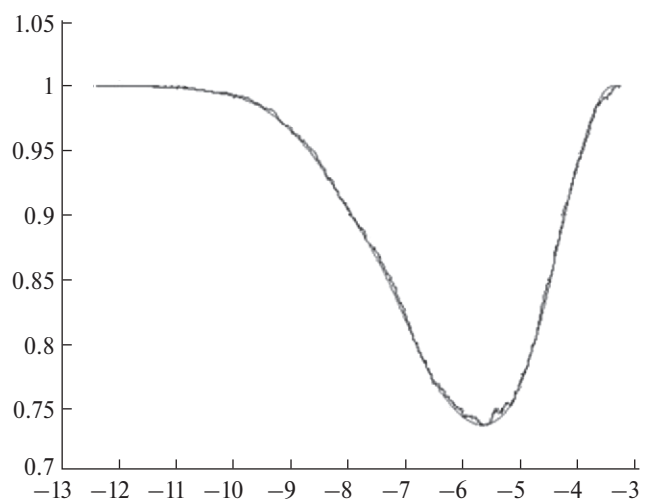


Рис. 5. Курилы. ОР-окно. Графики суммарной ошибки $F_{rand}(W) + (1 - F_{real}(W))$. Исходная кривая и сглаженный вариант.

Таблица 1

Регион, №	Метод NND		ОР-окно		Окно Гарднера– Кнопова		Окно Урхаммера	
	$\min F_{rand}(W) + (1-F_{real}(W))$	W_{min}	$\min F_{rand}(W) + (1-F_{real}(W))$	W_{min}	$\min F_{rand}(W) + (1-F_{real}(W))$	W_{min}	$\min F_{rand}(W) + (1-F_{real}(W))$	W_{min}
1	0.642	−6.44	0.623	−8.43	0.756	−0.357	0.738	−0.27
2	0.790	−5.57	0.806	−6.78	0.924	−0.014	0.935	−0.14
3	0.682	−4.43	0.698	−4.89	0.886	−0.068	0.870	−0.46
4	0.746	−4.51	0.813	−5.36	0.908	−0.015	0.911	−0.41
5	0.732	−5.29	0.732	−5.78	0.811	−0.087	0.855	−0.093
6	0.718	−4.89	0.711	−5.93	0.797	−0.230	0.796	−0.42
7	0.753	−4.83	0.849	−5.63	0.899	−0.107	0.907	−0.14
8	0.708	−4.98	0.680	−5.87	0.801	−0.217	0.798	−0.35
9	0.699	−5.19	0.703	−5.90	0.823	−0.141	0.799	−0.44
10	0.690	−4.64	0.705	−5.71	0.820	−0.111	0.817	−0.36
11	0.759	−4.04	0.720	−4.85	0.820	−0.216	0.817	−0.31
12	0.654	−4.23	0.765	−4.94	0.846	−0.179	0.838	−0.38
13	0.458	−4.45	0.437	−5.82	0.569	−0.291	0.611	−0.52
14	0.451	−4.43	0.427	−6.03	0.535	−0.215	0.539	−0.53
15	0.731	−6.27	0.747	−8.25	0.837	−0.100	0.883	−0.24
16	0.711	−6.02	0.669	−8.42	0.886	−0.061	0.917	−0.053
17	0.793	−6.19	0.712	−8.00	0.841	−0.106	0.868	−0.29

Таблица 2

Сейсмо тектонические зоны и номера регионов	Метод NND	ОР-окно	Окно Гарднера– Кнопова	Окно Урхаммера
	$\min F_{rand}(W) + (1-F_{real}(W))$	$\min F_{rand}(W) + (1-F_{real}(W))$	$\min F_{rand}(W) + (1-F_{real}(W))$	$\min F_{rand}(W) + (1-F_{real}(W))$
Зона Субдукции 1–12	0.710 ± 0.033	0.732 ± 0.070	0.842 ± 0.057	0.843 ± 0.063
Зона Континентов 13–14	0.581 ± 0.152	0.587 ± 0.180	0.693 ± 0.163	0.701 ± 0.149
Зона Рифтов 15–17	0.745 ± 0.042	0.703 ± 0.039	0.855 ± 0.087	0.889 ± 0.025
Все Зоны 1–17	0.686 ± 0.095	0.694 ± 0.113	0.809 ± 0.107	0.818 ± 0.106

