

Оптимизация подхода к систематическому прогнозу землетрясений

В.Г. Гитис, А.Б. Дерендяев

Институт проблем передачи информации, Российская академия наук, Москва, Россия

Поступила в редколлегию 06.03.2024 г. Принята 15.04.2024 г.

Аннотация—Систематический прогноз землетрясений производится регулярно в заранее выбранной зоне анализа. На каждой итерации загружаются новые данные о сейсмическом процессе, данные обрабатываются, преобразуются в сеточные пространственно-временные поля, производится машинное обучение и даётся прогноз на постоянный временной интервал. Результатом является карта зоны тревоги, в которой ожидаются эпицентры всех целевых землетрясений на данном интервале прогноза. Для прогноза используется метод минимальной области тревоги. В новой версии систематического прогноза землетрясений решение проблемы разделено на два этапа. На первом этапе алгоритм выявляет интервалы тревоги, содержащие целевые землетрясения с эпицентрами в зоне анализа. На втором этапе на интервалах тревоги алгоритм прогнозирует зоны тревоги, содержащие все эпицентры целевых землетрясений. Это позволяет оптимизировать оценку вероятности того, что при очередном прогнозе все эпицентры целевых землетрясений попадут в прогнозируемую зону тревоги. Рассмотрены примеры применения метода для прогноза землетрясений на Камчатке, в Калифорнии и островной части Японии.

КЛЮЧЕВЫЕ СЛОВА: систематический прогноз землетрясений, машинное обучение, метод минимальной области тревоги, временные ряды GPS, качество прогноза.

DOI: 10.53921/18195822_2024_24_1_42

1. ВВЕДЕНИЕ

Сильные землетрясения вызывают большой социальный и экономический ущерб. Поэтому необходим мониторинг сейсмического процесса, анализ его динамики, своевременное выявление критических состояний и систематический прогноз сильных землетрясений. Решение этих задач требует анализа больших объемов пространственно-временных данных, которые наряду со значениями тематических свойств, содержат информацию о пространственной и временной привязке, что значительно усложняет их внутреннюю структуру. Это определяет необходимость разработки специальных моделей и методов извлечения эмпирических закономерностей из разнотипных пространственно-временных данных.

Задача прогноза землетрясений состоит в том, чтобы указать ограниченную по размеру зону тревоги, в которой в течение определенного времени ожидается эпицентр землетрясения целевой магнитудой. Представления о возможности прогноза землетрясений базируются на данных физического моделирования процесса разрушения горных пород и реальных наблюдениях, которые показывают, что землетрясению предшествуют процессы, формирующие аномальные изменения геологической среды в области очага ожидаемого землетрясения [1,2,3]. Так, например, в области подготовки сильного землетрясения зафиксированы аномальные отклонения в частоте и силе землетрясений, в деформациях земной поверхности, в химическом составе флюидов, в уровне грунтовых вод, во времени прохождения сейсмических волн, в значениях электрических и геомагнитных полей и др. [1,3,4,5,6,7]. Это свидетельствует в пользу

предположения о том, что землетрясение может быть предсказано по локальным изменениям геологической среды.

Работы по прогнозу землетрясений ведутся во многих направлениях. Они включают исследование механизмов разрушения горных пород, изучение различных предвестников землетрясений, разработку математических моделей и методов прогноза. Однако успешность прогноза затруднена недостаточной изученностью моделей сейсмичности и сложностью получения корректных инструментальных данных о поведении сейсмического процесса. Некоторые авторы полагают гипотетичной саму возможность решения проблемы прогноза, аргументируя тем, что решение задачи на данном этапе не реализуемо [8, 9, 10].

В настоящее время появились новые системы сейсмологических, геодезических, электромагнитных и атмосферных наблюдений. Разрабатываются новые методы машинного обучения предсказанию редких аномальных явлений. Накапливается статистика данных, относящихся к прогнозу землетрясений. Это даёт надежду на то, что разработка новых специализированных моделей и методов прогноза землетрясений позволит получить серьёзное продвижение в решении данной проблемы.

Для прогноза землетрясений широко используются методы машинного обучения [11, 12, 13, 14, 15, 16, 17, 18]. В ряде работ используются искусственные нейронные сети [19, 20, 21], а также их гибридные и рекуррентные модификации [22, 23]. Эти методы требуют для обучения достаточно больших выборок целевых событий. В то же время известно, что для ряда сейсмически активных регионов количество сильных землетрясений на интервалах обучения невелико. Поэтому, некоторые более простые модели могут иметь аналогичные или лучшие предсказательные возможности.

В статье мы развиваем подход к систематическому прогнозу землетрясений, представленный в [24, 25, 26, 27, 28, 29, 30, 31, 32]. В нём прогноз землетрясений даётся регулярно на заданный временной интервал. Подход основан на методе минимальной области тревоги. Основой метода является математическая модель, которая позволяет по ретроспективным данным находить предшествующие целевым землетрясениям аномалии в свойствах сейсмического процесса, сравнивать эти аномалии друг с другом и затем использовать аналогичные аномалии при прогнозе.

Качество прогноза землетрясений обычно оценивается по успешности обнаружения целевых событий в серии прогнозов. Оценка представляет собой долю обнаруженных событий от всех событий в серии. Эта оценка определяет предсказательную силу метода прогнозирования и позволяет оценить эффективность свойств сейсмического процесса (признаков), используемых для прогноза. Мы оптимизировали метод для улучшения другой характеристики качества решения: вероятности того, что при очередном прогнозе все эпицентры целевых землетрясений попадут в прогнозируемую зону тревоги. Этот показатель важен для практического применения систематического прогноза землетрясений.

Основные элементы подхода представлены в разделе 2. В разделе 3 представлены результаты моделирования подхода на примерах прогноза землетрясений в регионах Камчатки, Калифорнии и Японии.

2. БАЗОВЫЕ ЭЛЕМЕНТЫ ПОДХОДА

2.1. Оптимизация подхода

Систематический прогноз землетрясений дается регулярно и на постоянный временной интервал. На каждой итерации загружаются новые исходные данные о свойствах сейсмического процесса, дополняется выборка эпицентров целевых землетрясений, вычисляются однотипные пространственно-временные сеточные поля признаков прогноза, производится машинное

обучение прогнозирующей функции, вычисляется карта зон тревоги, в которых на данном интервале прогноза ожидаются эпицентры целевых землетрясений. Следует рассматривать два типа оценок качества прогноза. В первом случае качество решения определяется оценкой вероятности обнаружения эпицентров землетрясений в серии испытаний. Во втором случае оценкой качества является вероятность того, что на очередном временном интервале однократный прогноз окажется успешным.

Оценка вероятности успешного однократного прогноза землетрясений определяется отношением числа интервалов прогноза, в которых все эпицентры целевых землетрясений попадают в зону тревоги, к числу интервалов прогноза. Известно, что в ряде сейсмически активных регионов число сильных землетрясений с целевыми магнитудами весьма ограничено. При относительно небольших временных интервалах прогноза и небольшом числе целевых землетрясений данная оценка оказывается слишком малой для принятия практических решений об объявлении тревоги.

Улучшение качества однократного прогноза возможно, если вычислять зоны тревоги не на каждом шаге прогноза, а только при условии, что на выбранных интервалах прогноза ожидаются землетрясения в зоне анализа. Для этого прогноз осуществляется в два этапа. Первый этап состоит в определении интервала тревоги, на котором ожидается появление эпицентров целевых землетрясений в зоне анализа. Второй этап состоит в вычислении зоны тревоги при условии, что очередной интервал прогноза является интервалом тревоги. Это позволяет оптимизировать оценку условной вероятности того, что при однократном прогнозе на интервале тревоги все эпицентры целевых землетрясения попадут в зону тревоги. Таким образом, мы приходим к выводу о необходимости рассматривать два типа систематического прогноза землетрясений: прогноз по времени и прогноз по времени и пространству. При прогнозе по времени требуется предсказать интервал прогноза, содержащий целевые землетрясения. При пространственно-временном прогнозе требуется определить зону тревоги, в которой на данном интервале ожидаются все эпицентры целевых землетрясений.

2.2. Обобщение метода минимальной области тревоги

Модель обобщенного метода минимальной области тревоги состоит из следующих положений.

1. *Поля признаки прогноза.* Свойства процессов, связанных подготовкой сильных землетрясений, могут быть представлены пространственными и пространственно-временными сеточными полями.
2. *Данные машинного обучения.* Обучение использует все доступные к моменту прогноза значения сеточных полей и эпицентры землетрясений с целевыми магнитудами. Значения сеточных полей в узлах сетки соответствуют векторам многомерного пространства признаков.
3. *Условие аномальности.* Сильным землетрясениям предшествуют аномалии, при которых значения некоторых сеточных полей, близки к максимальным или минимальным. Этим аномалиям соответствуют векторы пространства признаков, которые называются предвестниками землетрясений.
4. *Цилиндр предвестника.* Узлы сетки, соответствующие возможным предвестникам землетрясений, находятся в предшествующем этому землетрясению цилиндре с центром основания в точке эпицентра.
5. *Условие монотонности.* Пусть известны поля признаков, у которых аномальные значения близки к минимальным, и поля, у которых аномальные значения близки к максимальным. Векторы пространства признаков, абсолютные значения компонент которых для каждого типа полей по абсолютной величине больше или равны компонентам вектора предвестника, также могут предшествовать аналогичным событиям.

Метод минимальной области тревоги использует для обучения известные к моменту прогноза данные: сейсмогенные сеточные пространственно-временные поля и эпицентры землетрясений с целевыми магнитудами. Обучение начинается с того, что для каждого эпицентра целевого землетрясения определяется учебная выборка предвестников. Предвестниками являются векторы значений полей в узлах координатной сетки, предшествующих эпицентру землетрясения. Векторы значений полей во всех остальных узлах координатной сетки маркировать не удаётся. В результате имеется выборка предвестников целевых землетрясений и множество немаркированных векторов, представляющих значения полей во всех остальных узлах координатной сетки. По этим данным требуется научиться обнаруживать близкое к наибольшему число эпицентров целевых землетрясений при условии, что число выделенных немаркированных векторов ограничено. Алгоритмы машинного обучения в данной постановке относятся к методам одноклассовой классификации [33, 34, 35].

Условие 3 предполагает, что предвестник является аномальным значением векторного поля, которое предшествует близкому в пространстве-времени эпицентру землетрясений с целевой магнитудой. Условие 4 уточняет область поиска предвестников землетрясений. Вместе с условием 3 оно является алгоритмической основой машинного обучения в методе минимальной области тревоги.

С каждым предвестником $\mathbf{f}^{(q)}$ землетрясения q связано множество векторов пространства признаков $h^{(q)}$, все компоненты которых по абсолютной величине больше или равны компонентам вектора $\mathbf{f}^{(q)}$. Это множество будем называть ортантом $h^{(q)}$ с вершиной в точке $\mathbf{f}^{(q)}$. Ортанту соответствует множество узлов координатной сетки $H^{(q)}$. Среди узлов сетки имеются те, которые являются другими предвестниками землетрясений, и те, которые относятся к немаркированным векторам. Согласно условию монотонности 5 все эти векторы также могут предшествовать землетрясению аналогичному q . Условие монотонности позволяет ввести меру информативности предвестника землетрясений. Под информативностью мы понимаем эффективность использования предвестника при обнаружении целевых событий. Мера информативности позволяет упорядочить все векторы пространства признаков. Однако это условие соответствует идеальной ситуации, в которой пространственно-временное поле признаков содержит полную информацию о процессах подготовки землетрясений. В реальной ситуации эта информация не полна. Поэтому для некоторых предвестников землетрясений условие монотонности не выполняется. Такие предвестники, как правило, имеют пониженную меру информативности.

Далее для упрощения объяснения метода минимальной области тревоги без ограничения общности будем предполагать, что аномальные компоненты предвестников землетрясений принимают только значения полей признаков прогноза, близкие к максимальным.

При обучении для каждого эпицентра учебной выборки вычисляется группа векторов предвестников землетрясения. Далее для каждого предвестника оценивается его мера информативности. В простейшем случае мера информативности предвестника определяется объёмом тревоги его ортанта. При пространственно-временном прогнозе объём тревоги ортанта равен доле числа узлов сетки, соответствующих векторам ортанта, от всех узлов сетки области анализа (в координатах пространство-время). Объём тревоги ортанта при прогнозе по времени равен доле числа интервалов тревоги от всех интервалов прогноза. Данная мера информативности предвестника тем больше, чем меньше объём тревоги ортанта.

При систематическом прогнозе землетрясений обучение выполняется на каждом шаге по всем данным от начала работы системы до момента очередного прогноза. Схема обучения прогнозу состоит из 7 операций.

1. *Определение учебной выборки предвестников целевых землетрясений.* Предвестниками целевого землетрясения являются все векторы пространства признаков, соответствующие зна-

чениям полей в узлах сетки из цилиндра предвестника с центром основания в эпицентре землетрясения, радиусом R и образующей T , где R и T являются параметрами алгоритма. Атрибутами предвестника землетрясений являются пространственно-временные координаты соответствующего узла сетки и события и его магнитуда.

2. *Вычисление ортантов предвестников.* Каждый предвестник является вершиной ортанта. Векторы, относящиеся к ортанту предвестника, согласно условию монотонности, также могут предшествовать аналогичным землетрясениям.
3. *Оценивание информативности предвестников.* Мера информативности предвестника оценивается по его ортанту. При пространственно-временном прогнозе она зависит от соотношения маркированных и немаркированных узлов координатной сетки, соответствующих векторам ортанта. При прогнозе по времени она зависит от соотношения числа интервалов, имеющих и не имеющих эпицентры землетрясений в зоне анализа.
4. *Упорядочивание предвестников по мере их информативности.*
5. *Вычисление прогнозирующей функции в пространстве признаков* (идея метода рассматривается ниже).
6. *Определение прогнозирующей функции в географических координатах.* Значения прогнозирующей функции переносятся на соответствующие интервалы прогноза при прогнозе по времени или на соответствующие узлы координатной сетки при пространственно-временном прогнозе.
7. *Вычисление порога прогнозирующей функции.* Порог определяет является ли данный интервал прогноза интервалом тревоги или попадает ли данный воксел координатной сетки в зону тревоги. При прогнозе по времени, в общем случае, порог выбирается по оценке минимума суммы ущерба от пропуска интервалов тревоги с целевыми событиями и ущерба от проведения антисейсмических мероприятий на интервалах тревоги. Порог при пространственно-временном прогнозе выбирается аналогично. Он определяется оценкой минимума суммы ущерба от пропуска целевых событий в зонах тревоги и ущерба от проведения антисейсмических мероприятий в зонах тревоги.