Варьирование показателя d

Как мы упоминали, для разделения на кластеры связанных между собой «генетически» и несвязанных событий необходимо ввести какое-то обобщенное расстояние. *NND*-метод принимает это расстояние в виде (4), а ОР-метод в виде (11). Для показателя d обычно используют значение $d = 1.6$, ссылаясь при этом на корреляционную размерность поля эпицентров [Molchan, Kronrod, 2009]. Проверим, является ли такое задание показателя d оправданным. Если взять логарифм нормированного обобщенного расстояния (6)

$$\lg(t_i - t_k) + d \cdot \lg(r) - b \cdot m_k, \quad (12)$$

то мы получим показатель d как множитель при логарифме расстояния. Таким образом, этот показатель определяет вес, который пространственная компонента имеет по сравнению с временной. Чем больше d , тем больший вес мы придаем пространственной компоненте. В качестве критерия, который позволит выбрать оптимальное в определенном смысле значение d в задаче выделения кластеров, выберем минимальную суммарную вероятность ошибки, рассмотренную в предыдущем разделе. Обозначим ее через p :

$$p = \min[F_{rand}(W) + (1 - F_{real}(W))]. \quad (13)$$

Посчитаем с помощью ОР-окна вероятности p для значений $d = 0; 0.5; 1.0; 1.5; 2.0$ для всех 17 регионов, выбирая для каждого из них соответствующее значение b с теми же параметрами, которые мы использовали при составлении табл. 1, табл. 2. В качестве иллюстрации на рис. 6 показаны кривые суммарных вероятностей $F_{rand}(W) + (1 - F_{real}(W))$ для региона Японии.

Средние по 17 регионам значения p для разных значений d показаны в табл. 3.

Ради экономии места мы приводим только средние значения p , заметив при этом, что вариабельность значений p по регионам

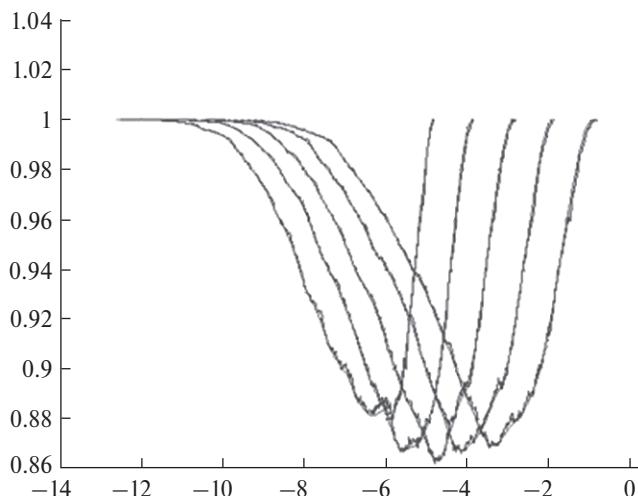


Рис. 6. Суммарная вероятность ошибки для региона Японии при разных d . Крайняя левая кривая соответствует $d = 0$. Затем вправо по порядку: $d = 0.5; 1.0; 1.5; 2.0$. Приведена исходная кривая (неровная) и сглаженный вариант. По оси X порог W .

довольно велика (среднее стандартное отклонение составляет 0.107). При этом для каждого из регионов при изменении d вероятность p изменяется менее существенно (среднее стандартное отклонение для вариаций p внутри региона составляет 0.088). Формально проинтерполированное среднее минимальное значение $p = 0.728$ соответствует размерности $d = 1.34$, что достаточно близко к принятому значению $d = 1.6$. При этом при других значениях d минимальные p отличаются в среднем не более, чем на 0.025, что не является значимым. Таким образом, при использовании критерия p выбор значения d в ОР-окне не существенен и можно использовать стандартное значение $d = 1.6$ без потери эффективности. Отметим также, что при использовании ОР-окна существенна подгонка параметров окна к данному региону (наклон графика Гутенберга–Рихтера b , порог W , соответствующий минимальному p , нижний порог магнитуд m_0).

Таблица 3. Минимальные вероятности p для ОР-окна

Коэффициент d	0	0.5	1.0	1.5	2.0
Минимальная вероятность $p \pm std$	0.735 ± 0.094	0.731 ± 0.102	0.729 ± 0.106	0.729 ± 0.114	0.735 ± 0.116

**Степень стационарности потока
главных толчков для разных методов
декластеризации**

Одно из основных требований, предъявляемых к различным методам декластеризации сейсмического потока, является требование стационарности последовательности главных толчков. Сравним по этому критерию стандартные окна Гарднера–Кнопова, Урхаммера и ОР-окно (9)–(11), а также метод *NND*. Для первых двух стандартный порог $W=0$, для ОР-окна и метода *NND* $W=-5$. Для идеального Пуассоновского стационарного потока распределение моментов событий на нормированном к единице интервале времени должно быть равномерным на отрезке $[0; 1]$. В качестве критерия стационарности мы будем использовать стандартный критерий, основанный на расстоянии Колмогорова KD . Если обозначить моменты времени главных толчков $(t_1, t_2, \dots, t_n)$, то расстояние KD будет иметь вид:

$$KD = \sqrt{n} \cdot \max |F_n(t) - t|, \quad (14)$$

где $F_n(t)$ — функция эмпирического распределения выборки $(t_1, t_2, \dots, t_n)$ и максимум берется по всем t . KD имеет стандартное распределение Колмогорова, из которого для конкретного расстояния KD_0 можно получить вероятность его превышения при выполнении гипотезы равномерного распределения моментов событий на заданном отрезке времени, которую будем обозначать pKD . Если pKD очень мало (расстояние Колмогорова очень велико), то гипотезу равномерности отвергают. Обычно в качестве порогового значения берут 0.1.

В табл. 4 приведены значения KD , pKD для 17 регионов, рассмотренных выше, а также две характеристики декластеризации: доля главных толчков Cm и доля одиночных толчков Cs среди всех главных толчков. В значениях pKD мы приводим две значащие цифры. Для окон использовались укороченные каталоги, с верхней границей 05.08.2005; для метода *NND* укороченный каталог начинался с 19.11.1992 г.

В табл. 5 приведены медианы значений pKD по всем 17 регионам для *NND*-метода и трех окон (медианы в данном случае являются более

Таблица 4

	Метод <i>NND</i>		ОР-окно		Окно Гарднера–Кнопова		Окно Урхаммера	
	KD/pKD	Cm/Cs	KD/pKD	Cm/Cs	KD/pKD	Cm/Cs	KD/pKD	Cm/Cs
1	0.94/0.35	0.38/0.64	1.02/0.25	0.49/0.75	1.23/0.097	0.76/0.80	1.30/0.070	0.82/0.88
2	0.94/0.35	0.51/0.72	0.84/0.48	0.65/0.81	1.18/0.13	0.57/0.71	0.97/0.30	0.70/0.84
3	0.76/0.61	0.60/0.82	0.86/0.45	0.66/0.83	0.86/0.45	0.44/0.66	1.32/0.0062	0.48/0.79
4	0.71/0.69	0.65/0.85	0.74/0.64	0.71/0.84	0.56/0.91	0.41/0.60	1.11/0.17	0.51/0.82
5	2.01/0.0001	0.47/0.83	2.98/0	0.56/0.88	2.95/0	0.48/0.73	1.45/0.030	0.35/0.83
6	0.64/0.80	0.64/0.77	0.69/0.73	0.70/0.81	0.74/0.65	0.53/0.68	0.59/0.87	0.62/0.80
7	1.32/0.061	0.49/0.83	1.30/0.068	0.53/0.86	1.36/0.050	0.42/0.68	2.82/0	0.33/0.85
8	0.82/0.52	0.55/0.79	1.10/0.18	0.62/0.81	0.61/0.86	0.47/0.70	0.92/0.37	0.51/0.79
9	0.83/0.49	0.84/0.89	0.78/0.57	0.86/0.92	0.67/0.76	0.66/0.79	0.60/0.86	0.81/0.91
10	1.27/0.079	0.68/0.82	1.55/0.016	0.73/0.85	1.11/0.18	56/0.75	1.39/0.043	0.65/0.84
11	1.06/0.22	0.83/0.86	0.92/0.33	0.87/0.91	0.64/0.80	0.66/0.75	0.81/0.53	0.77/0.88
12	2.47/0	0.62/0.86	2.76/0	0.67/0.89	1.75/0.0044	0.48/0.73	1.06/0.21	0.38/0.84
13	1.04/0.23	0.80/0.89	0.89/0.40	0.82/0.94	1.07/0.21	0.77/0.86	0.88/0.43	0.78/0.87
14	1.29/0.072	76/0.82	1.27/0.079	0.80/0.83	1.21/0.107	0.71/0.76	1.31/0.066	0.78/0.82
15	1.55/0.016	0.60/0.74	1.53/0.019	0.74/0.82	1.60/0.012	0.87/0.89	1.67/0.0075	0.93/0.93
16	0.52/0.95	0.78/0.84	0.62/0.83	0.84/0.88	0.92/0.37	0.91/0.91	1.15/0.14	0.95/0.95
17	0.76/0.61	0.74/0.79	0.82/0.54	0.83/0.86	0.96/0.32	0.84/0.86	1.19/0.12	0.92/0.94

Примечание: жирным шрифтом выделены значения $pKD < 0.1$, не удовлетворяющие требованию статистически равномерного распределения основных толчков по времени.