Схемы обучения метода минимальной области тревоги по времени и по времени-пространству принципиально совпадают. Отличия состоят в том, что при временном прогнозе векторы пространства признаков соответствуют временному ряду интервалов прогноза, а при пространственно-временном прогнозе векторы пространства признаков соответствуют узлам координатной сетки области анализа в координатах пространство-время.

Алгоритм вычисления прогнозирующей функции в обобщенном методе минимальной области тревоги (пункт 5) имеет три существенных отличия от работ [24, 25, 26, 27, 28, 29, 30, 31, 32]. Первое состоит в том, что решение задачи пространственно-временного прогноза землетрясений разбито на этапы: (1) прогноз интервалов тревоги, на которых ожидается появление землетрясений с эпицентрами в зоне анализа, и (2) прогноз зон тревоги, в которых на найденных интервалах тревоги ожидаются все эпицентры землетрясений. Второе отличие относится к выбору наиболее информативных предвестников землетрясений (пункт 4). В предыдущей версии для каждого эпицентра землетрясений выбирался один предвестник с наивысшей мерой информативности. Затем все эти предвестники упорядочивались по мере их информативности. Теперь упорядочивание предвестников по мере информативности производится для всех векторов, соответствующих узлам сетки цилиндров предвестников всех эпицентров землетрясений. Третье отличие относится к определению прогнозирующей функции в географических координатах (пункт 6). В предыдущей версии алгоритма обучения пространственно-временному прогнозу мы сглаживали зоны тревоги. В новой версии прогнозирующая функция в географических координатах строится только по узлам сетки, соответствующим векторам ортантов.

Рассмотрим алгоритм вычисления прогнозирующей функции для случая, когда для предвестника эпицентра землетрясения q мера информативности $G^{(q)}$ определяется по его ортанту $H^{(q)}$, как:

$$G^{(q)} = 1 - \nu^{(q)}, \quad (1)$$

где $\nu^{(q)}$ — объём тревоги ортанта $H^{(q)}$. При пространственно-временном прогнозе прогнозирующая функция $\Phi(\mathbf{f})$ представляет собой пространственно-временное поле, а при прогнозе по времени — временной ряд.

1. Предвестники упорядочиваются по мере информативности их ортантов.
2. Всем векторам пространства признаков присваивается значение 0.
3. Вычисляется прогнозирующая функция. Пусть узел сетки n соответствует вектору $f^{(n)}$ пространства признаков, который принадлежит подмножеству ортантов O . Тогда значение прогнозирующей функции на векторе $f^{(n)}$ определяется равным максимальной мере информативности ортантов множества O ,

$$\Phi(\mathbf{f}^{(n)}) = \max_{q \in O} G^{(q)} \quad (2)$$

4. Значения прогнозирующей функции переносятся при пространственно-временном прогнозе на узлы координатной сетки зоны анализа, соответствующие векторам пространства признаков, а при временном прогнозе на все интервалы прогноза.

Таким образом, к моменту прогноза алгоритм обучения установил принадлежность каждого вектора пространства признаков к ортантам и определил значения прогнозирующей функции: соответственно при пространственно-временном прогнозе на всех пикселах области анализа в координатах пространство-время, а при временном прогнозе на всех интервалах прогноза. Упорядоченность предвестников по мере информативности обеспечивает первым предвестникам наиболее значимое влияние на качество прогноза. Предвестники с малыми мерами информативности на результаты прогноза практически не влияют.

Для оценивания качества прогноза по результатам обучения вычисляется функции объёма тревоги $V(\mathbf{f}^{(n)})$. При пространственно-временном прогнозе функция объёма тревоги представляет пространственно-временное поле. Для любого порога V_0 определяется значение объёма тревоги. Оно равно доле числа узлов сетки со значениями объёма тревоги меньшими или равными V_0 от числа узлов сетки зоны анализа в координатах пространство-время. Для прогноза по времени значение объёма тревоги равно доле числа интервалов тревоги, выбранных по порогу V_0 , от числа всех интервалов прогноза.

Для вычисления функции объёма тревоги используется следующий алгоритм.

1. Предвестники упорядочиваются по мере информативности их ортантов.
2. Всем векторам пространства признаков присваивается значение 1.
3. Выбирается предвестник \mathbf{a} с наивысшей мерой информативности из всех предвестников учебной выборки. В нашем случае выбирается предвестник с наименьшим объёмом тревоги ортанта.
4. Векторам этого ортанта $H^{(a)}$, присваивается значение ψ_a , равное объёму области тревоги ортанта $H^{(a)}$, соответственно в контексте пространственно-временного или временного прогнозов. Затем выбирается предвестник \mathbf{b} с наибольшей мерой информативности среди оставшихся. Множеству векторов $H^{(a)} \setminus H^{(b)}$ присваивается значение ψ_b , равное объёму тревоги объединения ортантов $H^{(a)} \cup H^{(b)}$. Повторяя рассмотренную процедуру для остальных ортантов, алгоритм вычисляет функцию объёма тревоги на всех векторах пространства признаков.

5. Значения объёма тревоги функции переносятся при пространственно-временном прогнозе на узлы координатной сетки зоны анализа, соответствующие векторам пространства признаков, а при временном прогнозе на все интервалы прогноза.

Алгоритм принимает решение об объявлении тревоги на очередном интервале прогноза $(t, t + \Delta t)$ на основании сравнения значения объёма тревоги V с величиной порога V_0 . В результате при пространственно-временном прогнозе эпицентров землетрясений на интервале $(t, t + \Delta t)$ выделяется зона тревоги. Прогноз является успешным, если все эпицентры целевых землетрясений на этом интервале попадают в зону тревоги. При прогнозе по времени алгоритм принимает решение о том, относится ли интервал $(t, t + \Delta t)$ к интервалам тревоги или нет. Прогноз является успешным, если после прогноза на интервале тревоги происходят землетрясения, а на интервале без объявления тревоги землетрясения не происходят.

Рассмотрим алгоритм выбора значения порога принятия решения V_0 . Пусть функция $U(V)$ обозначает зависимость вероятности успешного прогноза землетрясений от величины объёма тревоги. Пусть при пространственно-временном прогнозе множество \mathcal{A} состоит из пикселей, для которых алгоритм относит к зоне тревоги, и множество \mathcal{B} состоит из пикселей, на которых нет тревоги. А для прогноза по времени аналогично пусть множество \mathcal{A} состоит из интервалов, которые алгоритм относит к интервалам тревоги $V \leq V_0$, и множество \mathcal{B} состоит из интервалов, на которых тревога не объявляется. В общем случае значение порога V_0 выбирается из условий минимума функции потерь:

$$V_0 = \mathit{Arg} \min_{\langle V \rangle} \{C_1[1 - U(V)] + C_2V\} = \mathit{Arg} \max_{\langle V \rangle} \{C_1U(V) - C_2V\}, \quad (3)$$

где: C_1 — ущерб при пропуске эпицентров целевых землетрясений в зонах тревоги при пространственно-временном прогнозе или на интервалах тревоги при прогнозе по времени, C_2 — ущерб от проведения антисейсмических мероприятий в зонах или на интервалах тревоги соответственно. Правило принятия решения по классификации интервалов прогноза имеет вид:

$$V \leq V_0 \rightarrow \text{интервал}(t^*, t^* + \Delta t] \ni \left\{ \begin{array}{l} \mathcal{A} \\ \mathcal{B} \end{array} \right\}. \quad (4)$$

3. МОДЕЛИРОВАНИЕ

Моделирование имитирует работу автоматической системы систематического прогноза землетрясений. На каждой итерации по всем известным ретроспективным данным вычисляется прогнозирующая функция и даётся прогноз на постоянный временной интервал. Прогноз выполняется в два этапа. На первом этапе принимается решение о том, ожидаются ли на очередном интервале прогноза целевые землетрясения с эпицентрами в зоне анализа. Если принято решение, что землетрясения ожидаются, то интервал прогноза называется интервалом тревоги. В этом случае выполняется второй этап прогноза. На этом этапе вычисляется зона тревоги, в которой на данном интервале тревоги ожидаются эпицентры всех целевых землетрясений.

Критериями качества серии прогнозов являются оценка вероятности обнаружения эпицентров землетрясений, оценка вероятности обнаружения интервалов тревоги, оценка вероятности обнаружения интервалов тревоги, для которых все эпицентры землетрясений находятся в зонах тревоги, и оценки ложного прогноза. Критериями качества однократного прогноз являются оценка вероятности того, что на интервале тревоги произойдут землетрясения с эпицентрами в зоне анализа и оценка вероятности того, что на интервале тревоги все эпицентры землетрясений попадут в зону тревоги.

Объём тревоги вычисляется для двух типов данных. В первом случае он определяется при обучении на интервале от начала обучения до момента прогноза. Поскольку с течением времени ретроспективные данные обучения непрерывно обновляются, то на каждой итерации мера информативности предвестников учебной выборки может изменяться. Кроме того, изменения могут происходить из-за появления новых предвестников землетрясений. Поэтому на каждой итерации обучения значения поля тревоги могут изменяться на всех временных срезах от начала обучения до момента тестирования t . Последний срез поля тревоги в момент t определяет прогноз на очередной интервал $(t, t + \Delta t)$. Оценки объёма тревоги, полученные алгоритмом обучения, мы будем обозначать буквой V . Каждому значению V соответствует оценка успешного прогноза U . Она равна отношению числа успешных прогнозов к числу всех прогнозируемых событий при условии, что прогнозный объём тревоги не превышает V . Результаты обучения показывает зависимость $U(V)$, которая соответствует ROC кривой. Вторым типом данных являются результаты прогноза, полученные на интервале тестирования. Объём тревоги по данным тестирования мы будем обозначать W . Отличие от V состоит в том, что V – это ожидаемый объём, определяемый по обучению, а W – реально наблюдаемый объём по тестированию.

Моделирование прогноза выполнялось для регионов Камчатки, Калифорнии и Японии. Во всех трёх случаях мерой информативности ортанта являлся его объём тревоги на данных обучения. При пространственно-временном прогнозе объём тревоги ортанта равен доле числа узлов сетки, соответствующих всем векторам ортанта, от числа узлов сетки зоны анализа в координатах пространство-время. При прогнозе по времени объём тревоги ортанта равен доле числа интервалов прогноза, соответствующих всем векторам ортанта, от числа всех интервалов прогноза. При выборе порога по условию (3) предполагается $C_1 = C_2$.

3.1. Прогноз землетрясений Камчатки

Прогноз землетрясений Камчатки выполняется по данным каталога землетрясений Камчатского филиала Геофизической службы Российской Академии наук, <http://sdis.emsd.ru/info/earthquakes/catalogue.php> [36]. Сеточные поля прогноза вычисляются по эпицентрам землетрясений с 1986 г. с магнитудами $m \geq 3.5$ и глубинами гипоцентров $H \leq 160$ км в координатной сетке $\Delta x \times \Delta y \times \Delta t = 0.1^\circ \times 0.75^\circ \times 30$ суток. Целевыми событиями являются эпицентры землетрясений с магнитудами $m \geq 6.0$ и глубинами гипоцентров $H \leq 60$ км. Обучение начинается с 1995 г. Тестирование выполняется от 08.01.2013 до 14.07.2023 с интервалом 30 суток. За время тестирования было сделано $N = 128$ прогнозов. Зона анализа и эпицентры тестируемых землетрясений показаны на Рис. 1.

Мы анализировали около 30 сеточных полей. Эти поля представляют пространственно-временные частотные и энергетические характеристики сейсмического процесса. Наиболее существенными для прогноза землетрясений являются следующие поля.

S_1 – поле плотности эпицентров землетрясений, вычисленное по каталогу землетрясений с помощью метода гауссовского ядерного сглаживания. Значения поля плотности эпицентров вычислялись с экспоненциальной ядерной функцией:

$$d(x, y, t) = \sum_n \left[\exp \left(- \left(\frac{r_n}{R_0} \right)^2 \right) \times \exp \left(- \left(\frac{t_n}{T_0} \right)^1 \right) \right] \times 1 \left(\varepsilon - \frac{r_n}{R_0} \right) \times 1 \left(\varepsilon - \frac{t_n}{T_0} \right) \quad (5)$$

где $1(u) = \begin{cases} 1, & u \geq 0 \\ 0, & u < 0 \end{cases}$, (x, y, t) – координаты узла 3D сетки, n – номер события, m_n – магнитуда n -го землетрясения, r_n [км] – расстояние от узла сетки до эпицентра, t_n [сутки] –

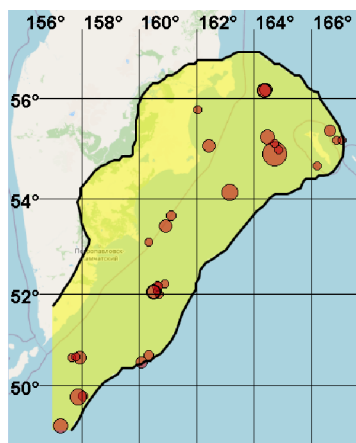


Рис. 1. Зона анализа и эпицентры тестовых землетрясений на интервале 08.01.2013 до 14.07.2023 с магнитудами $m \geq 6.0$ и глубинами эпицентров $H \leq 60$ км

временной интервал от узла сетки до эпицентра, $t - \varepsilon T \leq t_n \leq t$, радиус $R_0 = 50$ км и временной интервал $T_0 = 100$ суток – коэффициенты затухания, $\varepsilon = 2$ – пороговый коэффициент.

\mathbf{S}_2 – поле b-value, вычисленное аналогично полю \mathbf{S}_1 по каталогу землетрясений с помощью метода гауссовского ядерного сглаживания с экспоненциальной ядерной функцией с параметрами $R = 100$ км, $T = 365$ суток.