Таблица 5

	Метод <i>NND</i>	ОР-окно	Окно Гарднера— Кнопова	Окно Урхаммера
Медиана <i>pKD</i>	0.35	0.33	0.21	0.14
Средняя доля <i>Cm</i>	0.64	0.71	0.62	0.66
Средняя доля <i>Cs</i>	0.81	0.85	0.76	0.86

адекватной характеристикой среднего значения из-за большой асимметрии распределения). Там же приведены средние значения для долей главных толчков *Cm* (среди всех толчков) и долей одиночных толчков *Cs* среди главных толчков.

Из табл. 4, табл. 5 можно сделать следующие выводы. Метод *NND* и ОР-окно в целом обеспечивают режим главных толчков близкий к стационарному. Несколько хуже работает окно Гарднера—Кнопова. Для отдельных регионов и интервалов времени (в частности, практически всегда после гигантских землетрясений) возможны заметные отклонения от стационарности. Такие отклонения возникают при всех методах выделения афтершоков. Окно Урхаммера не обеспечивает достаточной стационарности потока главных толчков. Отметим, что ОР-окно оставляет большую долю событий в качестве главных толчков, не снижая показателя *pKD*. Это качество может оказаться привлекательным при решении некоторых задач, связанных с декластеризацией. Доля одиночных главных толчков *Cs* для всех рассмотренных методов довольно велика и составляет 75–85 %.

ОБСУЖДЕНИЕ

Для количественного измерения степени «генетической» связи толчков каталога (обобщенного расстояния между толчками) предлагаются разные модели [Gardner, Knopoff, 1974; Reasenber, 1985; Uhrhammer, 1986; Molchan, Dmitrieva, 1992; Zalyapin I and Ben-Zion, 2013; и др.]. В оконных вариантах решения задачи считается, что главный толчок каким-то способом определен (например, выбирается сильнейшее событие, которое по используемому определению не может быть афтершоком), а следующие за ним толчки проверяются с помощью соответствующего окна на принадлежность к его кластеру. При статистическом подходе в качестве обобщенного расстояния используют

монотонную функцию от правдоподобия (или от интенсивности точечного процесса). Естественно, что для получения функции правдоподобия нужна статистическая модель событий каталога в виде случайного точечного процесса. Процедура статистической проверки может включать в себя итерационные циклы и инкорпорировать известные сейсмологические закономерности (закон Гутенберга—Рихтера, закон Омори и др.).

В настоящей статье проводится сравнение традиционных оконных методов выделения афтершоков (окно Гарднера—Кнопова и окно Урхаммера) с недавно введенным методом *NND* и предлагаемом нами новым, основанном на обобщенном расстоянии между толчками ОР-окном. Показано также, что при расчете минимальных расстояний для уменьшения систематического смещения желательно использовать укороченный каталог. При этом для ОР-окна нужно укорачивать каталог с конца, т. е. использовать только главные толчки, не выходящие по времени за определенный предел, а для метода *NND* укорачивать каталог следует с его начала.

При сравнении эффективности методов с помощью случайной перестановки моментов толчков показано, что метод *NND* и ОР-окно имеют близкие показатели (суммарные вероятности ошибочного решения), в то время как окна Гарднера—Кнопова и Урхаммера заметно им уступают (табл. 2). Заметим, что при этом сравнении мы не использовали каких-либо порогов для промежутка времени и расстояния между толчками. Отметим также, что метод *NND* и ОР-окно используют некоторые региональные характеристики (адаптация к условиям конкретного региона), что дает им определенное преимущество.

При изменении параметра *d* в методе *NND* и ОР-окне мы не обнаружили существенного влияния этого параметра на эффективность различения реального и псевдослучайного каталогов в диапазоне $0 \leq d \leq 2$ (при том, что оптимальное значение $d = 1.34$ достаточно близко к обычно используемому значению $d = 1.6$). Отсюда заключаем, что можно пользоваться общепринятым значением $d = 1.6$ без потери эффективности.