\mathbf{S}_3 – поле AWS плотности эпицентров, вычисленное с помощью метода AWS по полю \mathbf{S}_1 . Метод адаптивного весового сглаживания пространственных и пространственно-временных полей AWS (Adaptive Weight Smoothing или Propagation-Separation) предложен и исследован в работах [37, 38]. Метод ориентирован на сглаживание полей, которые включают в себя области с локально постоянными значениями. Отличие от стандартных методов состоит в том, что границы между локальными областями с различными постоянными значениями сохраняются, а качество сглаживания участков поля с постоянными значениями соответствует стандартным локально-ядерным методам. AWS обобщает методы локально-ядерного сглаживания, учитывая не только локальные особенности пространственно-временных данных, но и их статистическую структуру. Для учета локальных статистических свойств данных в методе вводится статистическая мера контраста, основанная на оценке максимума локально взвешенного правдоподобия. В работе [38] показано, что для достаточно широкого класса вероятностных моделей, принадлежащих семейству распределений экспоненциального типа, мерой контраста является расстояние (расхождение) Кульбака-Лейблера [39]. В [40, 41] метод обобщен для оценивания полей параметров сейсмического процесса по маркированным точечным полям, которыми являются, в частности каталоги землетрясений.

\mathbf{S}_4 – поле изменения AWS плотности эпицентров во времени. Значения поля $s_4(t)$ в момент t равно отношению разности средних значений AWS плотности $\bar{s}_3(t)$ и $\bar{s}_3(t - T_2)$ в двух последовательных интервалах $T_1 = 3000$ суток и $T_2 = 121$ суток к оценке среднеквадратичного отклонения этой разности $\sigma(s_3)$,

$$s_4(t) = (\bar{s}_3(t) - \bar{s}_3(t - T_2)) / \sigma(s_3) \quad (6)$$

где $\bar{s}_3(t)$ вычисляется по значениям поля \mathbf{S}_3 на интервале $(t - T_2, t)$, а $\bar{s}_3(t - T_2)$ – на интервале $(t - T_2 - T_1, t - T_2)$.

\mathbf{S}_5 – поле AWS b-value, вычисленное по полю \mathbf{S}_2 с помощью метода AWS.

$\mathbf{S}_6 = \mathbf{S}_4 \times \mathbf{S}_5$ – поле произведений значений полей \mathbf{S}_4 и \mathbf{S}_5 .

S_7 – поле площадей сечений. Значения поля в узлах сетки равны площадям сечений пространственных срезов поля S_1 горизонтальной плоскостью со значением плотности эпицентров $s_1 = 1$.

Пространственно-временной прогноз эпицентров землетрясений

Для прогноза использовались близкие к максимальным значения поля S_4 . Параметры цилиндра предвестника $R = 8$ км, $T = 91$ суток. Тревога в узле сетки продолжается $T = 91$ сутки.

Качество пространственно-временного прогноза определяют следующие показатели.

Оценка вероятности обнаружения эпицентров землетрясений

$$U = \frac{Q^*}{Q}, \quad (7)$$

где Q^* – число эпицентров целевых землетрясений, обнаруженных в серии тестовых прогнозов, $Q = 30$ – число тестовых эпицентров землетрясений.

Оценка вероятности обнаружения интервалов прогноза, на которых все эпицентры целевых землетрясений попадают в прогнозную зону тревоги.

$$U' = \frac{M'}{M}, \quad (8)$$

где M' – число интервалов прогноза, на которых все эпицентры целевых землетрясений попадают в прогнозную зону тревоги, $M = 17$ – число всех интервалов, в которых были эпицентры землетрясений в зоне анализа.

Оценка вероятности однократного прогноза, при котором на интервале прогноза все эпицентры целевых землетрясений попадают в прогнозную зону тревоги

$$P_1 = \frac{M'}{N}, \quad (9)$$

где $N = 128$ – число интервалов прогноза.

Результаты обнаружения эпицентров целевых землетрясений при величине порога объёма тревоги по данным обучения $\theta(V) = 0.2$ показаны в таблице 1. Эпицентр землетрясений не обнаружен, если значение объёма тревоги V на эпицентре землетрясения более 0.2. Ошибки прогноза в таблице выделены жирным шрифтом.

На Рис. 2 показаны графики (ROC кривые) зависимостей $U(V)$ и $U(W)$, где V и W обозначают объёмы тревоги, равные отношению числа узлов сетки пространственно-временной области тревоги к числу узлов области анализа по данным обучения тестирования. На графиках можно видеть, что 4 эпицентра землетрясений из 30 не обнаружены. Остальные 26 эпицентров землетрясений обнаруживаются при пороге $\theta(V) = 0.18$ и соответствующем пороге $\theta(W) = 0.31$. На Рис. 3 показаны графики (ROC кривые) зависимостей $U'(V)$ и $U'(W)$. На графиках можно видеть, что на 4-х интервалах прогноза в зону тревоги попали не все эпицентры землетрясений. Остальные интервалы со всеми эпицентрами землетрясений в зонах тревоги обнаруживаются при пороге $\theta(V) = 0.18$ и соответствующем пороге $\theta(W) \leq 0.32$. На Рис. 2 и 3 можно видеть существенные отличия зависимостей $U(V)$ и $U(W)$, как и зависимостей $U'(V)$ и $U'(W)$ по величинам объёмов тревоги V и W . Возможными причинами этих отличий могут являться относительно как малые учебные выборки эпицентров целевых землетрясений при обучении, так и изменения сейсмического процесса.

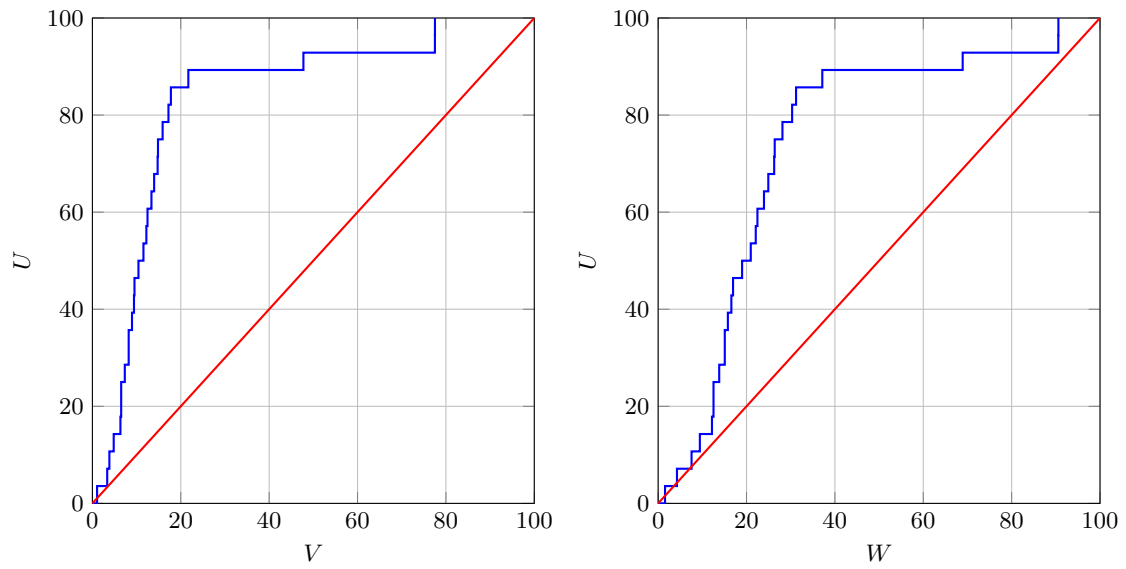


Рис. 2. Графики ROC кривых $U(V)$ (слева) и $U(W)$ (справа). Значения U , V и W даны в процентах.

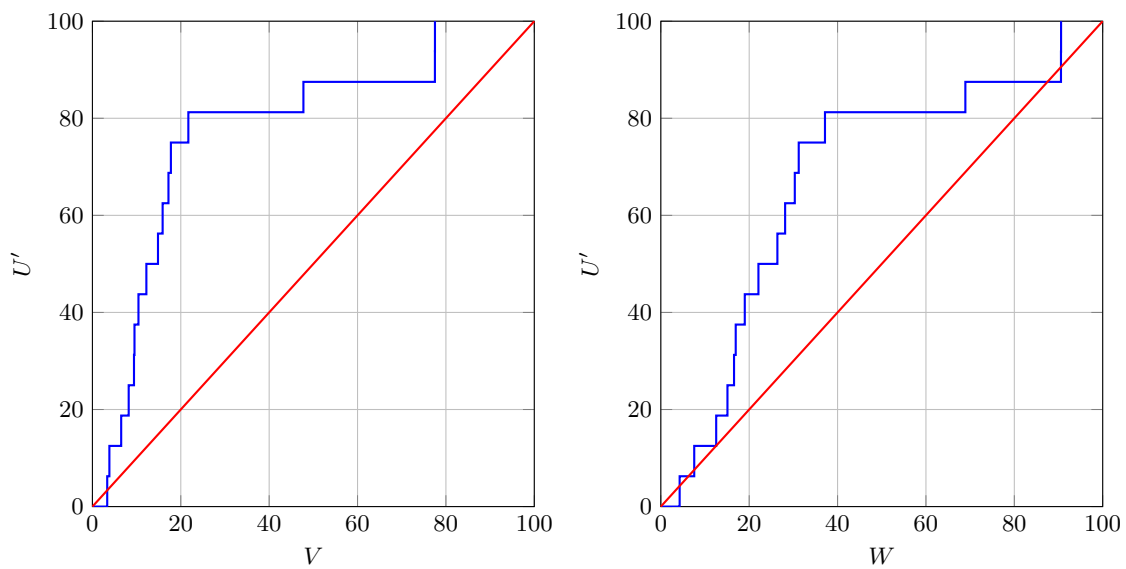


Рис. 3. Графики ROC кривых $U'(V)$ (слева) и $U'(W)$ (справа). Значения U' , V и W даны в процентах.

Таблица 1. Результаты пространственно-временного прогноза землетрясений Камчатки.

№	Дата	Долгота (градусы)	Широта (градусы)	Глубина (км)	Магнитуда	Объем тревоги %
1	2013/03/01	157.94	50.63	52.0	6.4	9.41
2	2013/03/04	157.66	50.63	51.0	6.1	8.96
3	2013/03/09	157.80	50.65	49.0	6.1	8.21
4	2013/03/24	160.33	50.68	58.0	6.2	8.21
5	2013/04/19	158.04	49.77	45.0	6.2	6.51
6	2013/04/20	157.88	49.74	39.0	6.7	6.51
7	2013/05/19	160.69	52.01	50.0	6.1	13.36
8	2013/05/19	160.65	52.08	42.0	6.0	13.97
9	2013/05/19	160.67	52.18	40.0	6.0	14.76
10	2013/05/21	160.89	52.22	59.0	6.1	15.89
11	2013/05/21	160.63	52.18	43.0	6.2	14.76
12	2013/05/21	160.49	52.05	48.0	6.5	12.47
13	2014/07/03	166.86	55.19	43.0	6.0	12.21
14	2014/07/03	167.06	55.18	40.0	6.0	11.55
15	2016/03/20	163.14	54.14	42.0	6.7	94.27
16	2016/04/14	161.11	53.66	48.0	6.2	14.84
17	2017/09/29	160.33	53.10	51.0	6.0	94.51
18	2018/01/25	166.65	55.37	46.0	6.3	3.35
19	2018/05/23	162.44	55.08	56.0	6.4	10.42
20	2018/10/10	157.26	49.09	41.0	6.6	94.61
21	2018/12/20	164.71	54.91	54.0	7.3	7.33
22	2018/12/20	164.85	54.99	54.0	6.0	4.82
23	2018/12/22	164.71	55.12	55.0	6.0	6.34
24	2018/12/24	164.46	55.25	51.0	6.6	9.52
25	2019/03/28	160.07	50.51	49.0	6.3	94.70
26	2019/06/25	164.41	56.18	57.0	6.4	1.06
27	2019/06/26	164.36	56.16	53.0	6.5	1.06
28	2019/08/09	162.04	55.78	60.0	6.0	3.84
29	2020/02/20	160.92	53.44	52.0	6.4	17.21
30	2020/04/04	166.21	54.67	37.0	6.0	17.76

Следуя графикам Рис. 2 и 3, выберем зоны тревоги при обучении со значениями $\theta(V) \leq 0.2$ и при тестировании со значениями $\theta(W) \leq 0.31$. Тогда качество прогноза характеризуется следующими показателями:

Оценка вероятности обнаружения эпицентров землетрясений

$$U = \frac{Q^*}{Q} = \frac{26}{30} = 0.87$$

Оценка вероятности обнаружения интервалов прогноза, на которых все эпицентры целевых землетрясений попадают в прогнозную зону тревоги.

$$U' = \frac{M'}{M} = \frac{13}{17} = 0.76$$

Оценка вероятности того, что на очередном интервале прогноза все эпицентры целевых землетрясений попадают в прогнозную зону тревоги.

$$P_1 = \frac{M'}{N} = \frac{13}{128} = 0.10$$

Прогноз интервалов тревоги

Функция, прогнозирующая интервалы тревоги, определяется по близким к максимальным значениям полей \mathbf{S}_3 и \mathbf{S}_5 . Параметрами цилиндра предвестника являются: радиус $R = 10$ км и образующая $T = 61$ суток. Тревога в узлах сетки продолжается $T = 61$ сутки.

Качество прогноза по времени определяют следующие показатели.

Оценка вероятности обнаружения интервалов тревоги, на которых эпицентры целевых землетрясений попадают в зону анализа.

$$U^* = \frac{M^*}{M}, \quad (10)$$

где M^* – число интервалов прогноза, на которых эпицентры целевых землетрясений попадают в зону анализа, $M = 17$ – число всех интервалов, в которых эпицентры землетрясений попадают в зону анализа.

При прогнозе по времени объём тревоги V^* вычисляется по данным от начала обучения до момента прогноза. На каждой итерации он равен доле интервалов тревоги $N^*(V^*)$ от всех интервалов $N_{\text{обучения}}$ на данный момент времени. Объём тревоги W^* вычисляется по данным тестирования. Он равен доле числа интервалов тревоги $N^*(W^*)$ от числа всех интервалов за время от начала тестирования до конца тестирования $N_{\text{тест}}$.

$$V^* = N^*(V^*)/N_{\text{обучения}} \quad (11)$$

$$W^* = N^*(W^*)/N_{\text{тест}} \quad (12)$$

Оценка вероятности однократного прогноза, при котором на интервале тревоги произойдут землетрясения в зоне анализа, определяется по результатам тестирования. Она равна отношению числа успешных прогнозов интервалов тревоги с эпицентрами землетрясений в зоне анализа M^* к числу всех интервалов тревоги за время тестирования

$$P_2(W^*) = \frac{M^*}{N^*(W^*)} \quad (13)$$

На Рис. 4 показаны графики (ROC кривые) зависимостей $U^*(V^*)$ и $U^*(W^*)$. Из графика $U^*(W^*)$ видно, что по результатам тестирования минимальная сумма оценки вероятности ошибки неверной классификации интервала тревоги (пропуск цели), равная $(1 - U^*)$, и объёма тревоги имеет место при величине объёма тревоги $W^* = 0.39$. Этому значению соответствует объём тревоги по обучению $V^* = 0.22$. Выбираем порог объёма тревоги $V^* = 0.22$. При этом в $M^* = 13$ интервалах тревоги из $M = 17$ имелись эпицентры целевых землетрясений в зоне анализа. Из (10) мы определяем, что для обнаружения $M^* = 13$ интервалов тревоги с землетрясениями алгоритм выбрал $N^*(W^*) = W^* \cdot N_{\text{тест}} = 0.39 \cdot 128 = 50$ интервалов тревоги из 128. Отсюда качество прогноза характеризуется следующими показателями:

Оценка вероятности обнаружения интервалов тревоги с эпицентрами землетрясений

$$U^* = \frac{M^*}{M} = \frac{13}{17} \approx 0.76$$

Оценка вероятности того, что на очередном интервале тревоги будут обнаружены эпицентры землетрясений в зоне анализа

$$P_2(W^*) = \frac{M^*}{N^*(W^*)} = \frac{13}{50} \approx 0.26$$

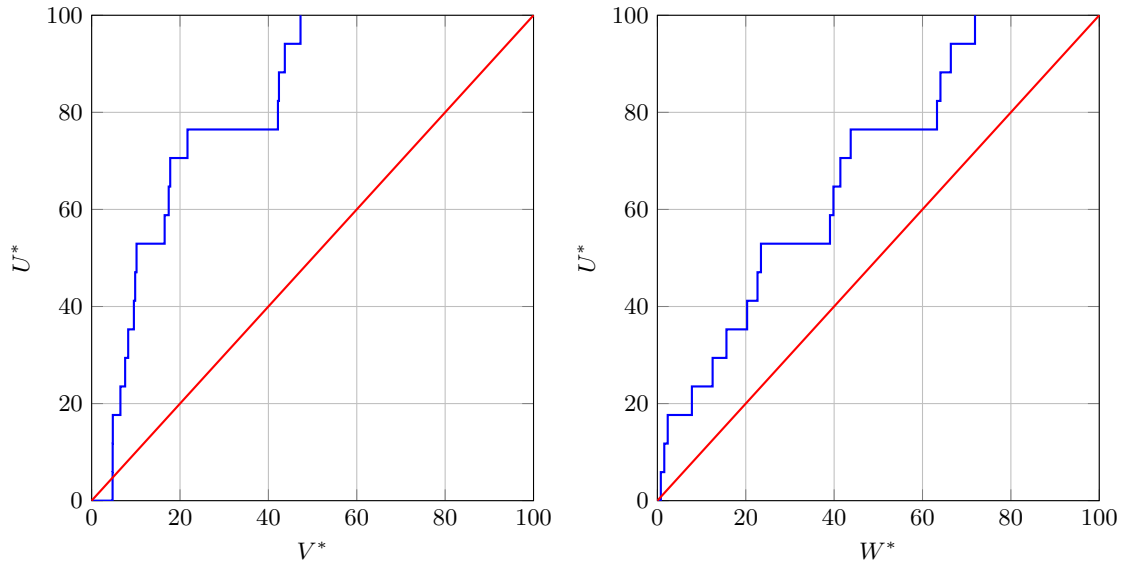


Рис. 4. Графики ROC кривых $U^*(V^*)$ (слева) и $U^*(W^*)$ (справа). Значения U^* , V^* и W^* даны в процентах.

Прогноз зон тревоги

На каждом интервале тревоги вычисляется зона тревоги. Для прогноза используются близкие к максимальным значения поля S_4 . Параметрами цилиндра предвестника являются радиус цилиндра $R = 8$ км и образующая $T = 91$ суток.

Оценка вероятности обнаружения интервалов тревоги, на которых все эпицентры целевых землетрясений попадают в зону тревоги имеет вид

$$U'' = \frac{M''}{M^*}, \tag{14}$$

где M'' — число интервалов при тестировании, на которых все эпицентры целевых землетрясений попадают в зону тревоги. Оценка вероятности однократного прогноза, при котором на интервале тревоги все эпицентры целевых землетрясений попадают в зоны тревоги с учетом (11), (12) имеет вид

$$P_3(W^*) = \frac{M''}{N^*(W^*)} \tag{15}$$

На Рис. 5 показана ROC кривая зависимости $U''(W)$. Можно видеть, что на $M^* = 13$ интервалах тревоги с землетрясениями в зоне анализа алгоритм обучения при пороге $\theta(V) = 0.18$ и при пороге $\theta(W) = 0.31$ построил $M'' = 11$ зон тревоги, в которые попали все эпицентры землетрясений. Таким образом, качество прогноза определяется следующими оценками:

$$U'' = \frac{M''}{M^*} = \frac{11}{13} \approx 0.85$$

$$P_3(W^*) = \frac{M''}{N^*(W^*)} = \frac{11}{50} \approx 0.22$$

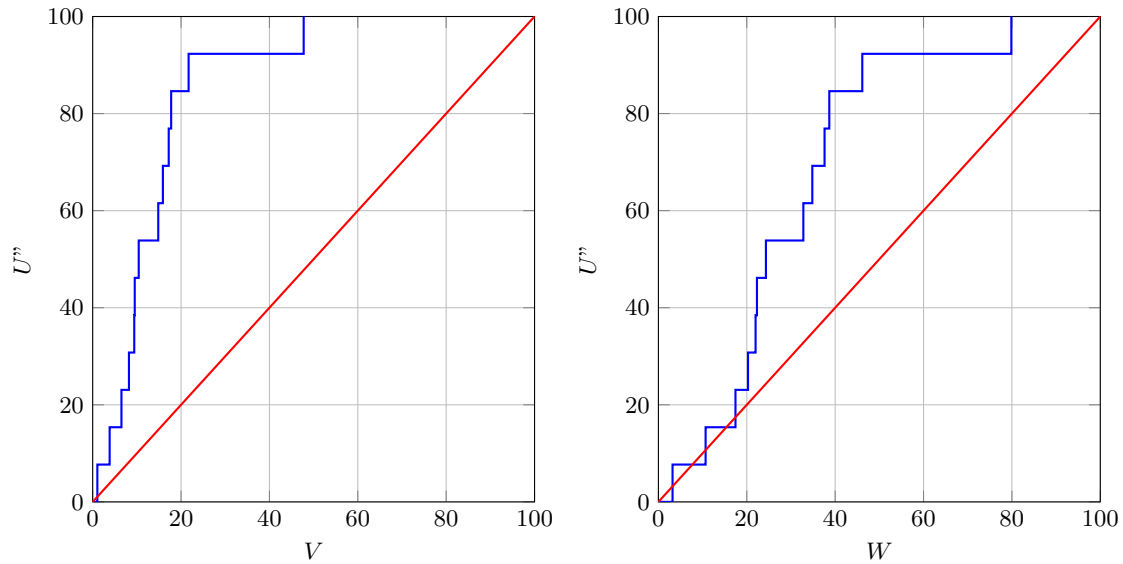


Рис. 5. Графики ROC кривых $U''(V)$ (слева) и $U''(W)$ (справа). Значения U'' , V и W даны в процентах.

3.2. Прогноз землетрясений Калифорнии

Для прогноза землетрясений Калифорнии анализировались два типа исходных данных: каталог землетрясений Национального центра данных (National Earthquake Information Center, NEIC) [42] на интервале 01.01.1995 – 23.09.2023 с магнитудами $m \geq 2.4$ и глубинами гипоцентров $H \leq 160$ км и временные ряды суточных горизонтальных смещений поверхности земли на интервале 01.01.2008 – 14.11.2023. Временные ряды получены из Nevada Geodetic Laboratory (NGL), <http://geodesy.unr.edu/about.php> [43]. Среднее расстояние между приемными станциями GPS составляет 9.38 км при стандартном отклонении 5.74 км. Сеточные поля прогноза вычисляются в координатной сетке $\Delta x \times \Delta y \times \Delta t = 0.1^\circ \times 0.75^\circ \times 30$ суток. Целевыми событиями являются эпицентры землетрясений с магнитудами $m \geq 5.3$ и глубинами гипоцентров $H \leq 60$ км. Обучение начинается с 07.08.2009 г. Тестирование выполняется от 08.01.2015 до 23.09.2023. Зона анализа и эпицентры тестируемых землетрясений показаны на Рис. 6. За время анализа в зоне анализа произошло 26 целевых землетрясений, из них 18 землетрясений произошли на интервале тестирования. Прогноз давался с интервалом 30 суток. На интервале тестирования было 107 прогнозов.

Мы анализировали около 40 сеточных полей. Часть из них вычислена по каталогу землетрясений и представляет пространственно-временные частотные и энергетические характеристики сейсмического процесса. Другая часть вычислена по временным рядам смещений поверхности Земли по данным GPS. Она представляет геодинамические процессы в земной коре. Наиболее существенными для прогноза землетрясений являются следующие поля.

\mathbf{S}_1 — поле плотности эпицентров землетрясений, вычисленное по каталогу землетрясений с помощью метода гауссовского ядерного сглаживания. Значения поля плотности эпицентров вычислялись с экспоненциальной ядерной функцией (5). Параметрами поля плотности эпицентров являются: радиус $R_0 = 50$ км, временной интервал $T_0 = 100$ суток, пороговый коэффициент $\epsilon = 2$.

\mathbf{S}_2 — поле значений b-value, вычисленное аналогично полю \mathbf{S}_1 по каталогу землетрясений с помощью метода гауссовского ядерного сглаживания с экспоненциальной ядерной функцией с параметрами $R = 150$ км, $T = 365$ суток.

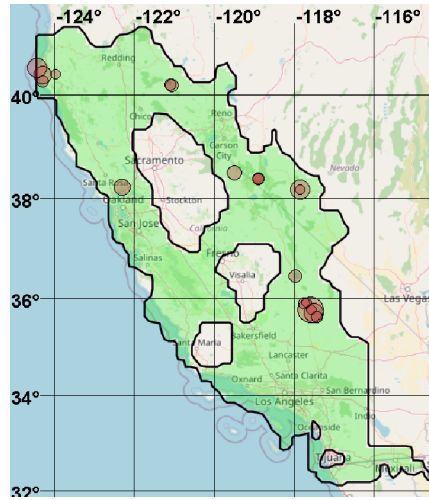


Рис. 6. Зона анализа и эпицентры тестовых землетрясений на интервале 08.01.2015 до 23.09.2023 с магнитудами $m \geq 5.3$ и глубинами эпицентров $H \leq 60$ км

S_3 — поле изменения плотности эпицентров во времени. Как и в (6) значения поля $s_3(t)$ в момент t равны отношению разности средних значений $\overline{s_1}(t)$ и $\overline{s_1}(t - T_2)$ в двух последовательных интервалах $T_1 = 2300$ суток и $T_2 = 121$ суток к оценке среднеквадратичного отклонения этой разности.

S_4 — поле изменения b-value во времени. Как и в (6) значения поля $s_4(t)$ в момент t равны отношению разности средних значений $\overline{s_2}(t)$ и $\overline{s_2}(t - T_2)$ в двух последовательных интервалах $T_1 = 2300$ суток и $T_2 = 121$ суток к оценке среднеквадратичного отклонения этой разности.

Метод использования временных рядов GPS для прогноза землетрясений подробно рассмотрен в [30]. Сеточные поля признаков прогноза вычисляются в два шага. Сначала выделяется полезный сигнал из временных рядов приемных станций GPS, а затем вычисляются пространственно-временные поля признаков прогноза. Мы предполагаем, что предвестники сильных землетрясений могут быть найдены в пространственно-временных полях аномальных изменений режима деформаций поверхности земли. Аномалии во времени вычислялись в трех полях инвариантов скорости деформаций.

F_1 — поле дивергенции скорости деформаций. Значение поля в точке с координатами (x, y) определяет относительные сокращения или расширения размеров малой горизонтальной площадки, расположенной в этой точке,

$$\operatorname{div}V_n = \frac{\partial V_{xn}}{\partial x} + \frac{\partial V_{yn}}{\partial y} \tag{16}$$

F_2 — сглаженное поле изменений дивергенции скорости деформаций. Значения поля $f_{2n}(t)$ равны пространственно сглаженным с радиусом $R = 20$ км значениям отношений разности средних значений дивергенции $\overline{div_{2n}} - \overline{div_{1n}}$ в двух последовательных интервалах T_1 и T_2 к стандартному отклонению этой разности $\sigma_n(div)$, $T_1 = T_2 = 360$ суток,

$$f_{2n}(t) = \frac{\overline{div_{2n}} - \overline{div_{1n}}}{\sigma_n(div)}, \tag{17}$$

где div_{2n} вычисляется по значениям поля F_1 на интервале $(t - T_2, t)$, а div_{1n} вычисляется на интервале $(t - T_2 - T_1, t - T_2)$.

\mathbf{F}_3 — поле ротора скорости деформаций. Значение поля в точке с координатами (x, y) определяет направление и интенсивность скручивания поля вокруг вертикальной оси, расположенной в этой точке

$$\text{rot}V_n = \frac{\partial V_{yn}}{\partial x} - \frac{\partial V_{xn}}{\partial y} \quad (18)$$

\mathbf{F}_4 — сглаженное поле изменений ротора скорости деформаций. Значения поля $f_{4n}(t)$ аналогичны значениям поля F_2 . Они равны пространственно сглаженным с радиусом $R = 20$ км значениям отношений разности средних значений дивергенции $\overline{\text{rot}_{2n}} - \overline{\text{rot}_{1n}}$ в двух последовательных интервалах T_1 и T_2 к стандартному отклонению этой разности $\sigma_n(\text{rot})$, $T_1 = T_2 = 360$ суток,

\mathbf{F}_5 — поле пространственных корреляций значений полей изменений дивергенции \mathbf{F}_2 и изменений ротора и \mathbf{F}_4 скорости деформаций в окне 150×150 км.

Пространственно-временной прогноз эпицентров землетрясений

Для прогноза использовались близкие к максимальным значения полей \mathbf{F}_2 , \mathbf{S}_1 и близкие к минимальным значения поля \mathbf{S}_4 . Параметры цилиндра предвестника составляют $R = 8$ км, $T = 31$ суток. Тревога в узле сетки продолжается $T = 31$ сутки.