Мы сравнили также рассмотренные 4 метода декластеризации по степени равномерности распределения по времени выделяемых данным методом главных толчков (стационарность потока главных толчков). В качестве критерия использовалось расстояние Колмогорова *KD* и соответствующая ему вероятность превышения *pKD*. Оказалось, что метод *NND* и ОР-окно приводят

к вполне приемлемым показателям стационарности (медиана значений pKD по всем 17 регионам равна 0.33–0.35), окно Гарднера–Кнопва дает несколько худший показатель: медиана $pKD = 0.21$, а окно Урхаммера приводит к потокам главных толчков с наибольшими отклонениями от стационарности: медиана $pKD = 0.14$; при этом в 7 регионах из 17 показатель pKD меньше 0.1, т. е., формально не удовлетворяют требованию равномерного распределения основных толчков по времени. Отметим, что для всех сравниваемых методов выделения афтершоков равномерность распределения основных толчков по времени нарушается в окрестности наиболее сильных событий ($M = 9.0+$ и близких к ним), как правило, сопровождаемых большим шлейфом афтершоков. Отметим, что выявленные нарушения стационарности могут порождаться как реальным нарушением режима фоновой сейсмичности в окрестности сильнейших землетрясений, так и издержками применения удобных оконных методов, заведомо отбраковывающих как афтершоки некоторое число основных событий. Последствия такой отбраковки будут особенно заметны для больших по площади зон влияния сильнейших землетрясений, происходящих при этом в областях с высоким уровнем фоновой сейсмичности.

Предлагаемый нами метод ОР-окна может оказаться практически удобным, ввиду простоты реализации и наибольшей близости результатов его применения к результатам наиболее логически обоснованного, но практически трудно реализуемого варианта NND -метода.

Примененный нами в настоящей работе метод случайной перестановки моментов толчков является полезным статистическим приемом, родственном известному бутстреп-методу, позволяющим анализировать эффективность выделения кластеров событий в сейсмическом потоке без дискуссионных априорных ограничений. При этом ОР-окно, за счет одновременного учета расстояний от главного события по пространству и по времени, а также учету локальных величин наклона графика повторяемости, представляется более гибким. В частности, это окно пропускает как афтершоки ранние, но довольно удаленные события; аргументы в пользу статистической связи таких событий с очагом главного землетрясения приводятся в работе [Freed, 2005; Felzer, Brodsky, 2007; Соболев, Закржевская, 2013; Lippiello et al., 2015].

Подавляющее большинство работ по афтершокам посвящены исследованиям статистических свойств афтершоковых кластеров [Баранов и др., 2018; Гульельми и др., 2014; 2015; 2017; Лутиков,

Родина 2013; Zalypin, Ben-Zion 2013; 2013a; Ogata, Zhuang, 2006] и вероятностному прогнозу афтершоковой активности [Баранов, Шебалин, 2016; 2017; 2018; 2018a]. При этом довольно часто авторы декларативно избегают сравнения качества разных алгоритмов выделения афтершоков (см., например, обобщающую работу [Stiphout et al., 2012]). В работе [Дещеревский и др., 2016] констатируется, что «анализ литературы показывает, что набор правил декластеризации, в том числе стохастической, в большей степени зависит от предпочтений исследователей, чем от каких-либо объективно определяемых признаков, обоснованных физическими моделями развивающегося процесса». Причина этого, по-видимому, в неопределенности понятия афтершока и отсутствия единого физического механизма возникновения афтершоков. Так, разными авторами афтершоки связываются с эффектом прохождения сейсмической волны от главного события, изменением напряженного состояния в результате этого события или с изменениями состояния некоторого объема геофизической среды. Отсюда афтершоки можно связать с эпизодами особенно быстрого изменения в состоянии геофизической среды, но тогда возникает вопрос, где проходит граница между этими более быстрыми и фоновыми изменениями. Заметим, что отмеченное выше нарушение стационарности потока основных событий в широкой окрестности сильнейших землетрясений также может быть связано с изменениями состояния среды при подготовке и реализации мегаземлетрясений.

Исследование физических аспектов афтершокового процесса — механизмов инициации афтершокового процесса и зависимости режима афтершоков от параметров теплового режима, флюидного режима и характера напряженного состояния литосферы — фактически только началось [Felzer, Brodsky, 2007; Hainzl et al., 2013; Lippiello et al., 2015; Shebalin, Narteau, 2017; Смирнов и др., 2019; и др.]. Мы здесь не касаемся этих вопросов. Отметим только, что выбор алгоритма выделения афтершоков является необходимым предварительным этапом исследования статистических закономерностей и физики афтершокового процесса. При этом выбор алгоритма влияет на статистические свойства исследуемой совокупности афтершоков и на физическую интерпретацию механизмов афтершокового процесса. В частности, при использовании предложенного выше алгоритма к афтершокам может быть отнесено определенное число ранних афтершоков, пространственно удаленных от основного землетрясения.

ФИНАНСИРОВАНИЕ РАБОТЫ

Исследования выполнены при частичной поддержке РФФИ, грант № 17-05-00351.

БЛАГОДАРНОСТИ

Авторы признательны П.Н. Шебалину, любезно ознакомившемуся со статьей и давшему рекомендации, способствовавшие ее улучшению.

ПРИЛОЖЕНИЕ

Оценка плотности вероятности ядерным методом (kernel method)

Пусть дана выборка $(x_1, \dots, x_n)$. Кernels-метод (ядерный метод) [Shawe, 2004], предлагает в качестве оценки плотности вероятности этой выборки функцию:

$$f^*(x) = \frac{1}{n} \sum_{k=1}^n \frac{1}{\sigma} g\left(\frac{x - x_k}{\sigma}\right), \quad (1A)$$

где: $g(x)$ — какая-нибудь симметричная, унимодальная плотность вероятности; σ — масштабный параметр. Мы предлагаем взять в качестве $g(x)$ стандартную, нормальную плотность:

$$g(x) = \frac{1}{2\pi} \exp\left(-\frac{x^2}{2}\right). \quad (2A)$$

Задача состоит в подходящем выборе масштабного параметра σ в зависимости от объема выборки n . При большом σ оценка (1A) может сильно сгладить плотность, а при малом σ в оценке могут появиться ложные пики. Для оптимального выбора масштабного параметра предположим, что неизвестная истинная плотность $f(x)$ является гладкой функцией и имеет вторую производную. Мы рассматриваем задачу асимптотически при $n \rightarrow \infty$, $\sigma \rightarrow 0$ и будем пренебрегать членами высшего порядка малости по $1/n$, σ . Математическое ожидание $f^*(x)$ равно:

$$\begin{aligned} Ef^*(x) &= \frac{1}{2\pi\sigma} \int f(y) \cdot \exp\left[-\frac{(x-y)^2}{2\sigma^2}\right] dy \cong \\ &\cong \frac{1}{2\pi\sigma} \int \left[f(x) + (y-x)f'(x) + \frac{1}{2}(y-x)^2 f''(x) \right] \times \\ &\times \exp\left[-\frac{(x-y)^2}{2\sigma^2}\right] dy = f(x) + \frac{\sigma^2}{2} f''(x). \end{aligned} \quad (3A)$$

Отсюда смещение $B(x)$ оценки $Ef^*(x)$ равно:

$$B(x) = Ef^*(x) - f(x) \cong \frac{\sigma^2}{2} f''(x). \quad (4A)$$

Точно так же, используя разложение в ряд Тейлора до членов второго порядка и отбрасывая члены более высокого порядка малости, получаем дисперсию оценки $f^*(x)$:

$$Var(f^*(x)) = \frac{1}{2\pi n \sigma} f(x) + 0(\sigma). \quad (5A)$$

Естественно требовать, чтобы квадрат смещения и дисперсия имели один и тот же порядок малости при $n \rightarrow \infty$, $\sigma \rightarrow 0$. В результате из уравнений (4A), (5A) получаем решение нашей задачи:

$$\sigma \sim C/n^{1/5}. \quad (6A)$$

Константу C целесообразно выбирать с учетом специфики конкретной задачи. Наш опыт привел к выбору $C \sim 0.3 \cdot std$, где std — стандартное отклонение выборки $(x_1, \dots, x_n)$.