Результаты обнаружения эпицентров целевых землетрясений при величине порога объёма тревоги $\theta(V) = 0.2$ показаны в таблице 2. Эпицентр землетрясений не обнаружен, если значение объёма тревоги V на эпицентре землетрясения более 0.2. Ошибки прогноза выделены жирным шрифтом.

Таблица 2. Результаты пространственно-временного прогноза землетрясений Калифорнии.

№	Дата	Долгота (градусы)	Широта (градусы)	Глубина (км)	Магнитуда	Объем тревоги %
1	2016/12/28	-118.89890	38.37550	11.3	5.6	4.72
2	2016/12/28	-118.89720	38.39040	12.2	5.6	4.72
3	2016/12/28	-118.89570	38.37770	8.8	5.5	4.72
4	2019/06/23	-124.30033	40.27350	9.4	5.6	6.12
5	2019/07/04	-117.50383	35.70533	10.5	6.4	2.40
6	2019/07/05	-117.57500	35.76033	7.0	5.4	2.40
7	2019/07/06	-117.59933	35.76950	8.0	7.1	2.40
8	2019/07/06	-117.74950	35.90117	5.0	5.5	2.40
9	2019/07/06	-117.68483	35.91017	7.4	5.4	2.40
10	2020/05/15	-117.84970	38.16890	2.7	6.5	100.00
11	2020/06/04	-117.42817	35.61483	8.4	5.5	0.45
12	2020/06/24	-117.97517	36.44683	4.7	5.8	3.09
13	2020/11/13	-117.85330	38.16930	4.8	5.3	0.59
14	2021/07/08	-119.49983	38.50750	7.5	6.0	40.66
15	2021/12/20	-124.29800	40.39017	27.0	6.2	8.63
16	2022/12/20	-124.42300	40.52500	17.9	6.4	0.56
17	2023/01/01	-123.97100	40.40900	30.6	5.3	0.57
18	2023/05/11	-121.10950	40.20417	5.8	5.5	15.00

На Рис. 7 показаны графики зависимостей $U(V)$ и $U(W)$, где V и W обозначают объёмы тревоги, равные отношениям числа узлов сетки пространственно-временной области тревоги к числу узлов области анализа по данным обучения тестирования. На графиках можно видеть, что 2 эпицентра землетрясений не обнаружены. Остальные 16 эпицентров землетрясений обнаруживаются при пороге $\theta(V) = 0.15$ и при пороге $\theta(W) = 0.22$. На Рис. 8 показаны графики

(ROC кривые) зависимостей $U'(V)$ и $U'(W)$. На графиках можно видеть, что на 2-х интервалах прогноза в зону тревоги попали не все эпицентры землетрясений. Остальные 8 интервалов со всеми эпицентрами землетрясений в зонах тревоги обнаруживаются при пороге $\theta(V) = 0.15$ и при пороге $\theta(W) = 0.22$. На Рис. 7 и 8 можно видеть, что зависимости $U(V)$ и $U(W)$, как и зависимости $U'(V)$ и $U'(W)$, в отличие от примера по региону Камчатки, мало отличаются друг от друга по величинам объёмов тревоги V и W .

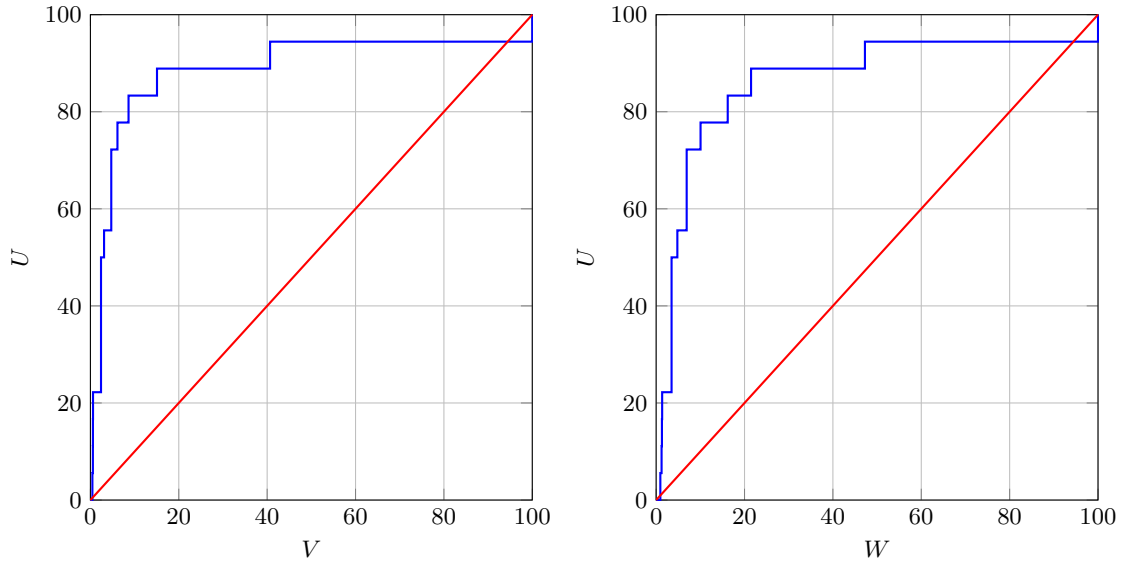


Рис. 7. Графики ROC кривых $U(V)$ (слева) и $U(W)$ (справа). Значения U , V и W даны в процентах.

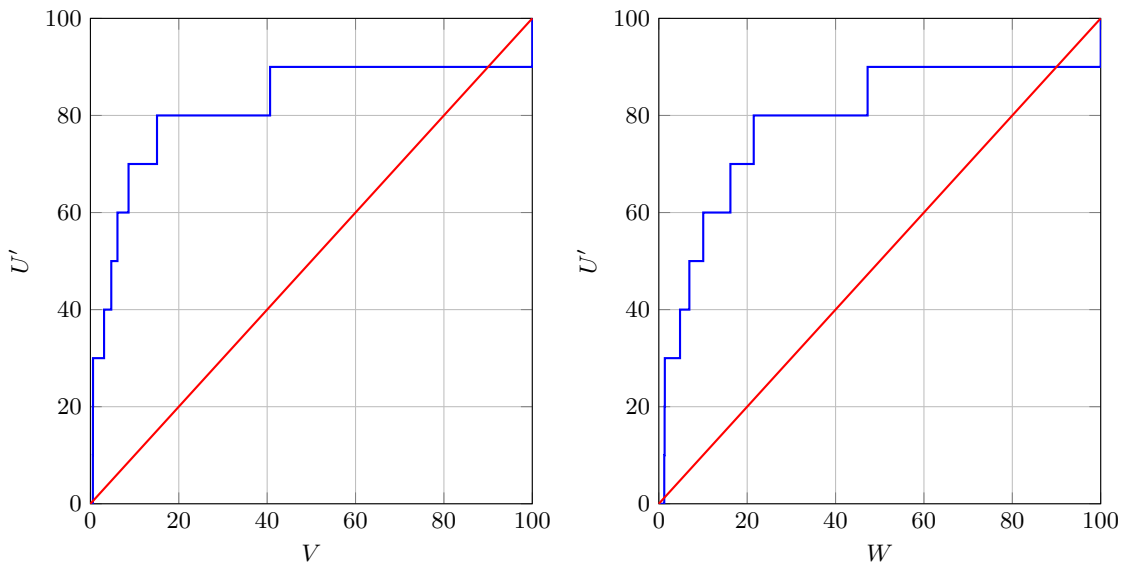


Рис. 8. Графики ROC кривых $U'(V)$ (слева) и $U'(W)$ (справа). Значения U' , V и W даны в процентах.

Следуя графикам Рис. 7 и 8, выберем зоны тревоги при прогнозе со значениями $V \leq 0.15$ и при тестировании со значениями $W \leq 0.22$. Тогда качество прогноза согласно (7), (8), (9) характеризуется следующими показателями:

Оценка вероятности обнаружения эпицентров землетрясений

$$U = Q^*/Q = 16/18 \approx 0.89.$$

Оценка вероятности обнаружения интервалов прогноза, на которых все эпицентры целевых землетрясений попадают в прогнозную зону тревоги.

$$U' = M'/M = 8/10 \approx 0.80.$$

Оценка вероятности того, что на очередном интервале прогноза все эпицентры целевых землетрясений попадают в прогнозную зону тревоги.

$$P_1 = M'/N = 8/107 \approx 0.075.$$

Прогноз интервалов тревоги

Для прогноза интервалов тревоги использовались значения полей \mathbf{S}_2 и \mathbf{F}_2 , близкие к максимальным, и значения полей \mathbf{S}_4 и \mathbf{F}_5 , близкие к минимальным. Параметры цилиндра предвестника составляют $R = 6$ км, $T = 31$ суток. Тревога в узле сетки продолжается $T = 31$ сутки.

На Рис. 9 показаны графики зависимостей $U^*(V^*)$ и $U^*(W^*)$. Из графика $U^*(V^*)$ видно, что по результатам тестирования минимальная сумма оценки вероятности ошибки неверной классификации интервала тревоги (пропуск цели), равная $(1 - U^*)$, и объёма тревоги имеет место при величине объёма тревоги $W^* = 0.29$. При этом в $M^* = 8$ интервалах тревоги из $M = 10$ имелись эпицентры целевых землетрясений в зоне анализа. Из (12) мы определяем, что для обнаружения $M^* = 8$ интервалов тревоги с землетрясениями алгоритм выбрал $N^*(W^*) = W^* \cdot N_{\text{тест}} = 0.29 \cdot 107 = 30$ интервалов тревоги из 107. Отсюда, следуя (10) и (13), получаем

$$U^* = M^*/M = 8/10 = 0.80.$$

$$P_2(W^*) = M^*/N^*(W^*) = 8/30 \approx 0.27.$$

Прогноз зон тревоги

Зоны тревоги вычисляются только на интервалах тревоги. Как и в случае пространственно-временного прогноза мы использовали значения полей \mathbf{F}_2 , \mathbf{S}_1 , близкие к максимальным, и значения поля \mathbf{S}_4 , близкие к минимальным. Параметры цилиндра предвестника составляют $R = 8$ км, $T = 61$ суток. Тревога в узле сетки продолжается $T = 61$ сутки.

На Рис. 10 показаны ROC кривые зависимостей $U''(V)$ и $U''(W)$. Можно видеть, что алгоритм вычислил прогнозные зоны тревоги, в которые при пороге $\theta(V) = 0.15$ и при пороге $\theta(W) = 0.32$ попали все эпицентры землетрясений на $M'' = 6$ интервалах тревоги из $M^* = 8$. Таким образом, качество прогноза определяется следующими оценками:

Из (14) оценка вероятности обнаружения интервалов тревоги, на которых все эпицентры целевых землетрясений попадают в зону тревоги

$$U'' = M''/M^* = 6/8 = 0.75.$$

Оценка вероятности однократного прогноза, при котором на интервале тревоги все эпицентры целевых землетрясений попадают в зоны тревоги (15), имеет вид

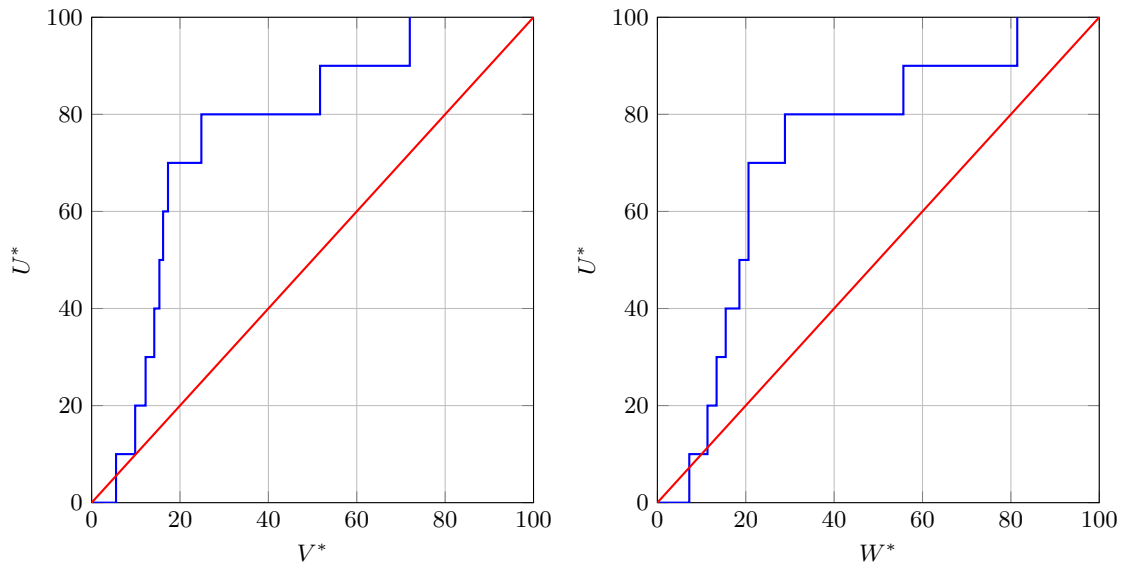


Рис. 9. Графики ROC кривых $U^*(V^*)$ (слева) и $U^*(W^*)$ (справа). Значения U^* , V^* и W^* даны в процентах.

$$P_3(W^*) = M''/N^*(W^*) = 6/32 \approx 0.19.$$

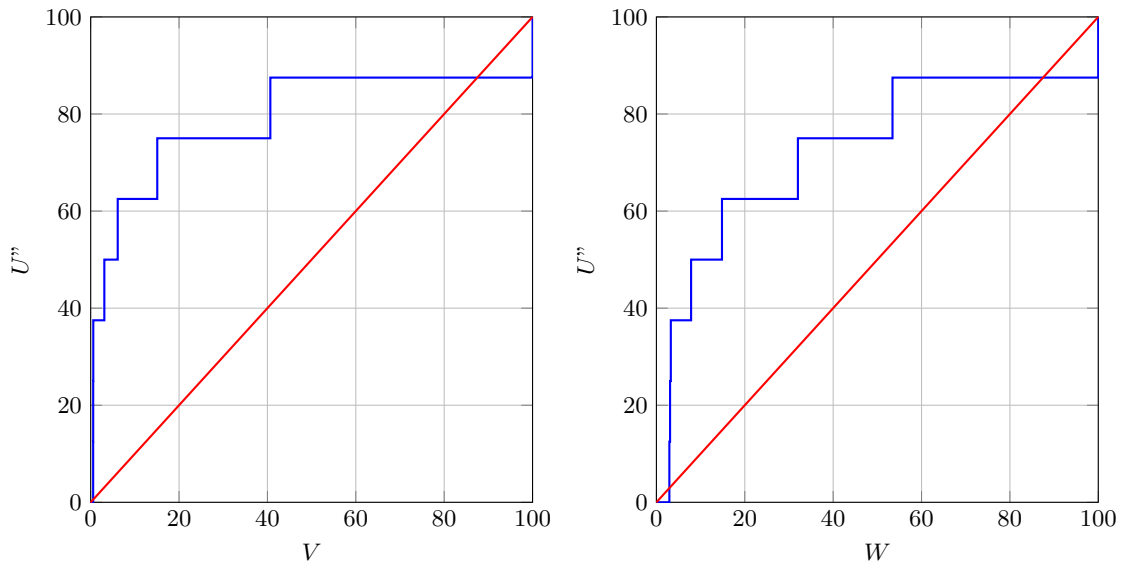


Рис. 10. Графики ROC кривых $U''(V)$ (слева) и $U''(W)$ (справа). Значения U'' , V и W даны в процентах.