СПИСОК ЛИТЕРАТУРЫ

- Баранов С.В., Шебалин П.Н., Дзедобоев Б.А. Закон повторности числа афтершоков // Докл. РАН. 2018. Т. 481. № 3.
- Баранов С.В., Шебалин П.Н. О прогнозировании афтершоковой активности 1. Адаптивные оценки на основе законов Омори и Гутенберга–Рихтера // Физика Земли. 2016. № 3. С. 82–101.
- Баранов С.В., Шебалин П.Н. О прогнозировании афтершоковой активности 2. Оценка области распространения сильных афтершоков // Физика Земли. 2017. № 2. С. 1–19.
- Баранов С.В., Шебалин П.Н. О прогнозировании афтершоковой активности 3. Динамический закон Бота // Физика Земли. 2018. № 6. С. 129–136.
- Баранов С.В., Павленко В.А., Шебалин П.Н. О прогнозировании афтершоковой активности. 4. Оценка максимальной магнитуды последующих афтершоков // Физика Земли. 2019. № 4. С. 15–32.
- Гульельми А.В., Завьялов А.Д., Зотов О.Д., Лавров И.П. Динамика афтершоков Суматра-Андаманского землетрясения // Физика Земли. 2014. № 1. С. 66–74.
- Гульельми А.В. Форшоки и афтершоки сильных землетрясений в свете теории катастроф // Успехи Физических наук. 2015. Т. 185. № 4. С. 415–429.
- Гульельми А.В., Завьялов А.Д., Зотов О.Д., Лавров И.П. Зависимость потока афтершоков от магнитуды главного удара // Физика Земли. 2017. № 1. С. 12–19.
- Дещеревский А.В., Мирзоев К.М., Лукк А.А. Критерии группирования землетрясений с учетом пространственной

- неоднородности сейсмичности // *Физика Земли*. 2016. № 1. С. 79–97.
- Кочарян Г.Г. Геомеханика разломов. М.: ГЕОС. 2016. 424 с.
- Лутиков А.И., Родина С.Н. Временные и энергетические характеристики афтершокового процесса на примере Курило-Камчатских землетрясений // *Геофизические исследования*. 2013. Т. 14. № 4. С. 23–35.
- Писаренко В.Ф., Родкин М.В. Распределения с тяжелыми хвостами: приложения к анализу катастроф // *Вычислительная сейсмология*. Вып. 38. М.: ГЕОС. 2007. 240 с.
- Смирнов В.Б., Пономарев А.В., Станчиц С.А., Потанина М.Г., Патонин А.В., Dresen G., Narteau C., Bernard P., Строганова С.М. Лабораторное моделирование афтершоковых последовательностей: зависимость параметров Омори и Гутенберга-Рихтера от напряжений // *Физика Земли*. 2019. № 1. С. 149–165.
- Соболев Г.А., Закржевская Н.А. К вопросу о влиянии удаленных землетрясений на сейсмичность // *Физика Земли*. 2013. № 4. С. 29–42.
- Baiesi M. and Paczuski M. Scale-free networks of earthquakes and aftershocks // *Phys. Rev.E*. 2004. V. 69. P. 66–106
- Console R., Jackson D.D., Kagan Y.Y. (2010) Using the ETAS Model for Catalog Declustering and Seismic Background Assessment // *Pure Appl. Geophys.* 2010. V. 167. P. 819–830.
- Efron B. Bootstrap Methods: Another Look at the Jackknife // *Ann. Statist.* 1979. V. 7. P. 1–26.
- Felzer K.R., Brodsky E.E. Decay of aftershock density with distance indicates triggering by dynamic stress // *Nature*. 2007. V. 441. 8 June 2006. doi:10.1038/nature04799
- Freed A.M. Earthquake triggering by static, dynamic, and postseismic stress transfer // *Annual Rev Earth and Planet Sci.* 2005. V. 33. P. 335–367.
- Gardner J.K., Knopoff L. Is the sequence of earthquakes in Southern California, with aftershocks removed, Poissonian? // *Bull. Seis. Soc. Am.* 1992. V. 64 (5). P. 1363–1367.
- Gomberg J. The failure of earthquake failure models // *J. Geophys. Res.* 2001. V. 106. P. 16253–16264.
- Hainzl S., Zakharchova O., Marsan D. Impact of Aseismic Transients on the Estimation of Aftershock Productivity Parameters // *BSSA*. 2013. V. 103 (3). P. 1723–1732. <https://doi.org/10.1785/0120120247>
- Lippiello E., Giacco F., Marzocchi W., Godano C. Mechanical origin of aftershocks // *Sci Rep.* 2015. V. 26;5. P. 15560. <https://doi.org/10.1038/srep15560>
- Molchan G., Dmitrieva O. Aftershock identification and new approaches // *Geophys.J. Int.* 192. V. 109. P. 501–516.
- Molchan G., Kronrod T. The fractal description of seismicity // *Geophys.J. Int.* 2009. V. 179. P. 1787–1799.
- Ogata Y., Zhuang J. Space–time ETAS models and an improved extension // *Tectonophysics*. 2006. V. 413. P. 13–23.
- Pisarenko V., Rodkin M. Heavy-Tailed Distributions in Disaster Analysis. *Advances in Natural and Technological Hazards Research*. Springer. Dordrecht-Heidelberg-London-New York. 2010. V. 30.
- Pisarenko V., Rodkin M. Statistical Analysis of Natural Disasters and Related Losses. *Springer Briefs in Earth Sciences*. Springer, Dordrecht-Heidelberg-London-New York. 2013.
- Pisarenko V.F., Rodkin M.V. New approach to estimation of probability of rare extreme events for the case of small samples: Application to seismic zoning problem. 17th International Conference on Geoinformatics – Theoretical and Applied Aspects, European Association of Geoscientists and Engineers (Kiev, Ukraine, 14–17 May 2017). 2018. doi: 10.3997/2214-4609.201801799
- Reasenber P. Second-order moment of central California seismicity, 1969–82 // *J. Geophys. Res.* 1985. V. 90. P. 5479–5495.
- Rodkin M.V., Kaftan V.I. Post-seismic relaxation from geodetic and seismic data // *Geodesy and Geodynamics*. 2017. V. 8. P. 13–16.
- Rodkin M., Pisarenko V. New approach to the general seismic zoning problem based on the theory of extreme values. Book of Abstracts of the 36 th General Assembly of the European Seismological Commission (2–7 September 2018, Valletta – Malta). ESC2018-S41–331. <https://drive.google.com/file/d/1PI3IMFHJuA1Wv5LHhNMpSw4I2yQm3nBk/view>
- Shawe-Taylor J., Cristiani C. Kernel-methods for Pattern Recognition? Cambridge University Press. 2004.
- Shebalin P., C. Narteau Depth dependent stress revealed by aftershocks. *NATURE COMMUNICATIONS*. 2017. doi: 10.1038/s41467-017-01446-y
- Stiphout T., Zhuang J., Marsan D. Seismicity declustering. Community Online Resource for Statistical Seismicity Analysis. 2012. Available at <http://www.corssa.org>. doi: 10.5078/corssa-52382934
- Uhrhammer R. Characteristics of Northern and Central California Seismicity // *Earthquake Notes*. 1986. V. 57 (1). P. 21.
- Zaliapin I., Gabrielov A., Keilis-Borok V., Wong H. Clustering Analysis of Seismicity and Aftershock Identification // *Phys. Rev. Lett.* 2008. V. 101 (1). P. 1–4.
- Zalyapin I., Ben-Zion Y. Earthquake clusters in Southern California I: Identification and stability // *Journ. Geophys. Res.* 2013. V. 118. P. 2847–2864.
- Zalyapin I., Ben-Zion Y. Earthquake clusters in Southern California I: Classification and relation to physical properties of the crust // *Journ. Geophys. Res.* 2013a. V. 118. P. 2865–2877.
- Zhuang J., Werner M.J., Hainzl S., Harte D., Zhou S. Basic models of seismicity: spatiotemporal models. Community Online Resource for Statistical Seismicity Analysis. 2011. doi: 10.5078/corssa-07487583. Available at <http://www.corssa.org>