3.3. Прогноз землетрясений островной части Японии

Для прогноза землетрясений островной части Японии мы анализировали данные каталога землетрясений и суточные горизонтальные смещения поверхности земли по данным космической геодезии. Сейсмологические данные представлены землетрясениями на интервале 02.06.2002 – 23.09.2023 с магнитудами $m \geq 2.0$, с глубинами гипоцентров $H \leq 160$ км, взятыми из каталога Japan Meteorological Agency [44, 45]. Смещения поверхности земли представлены временными рядами GPS на интервале 01.01.2008–14.11.2023, полученными из Nevada

Geodetic Laboratory (NGL), <http://geodesy.unr.edu/about.php> [43]. Средние минимальные расстояния между приёмными станциями GPS составляют 12,8 км, стандартное отклонения этих значений составляет 5,4 км. Обучение начинается с 02.03.2011 г. Тестирование выполняется от 10.04.2015 до 28.04.2023 с интервалом 30 суток. Зона анализа и эпицентры тестируемых землетрясений показаны на Рис. 11. За время тестирования в зоне анализа произошло 20 целевых землетрясений. На интервале тестирования было 99 прогнозов.

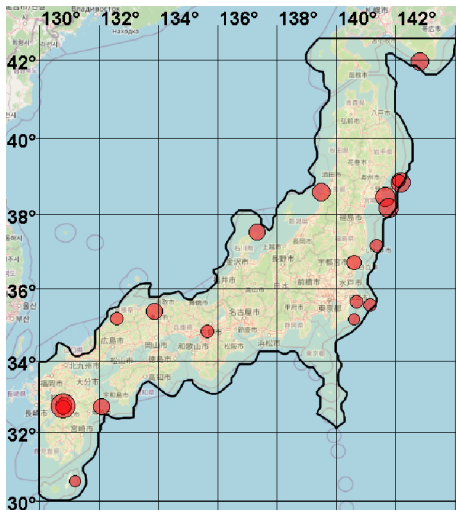


Рис. 11. Зона анализа и эпицентры тестовых землетрясений на интервале 10.04.2015 до 23.04.2023 с магнитудами $m \geq 6.0$ и глубинами эпицентров $H \leq 60$ км

Мы анализировали около 45 сеточных полей, которые представляют пространственно-временные характеристики сейсмического и геодинамического процессов. Наиболее существенными для прогноза землетрясений оказались поля, вычисленные по каталогу землетрясений.

S_1 — поле плотности эпицентров землетрясений, вычисленное по каталогу землетрясений с помощью метода гауссовского ядерного сглаживания. Значения поля плотности эпицентров вычислялись с экспоненциальной ядерной функцией (3). Параметрами поля плотности эпицентров являются: радиус $R_0 = 30$ км, временной интервал $T_0 = 60$ суток, пороговый коэффициент $\epsilon = 2$.

S_2 — поле значений b-value, вычисленное аналогично полю S_1 по каталогу землетрясений с помощью метода гауссовского ядерного сглаживания с экспоненциальной ядерной функцией с параметрами $R = 100$ км, $T = 150$ суток.

S_3 — поле скорости изменения плотности эпицентров во времени. Значения поля в момент t равны $s_3(t) = s_1(t) - s_1(t-1)$.

S_4 — поле изменения b-value во времени. Как и в (6) значения поля $s_4(t)$ в момент t равны отношению разности средних значений $\bar{s}_2(t)$ и $\bar{s}_2(t-T_2)$ в двух последовательных интервалах $T_1 = 1500$ суток и $T_2 = 241$ суток к оценке среднеквадратичного отклонения этой разности.

Пространственно-временной прогноз эпицентров землетрясений

Для прогноза использовались близкие к максимальным значения поля S_1 . Параметры цилиндра предвестника составляют $R = 8$ км, $T = 61$ суток. Тревога в узле сетки продолжается $T = 61$ сутки. Результаты обнаружения эпицентров целевых землетрясений при величине порога объёма тревоги $\theta(V) = 0.2$ показаны в таблице 3. Эпицентр землетрясений не обнаружен,

если значение объёма тревоги V на эпицентре землетрясения более 0.2. Ошибки прогноза выделены жирным шрифтом.

Таблица 3. Результаты пространственно-временного прогноза землетрясений островной части Японии.

№	Дата	Долгота (градусы)	Широта (градусы)	Глубина (км)	Магнитуда	Объем тревоги %
1	2015/05/12	142.15000	38.86300	46.2	6.8	3.35
2	2016/01/14	142.80100	41.97000	51.5	6.7	13.34
3	2016/04/14	130.80900	32.74200	11.4	6.5	11.93
4	2016/04/14	130.77800	32.70100	6.7	6.4	17.62
5	2016/04/15	130.76300	32.75500	12.4	7.3	11.93
6	2016/10/21	133.85600	35.38000	10.6	6.6	9.86
7	2016/11/23	141.34600	37.17500	23.8	6.2	3.57
8	2016/12/28	140.57400	36.72000	10.8	6.3	6.85
9	2018/04/08	132.58700	35.18500	12.1	6.1	54.78
10	2018/06/17	135.62200	34.84400	13.0	6.1	39.79
11	2018/07/07	140.59200	35.16500	56.8	6.0	9.49
12	2019/01/08	131.16500	30.57300	30.1	6.0	9.40
13	2019/06/18	139.47900	38.60800	14.0	6.7	67.46
14	2020/04/19	142.09900	38.88800	46.1	6.2	4.81
15	2020/06/24	141.11300	35.55300	36.1	6.1	5.55
16	2021/03/20	141.62800	38.46800	59.5	6.9	8.71
17	2021/05/01	141.74000	38.17400	51.4	6.8	1.08
18	2022/01/21	132.07200	32.71600	44.6	6.6	19.60
19	2023/05/05	137.30400	37.53900	12.1	6.5	1.45
20	2023/05/26	140.67200	35.64100	49.7	6.2	4.84

На Рис. 12 показаны графики зависимостей $U(V)$ и $U(W)$, где V и W обозначают объёмы тревоги, равные отношениям числа узлов сетки пространственно-временной области тревоги к числу узлов области анализа по данным обучения тестирования. На графиках можно видеть, что 3 эпицентра землетрясений не обнаружены. Остальные 17 эпицентров землетрясений обнаруживаются при пороге $\theta(V) = 0.2$ и при пороге $\theta(W) = 0.23$. На Рис. 13 показаны графики зависимостей $U'(V)$ и $U'(W)$. На графиках можно видеть, что на 3-х интервалах прогноза в зоны тревоги попали не все эпицентры землетрясений. Остальные интервалы со всеми эпицентрами землетрясений в зонах тревоги обнаруживаются при пороге $\theta(V) = 0.2$ и при пороге $\theta(W) \leq 0.22$. На Рис. 12 и 13 можно видеть, что зависимости $U(V)$ и $U(W)$, как и зависимости $U'(V)$ и $U'(W)$, в отличие от примера по региону Камчатки, практически не отличаются друг от друга.

Следуя графикам Рис. 12 и 13, выберем зоны тревоги при прогнозе со значениями $V \leq 0.2$ и при тестировании со значениями $W \leq 0.23$. Тогда качество прогноза согласно (7), (8), (9) характеризуется следующими показателями

Оценка вероятности обнаружения эпицентров землетрясений

$$U = Q^*/Q = 17/20 = 0.85.$$

Оценка вероятности обнаружения интервалов прогноза, на которых все эпицентры целевых землетрясений попадают в прогнозную зону тревоги.

$$U' = M'/M = 14/17 = 0.82.$$

Оценка вероятности того, что на очередном интервале прогноза все эпицентры целевых землетрясений попадают в прогнозную зону тревоги.

$$P_1 = M'/N = 14/99 = 0.14.$$

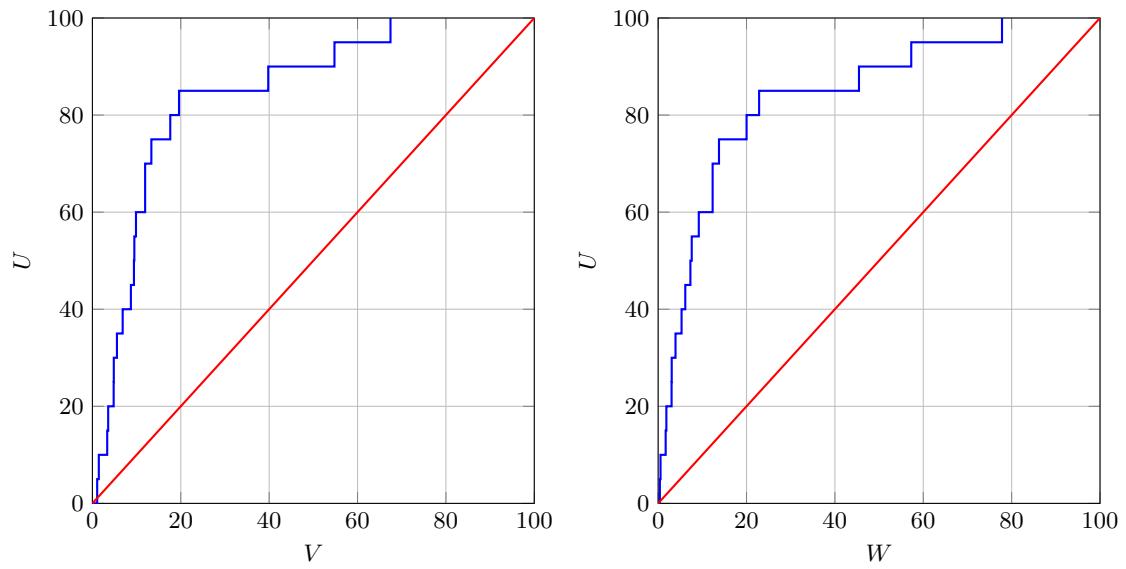


Рис. 12. Графики ROC кривых $U(V)$ (слева) и $U(W)$ (справа). Значения U , V и W даны в процентах.

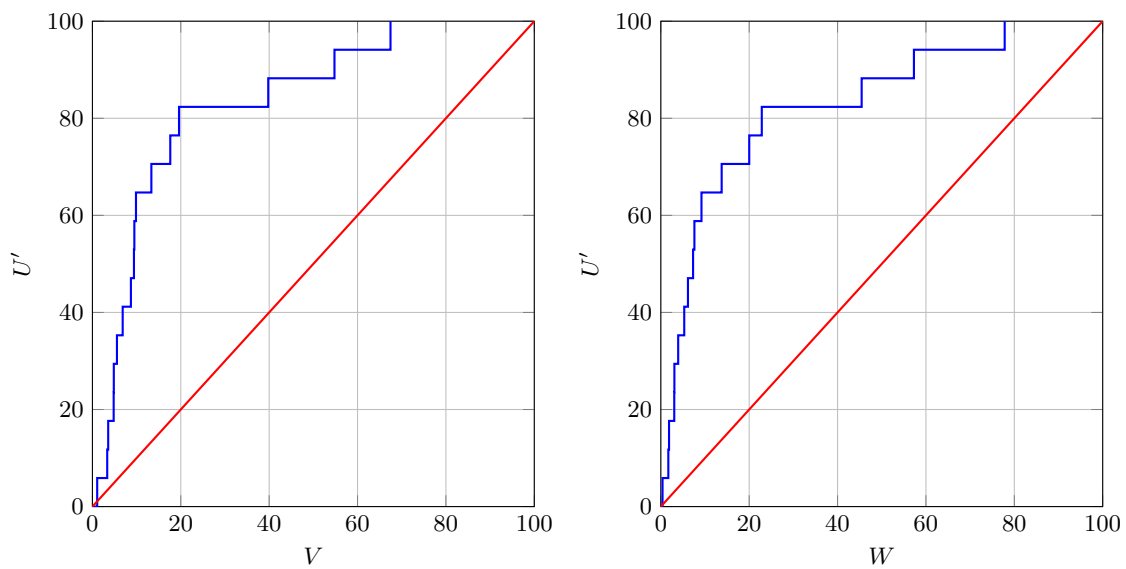


Рис. 13. Графики ROC кривых $U'(V)$ (слева) и $U'(W)$ (справа). Значения U' , V и W даны в процентах.

Прогноз интервалов тревоги

Для прогноза интервалов тревоги использовались близкие к минимальным значения поля S_2 и близкие к максимальным значения поля S_3 . Параметры цилиндра предвестника составляют $R = 16$ км, $T = 61$ суток. Тревога в узле сетки продолжается $T = 61$ сутки.