Declustering of Seismicity Flow: Statistical Analysis

V. F. Pisarenko^a and M. V. Rodkin^{a,b,*}

^a*Institute of Earthquake Prediction Theory and Mathematical Geophysics, Russian Academy of Sciences, Moscow, 117997 Russia*

^b*Institute of Marine Geology and Geophysics, Far Eastern Branch, Russian Academy of Sciences,
Yuzhno-Sakhalinsk, 693022 Russia*

**e-mail: rodkin@mitp.ru*

Received December 25, 2018

The known methods for identifying the clusters of seismic events that are mainly formed by the aftershocks frequently include debatable initial assumptions or a complicated system of successive approximations. One of the most recent and the most logically consistent method for identifying the clusters of aftershocks is the nearest neighbor distance (NND) method which is however algorithmically most challenging. In this paper, we propose a new declustering method based on a generalized distance (GD) metric, which employs some assumptions of the NDD method but is as simple in practical implementation as the window methods previously proposed for this purpose. In analyzing and substantiating this new method, a procedure of random shuffling of seismic events with respect to time is used for generating a real catalog which is however free of genetic relations between different events. The efficiency of the existing window methods, the GD method, and the NND methods is compared by a number of the criteria for 17 regions. It is shown that the GD method is, on average, noticeably more efficient than the standard window methods and compares very favorably with the NND method. In the conclusions, a certain speculativeness of separating the events into main and dependent shocks is discussed.

Keywords: Algorithms of aftershock identification, random shuffling, generalized-distance window