На Рис. 14 показаны графики зависимостей $U^*(V^*)$ и $U^*(W^*)$. Из графика $U^*(V^*)$ видно, что по результатам тестирования минимальная сумма оценки вероятности ошибки неверной классификации интервала тревоги (пропуск цели), равная $(1 - U^*)$, и объёма тревоги имеет место при величине объёма тревоги $W^* = 0.36$. При этом в $M^* = 13$ интервалах тревоги из $M = 17$ имелись эпицентры целевых землетрясений в зоне анализа. Из (12) мы определяем, что для обнаружения $M^* = 13$ интервалов тревоги с землетрясениями алгоритм выбрал $N^*(W^*) = W^* \cdot N_{\text{тест}} = 0.36 \cdot 99 = 36$ интервалов тревоги из 99. Отсюда, следуя (10) и (13), получаем

$$U^* = M^*/M = 13/17 \approx 0.76.$$

$$P_2(W^*) = M^*/N^*(W^*) = 13/36 \approx 0.36.$$

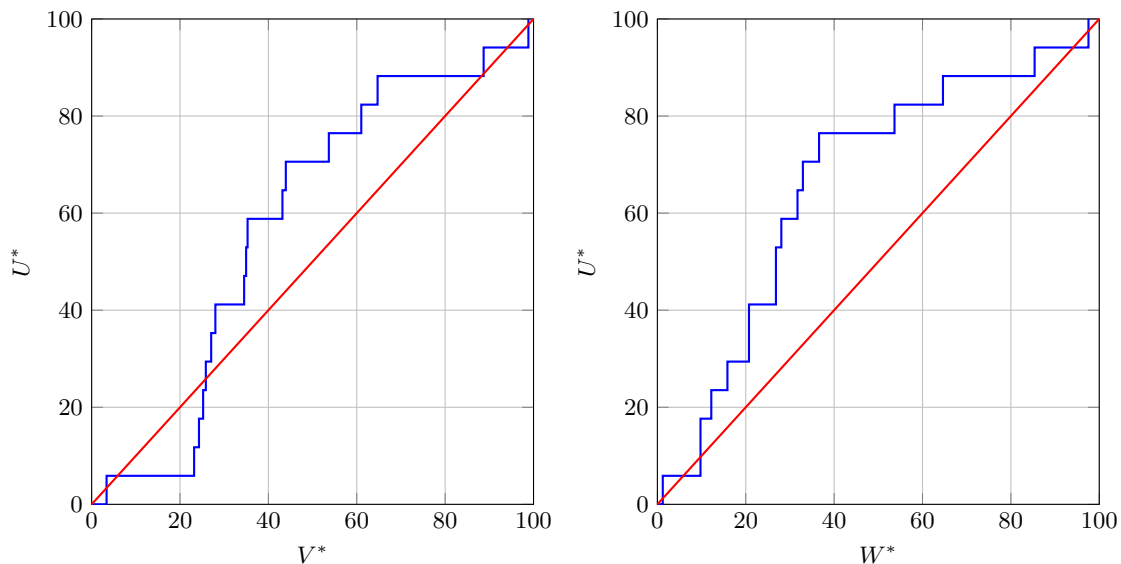


Рис. 14. Графики ROC кривых $U^*(V^*)$ (слева) и $U^*(W^*)$ (справа). Значения U^* , V^* и W^* даны в процентах.

Прогноз зон тревоги

Зоны тревоги вычисляются только на интервалах тревоги. Как и в случае пространственно-временного прогноза мы использовали близкие к максимальным значения поля S_1 . Параметры цилиндра предвестника составляют $R = 8$ км, $T = 61$ суток. Тревога в узле сетки продолжается $T = 61$ сутки.

На Рис. 15 показана ROC кривая зависимости $U''(V)$. Можно видеть, что алгоритм вычислил прогнозные зоны тревоги, в которые при пороге $V = 0.2$ попали все эпицентры землетрясений на $M'' = 11$ интервалах тревоги из $M^* = 13$. Таким образом, качество прогноза определяется следующими оценками:

Из (14) оценка вероятности обнаружения интервалов тревоги, на которых все эпицентры целевых землетрясений попадают в зону тревоги

$$U'' = M''/M^* = 11/13 = 0.85.$$

Оценка вероятности однократного прогноза, при котором на интервале тревоги все эпицентры целевых землетрясений попадают в зоны тревоги (15), имеет вид

$$P_3(W^*) = M''/N^*(W^*) = 11/36 \approx 0.31.$$

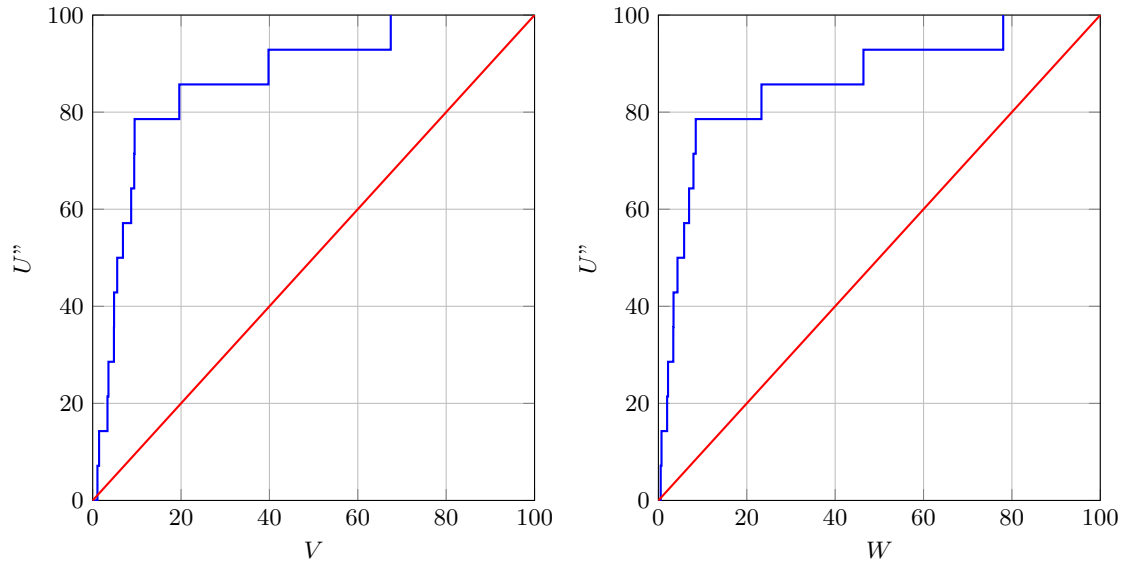


Рис. 15. Графики ROC кривых $U''(V)$ (слева) и $U''(W)$ (справа). Значения U'' , V и W даны в процентах.

4. ДИСКУССИЯ

Моделирование обобщённого метода минимальной области тревоги выполнено для трёх регионов. Во всех трех случаях тестовые испытания включали в себя около 100 тридцатидневных интервалов прогноза. Качество серии прогнозов оценивалось по вероятности обнаружения эпицентров целевых землетрясений, по вероятности обнаружения интервалов тревоги, содержащих землетрясения в зонах анализа, и по вероятности обнаружения всех эпицентров землетрясений в зонах тревоги. Эти вероятности находятся в диапазоне от 0.75 до 0.89. Качество однократного прогноза интервалов тревоги оценивалось по вероятности того, что прогнозируемый интервал тревоги будет содержать землетрясения в зоне анализа, а качество однократного прогноза зон тревоги оценивалось по вероятности того, что все эпицентры землетрясений будет содержать прогнозируемая зона тревоги. Эти вероятности принимают значения от 0.19 до 0.36. При этом по сравнению с одноэтапной процедурой пространственно-временного прогноза условные вероятности однократного прогноза увеличились в 2–3 раза.

Полученный результат моделирования отчасти объясняется самой постановкой задачи прогноза землетрясений. При прогнозе алгоритм машинного обучения должен классифицировать два множества объектов. Обычно в задачах классификации существуют объективные причины, по которым объекты, относящиеся к разным классам, являются сходными между собой и отличаются друг от друга. В нашем случае разделение на классы определяется величиной пороговой магнитуды землетрясений, следуя которой алгоритм прогноза должен отличать землетрясения с большими магнитудами от землетрясений с меньшими магнитудами. Поскольку пороговая магнитуда назначается не на основе физики сейсмического процесса, то само

определение классов не является объективным. Это обстоятельство в значительной степени осложняет задачу прогноза.

Результаты моделирования показывают, что пространственно-временной прогноз эпицентров землетрясений или зон тревоги и прогноз интервалов тревоги с землетрясениями в зоне анализа используют разные сеточные поля признаков. В то же время из раздела 2 известно, что алгоритмы метода минимальной области тревоги и в том и в другом случаях обучаются по предвестникам землетрясений, которыми являются векторы значений полей в узлах сетки, предшествующих эпицентрам землетрясений. Отличия в полях признаков обусловлены тем, что при обучении предвестники землетрясений оцениваются по различным мерам информативности. В первом случае мера информативности предвестника землетрясений определяется по узлам сетки, соответствующим векторам его ортанта, а во втором случае - по интервалам прогноза, соответствующим векторам его ортанта. Далее эти меры информативности учитываются при упорядочивании предвестников. Это объясняет отличия в выборе наиболее информативных полей при пространственно-временном прогнозе и прогнозе по времени.

Результаты прогноза иллюстрированы графиками ROC кривых, которые показывают зависимости вероятности успешного прогноза от объёмов тревоги V , полученных при обучении, и объёмов тревоги W , вычисленных по данным тестирования. На Рис (2)–(5) можно видеть, что эти зависимости сильно отличаются друг от друга только для региона Камчатки. Обычно значительное отличие качества прогноза на учебных и тестовых данных свидетельствует о переобучении. Переобучение бывает при применении решающих правил, у которых число оцениваемых при обучении параметров слишком велико по сравнению с числом прецедентов учебной выборки. Однако пространственно-временной прогноз эпицентров землетрясений (Рис. 2) и прогноз зон тревоги (Рис. 3, 5), в которые попадают все эпицентры землетрясений на интервале прогноза, вычислялся по одному полю признаков. В этом случае наш алгоритм обучения вычисляет область тревоги только по величине порога значений этого поля. Мы предполагаем, что число прецедентов учебной выборки вполне достаточно для оценивания единственного параметра прогнозирующей функции. Следует добавить, что при моделировании мы выбирали наиболее информативные поля признаков и параметры цилиндра предвестника. Но мы считаем эти действия оказываю значительно меньшее влияние на отличия в результатах прогноза по учебному и тестовому материалам.

На Рис. 16 показаны временные диаграммы обучения: график изменения значений объёма тревоги $V^*(t)$, магнитуды целевых землетрясений в зоне анализа, интервал тестирования и порог выбора интервалов тревоги $\theta(V^*) = 0.22\%$. Диаграмма построена по каталогу сейсмических событий Камчатки со следующими параметрами: $m \geq 3.5$, $H \leq 160$, время после 01.01.1998. Интервалам тревоги соответствуют участки графика со значениями меньшими величины порога. Внизу показана гистограмма числа землетрясений, которые использовались для вычисления полей признаков. На рисунке видно, что характер графика и гистограммы существенно изменились в 2010–2011 г. Можно предположить, что значительное увеличение числа землетрясений почти перед началом тестирования 01.01.2013 г. могло быть одной из причин существенных отличий между результатами прогноза, вычисляемыми по данным обучения и по данным тестирования.

Зависимости $U^*(V^*)$ и $U^*(W^*)$ позволяют оценить вероятности обнаружения интервалов тревоги с эпицентрами землетрясений в зоне анализа в ситуации реального прогноза. Обе эти зависимости известны к моменту очередного прогноза. Пусть эти зависимости совпадают с графиками Рис. 4. Допустим, что значение объёма тревоги на очередной интервал прогноза $V^* = 0.1$. Алгоритм принимает решение, что очередной интервал прогноза является интервалом тревоги, и все интервалы прогноза со значениями $V^* \leq 0.1$ являются интервалами тревоги. По графику $U^*(V^*)$ получаем, что порогу по объёму тревоги $V^* = 0.1$ соответствуют

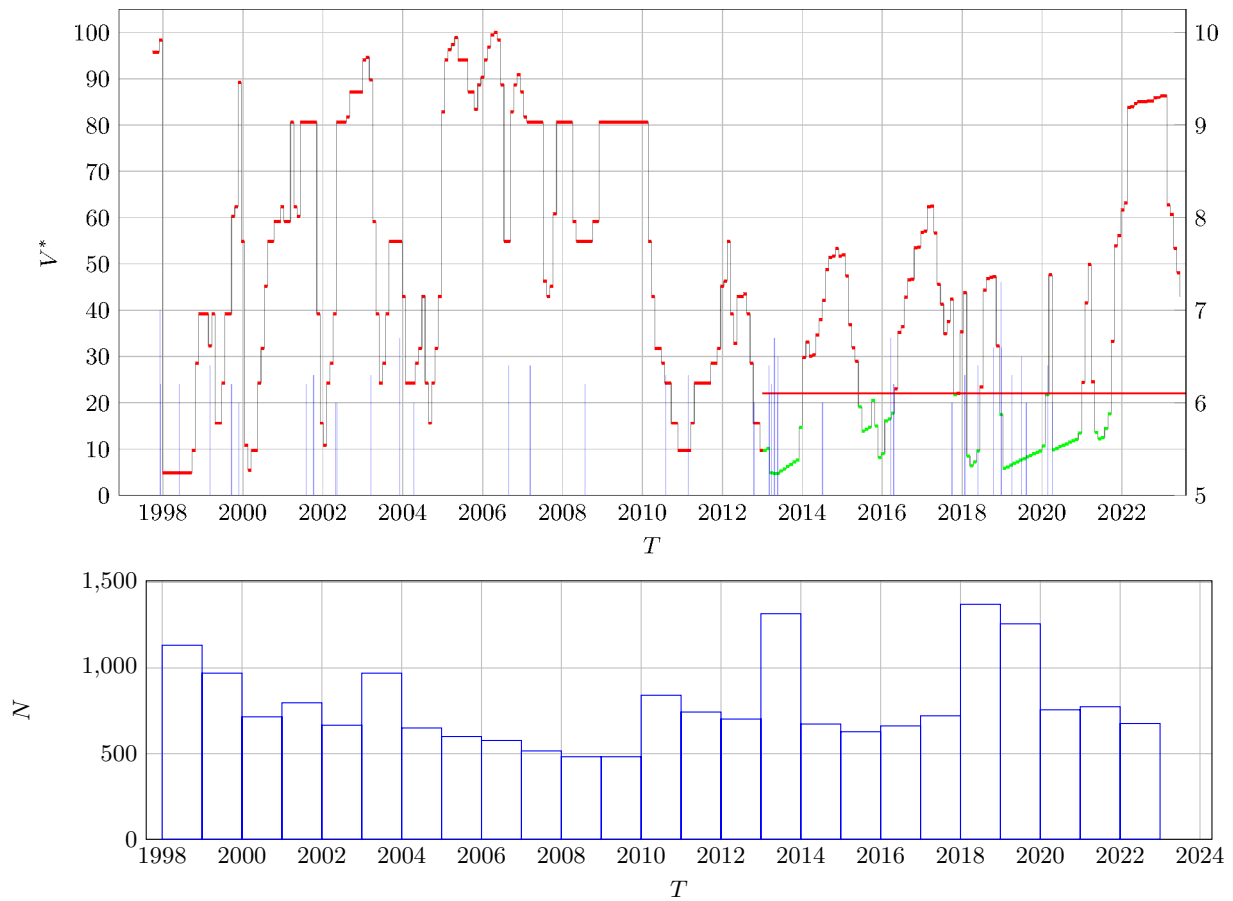


Рис. 16. График изменения значений объёма тревоги при обучении V^* в процентах и последовательность магнитуд целевых землетрясений в зоне анализа. Прямоугольная область показывает интервал тестирования и порог выбор интервалов тревоги $\theta(V^*) = 0.22$. Внизу гистограмма числа землетрясений с эпицентрами в зоне анализа. Значения V даны в процентах.

значения $M^* = 9$ и $W^* = 0.19$. Отсюда согласно (8) получаем, что число интервалов тревоги на тестовых данных

$$N^*(W^*) = W^* \cdot N_{\text{тест}} = 0.19 \cdot 128 = 24$$

Согласно (9) получаем оценку вероятности того, что обнаружения интервалов тревоги с эпицентрами землетрясений в зоне анализа

$$P_2(W^*) = M^*/N^*(W^*) = 9/24 = 0.375.$$

Рассматриваемый интервал прогноза является интервалом тревоги. Поэтому для него и для всех интервалов тревоги, отобранных при значениях объёмов тревоги $V^* \leq 0.1$, алгоритм выполняет второй этап прогноза: вычисление на интервалах тревоги зон тревоги. Графики ROC кривых зависимостей $U''(V)$ и $U''(W)$ показаны на Рис. 17. Из графика $U''(V)$ можно видеть, что зоны тревоги, которые содержат все эпицентры землетрясений, алгоритм построил на $M'' = 8$ интервалах тревоги. Следовательно, если интервалы тревоги выбираются при величине объёма тревоги $V^* = 0.1$, то оценка вероятности того, что данный интервал будет содержать все эпицентры землетрясений в зоне тревоги

$$P_3(W^*) = M''/N^*(W^*) = 8/24 \approx 0.33.$$

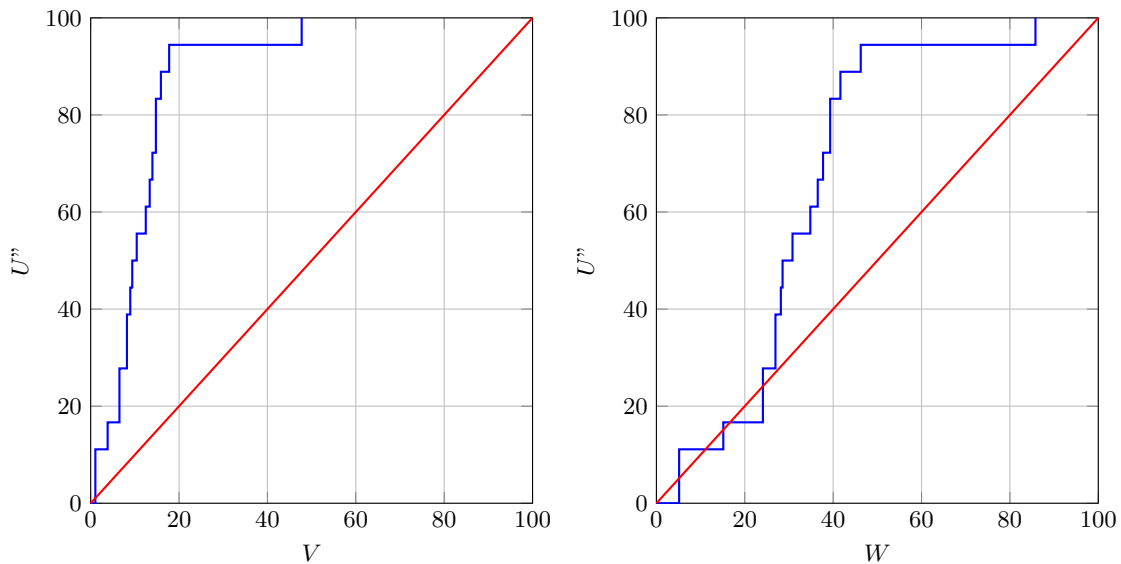


Рис. 17. Графики ROC кривых $U''(V)$ (слева) и $U''(W)$ (справа). Значения U'' , V и W даны в процентах.

5. ЗАКЛЮЧЕНИЕ

Мы рассмотрели основы обобщённого подхода к систематическому прогнозу землетрясений. В нём для определения качества прогноза используются два типа оценок: (1) вероятность обнаружения целевых событий, равная доле обнаруженных событий в серии прогнозов, и (2) вероятность однократного прогноза целевых событий, которая определяет возможность появления всех эпицентров целевых землетрясений в зоне тревоги. Первая оценка определяет эффективность исходных данных, методов их обработки и метода прогноза. Вторая оценка нужна для принятия практических решений, связанных с прогнозом сейсмической опасности.

Решение второй задачи потребовало существенной модернизации метода минимальной области тревоги. Обобщенный метод минимальной области тревоги состоит из двух этапов. На первом этапе прогнозируется возможность появления эпицентра землетрясения с целевой магнитудой в зоне анализа, и принимается решение об объявлении тревоги на очередном интервале прогноза. На втором этапе на интервале тревоги вычисляется зона тревоги, в которой на данном интервале прогноза ожидается появление всех эпицентров землетрясений.

Моделирование нового метода минимальной области тревоги выполнено для регионов Камчатки, Калифорнии и Японии. Для каждого региона оценивалось качество двухэтапного прогноза зон тревоги и качество предыдущей версии одноэтапного пространственно-временного прогноза. Сравнение показало, что во всех трёх случаях оценки вероятности того, что на очередном интервале тревоги все эпицентры целевых землетрясений попадут в зоны тревоги в 2–3 раза выше, чем аналогичные оценки вероятности для одноэтапного прогноза. Мы полагаем, что улучшение качества однократного прогноза следует искать в совершенствовании методов прогноза интервалов тревоги и в исследованиях новых пространственно-временных полей, представляющих свойства сейсмогенных процессов подготовки землетрясений.

СПИСОК ЛИТЕРАТУРЫ

1. Соболев Г. А. Основы прогноза землетрясений. — Общество с ограниченной ответственностью Международная академическая издательская компания "Наука/Интерпериодика 1993.
2. Соболев Г. А., Пономарев А. В. Физика землетрясений и предвестники. — Общество с ограниченной ответственностью Международная академическая издательская компания "Наука/Интерпериодика 2003.
3. Завьялов А. Д. Среднесрочный прогноз землетрясений: основы, методика, реализация. — М.: Наука, 2006. — Т. 254.
4. Lighthill J. A critical review of VAN: earthquake prediction from seismic electrical signals. — 1996.
5. King C. Y. Gas geochemistry applied to earthquake prediction: An overview //Journal of Geophysical Research: Solid Earth. — 1986. — Т. 91. — №. B12. — С. 12269-12281.
6. Matsumoto N., Koizumi N. Recent hydrological and geochemical research for earthquake prediction in Japan //Natural hazards. — 2013. — Т. 69. — С. 1247-1260.
7. Murai S. Can we predict earthquakes with GPS data? //International Journal of Digital Earth. — 2010. — Т. 3. — №. 1. — С. 83-90.
8. Geller R. J. et al. Earthquakes cannot be predicted //Science. — 1997. — Т. 275. — №. 5306. — С. 1616-1616.
9. Короновский Н. В., Наймарк А. А. Непредсказуемость землетрясений как фундаментальное следствие нелинейности геодинамических систем //Вестник Московского университета. Серия 4. Геология. — 2012. — №. 6. — С. 3-12.
10. Гуфельд И. Л., Матвеева М. И., Новоселов О. Н. Почему мы не можем осуществить прогноз сильных коровых землетрясений //Геодинамика и тектонофизика. — 2011. — Т. 2. — №. 4. — С. 378–415.
11. Keilis-Borok V., Soloviev A. A. (ed.). Nonlinear dynamics of the lithosphere and earthquake prediction. — Springer Science and Business Media, 2002.
12. Kossobokov V., Shebalin P. Earthquake prediction //Nonlinear dynamics of the lithosphere and earthquake prediction. — Berlin, Heidelberg : Springer Berlin Heidelberg, 2003. — С. 141-207.
13. Corbi F. et al. Machine learning can predict the timing and size of analog earthquakes //Geophysical Research Letters. — 2019. — Т. 46. — №. 3. — С. 1303–1311.
14. Shebalin P. N. et al. Combining earthquake forecasts using differential probability gains //Earth, Planets and Space. — 2014. — Т. 66. — №. 1. — С. 1–14.

15. Ainei A., Fu W., Ho C. H. Time series analysis for predicting the occurrences of large scale earthquakes //International Journal of Applied Science and Technology. — 2012. — Т. 2. — №. 7.
16. Marzocchi W., Zechar J. D. Earthquake forecasting and earthquake prediction: different approaches for obtaining the best model //Seismological Research Letters. — 2011. — Т. 82. — №. 3. — С. 442–448.
17. Moustira M., Avraamides M., Christodoulou C. Artificial neural networks for earthquake prediction using time series magnitude data or seismic electric signals //Expert systems with applications. — 2011. — Т. 38. — №. 12. — С. 15032–15039.
18. Rhoades D. A. Mixture models for improved earthquake forecasting with short-to-medium time horizons //Bulletin of the Seismological Society of America. — 2013. — Т. 103. — №. 4. — С. 2203–2215.
19. Alves E. I. Earthquake forecasting using neural networks: results and future work //Nonlinear Dynamics. — 2006. — Т. 44. — С. 341–349.
20. Priambodo B., Mahmudy W. F., Rahman M. A. Earthquake magnitude and grid-based location prediction using backpropagation neural network //Knowledge Engineering and Data Science. — 2020. — Т. 3. — №. 1. — С. 28–39.
21. Mignan A., Broccardo M. Neural network applications in earthquake prediction (1994–2019): Meta-analytic and statistical insights on their limitations //Seismological Research Letters. — 2020. — Т. 91. — №. 4. — С. 2330–2342.
22. Asim K. M. et al. Earthquake prediction model using support vector regressor and hybrid neural networks //PloS one. — 2018. — Т. 13. — №. 7. — С.
23. Panakkat A., Adeli H. Neural network models for earthquake magnitude prediction using multiple seismicity indicators //International journal of neural systems. — 2007. — Т. 17. — №. 01. — С. 13–33.
24. Gitis V. G., Derendyaev A. B. An approach to the automatic prediction of earthquakes //Journal of Communications Technology and Electronics. — 2016. — Т. 61. — С. 1466–1473.
25. Gitis V. G., Derendyaev A. B. Earthquake prediction learning using the least alarm method //Journal of Communications Technology and Electronics. — 2018. — Т. 63. — С. 680–690.
26. Gitis V. G., Derendyaev A. B. Web-Based GIS platform for automatic prediction of earthquakes //International Conference on Computational Science and Its Applications. — Cham : Springer International Publishing, 2018. — С. 268–283.
27. Gitis V. G., Derendyaev A. B. Machine learning methods for seismic hazards forecast //Geosciences. — 2019. — Т. 9. — №. 7. — С. 308.
28. Gitis V., Derendyaev A. From monitoring of seismic fields to the automatic forecasting of earthquakes //International journal of web information systems. — 2019. — Т. 15. — №. 5. — С. 535–549.
29. Gitis V., Derendyaev A. The method of the minimum area of alarm for earthquake magnitude prediction //Frontiers in Earth Science. — 2020. — Т. 11. — С. 585317.
30. Gitis V., Derendyaev A., Petrov K. Analyzing the performance of GPS data for earthquake prediction //Remote Sensing. — 2021. — Т. 13. — №. 9. — С. 1842.
31. Gitis V. G., Derendyaev A. B., Petrov K. N. Earthquake prediction based on combined seismic and GPS monitoring data //International Conference on Computational Science and Its Applications. — Cham : Springer International Publishing, 2021. — С. 601–612.
32. Gitis V., Derendyaev A. A Technology for Seismogenic Process Monitoring and Systematic Earthquake Forecasting //Remote Sensing. — 2023. — Т. 15. — №. 8. — С. 2171.
33. Bishop, C.M. Pattern recognition and machine learning. — New York : springer, 2006
34. Kotsiantis S. B. et al. Supervised machine learning: A review of classification techniques //Emerging artificial intelligence applications in computer engineering. — 2007. — Т. 160. — №. 1. — С. 3–24.
35. Khan S. S., Madden M. G. A survey of recent trends in one class classification //Artificial Intelligence and Cognitive Science: 20th Irish Conference, AICS 2009, Dublin, Ireland, August 19–21, 2009, Revised Selected Papers 20. — Springer Berlin Heidelberg, 2010. — С. 188–197.

36. Чебров В. Н., Салтыков В. А., Серафимова Ю. К. Опыт выявления предвестников сильных (м 6.0) землетрясений на Камчатке в 1998–2011 гг. // Вулканонология и сейсмология. — 2013. — №. 1. — С. 85–85.
37. Polzehl J., Spokoiny V. G. Adaptive weights smoothing with applications to image restoration // Journal of the Royal Statistical Society: Series B (Statistical Methodology). — 2000. — Т. 62. — №. 2. — С. 335–354.
38. Polzehl J., Spokoiny V. Propagation–separation approach for local likelihood estimation // Probability Theory and Related Fields. — 2006. — Т. 135. — С. 335–362.
39. Kullback S. Information theory and statistics. — Courier Corporation, 1997.
40. Gitis V. G. et al. Adaptive estimation of seismic parameter fields from earthquake catalogs // Journal of Communications Technology and Electronics. — 2015. — Т. 60. — С. 1459–1465.
41. Gitis V. G. et al. Earthquake prediction using the fields estimated by an adaptive algorithm // Proceedings of the 7th International Conference on Web Intelligence, Mining and Semantics. — 2017. — С. 1–8.
42. Barnhart W. D., Hayes G. P., Wald D. J. Global earthquake response with imaging geodesy: Recent examples from the USGS NEIC // Remote Sensing. — 2019. — Т. 11. — №. 11. — С. 1357.
43. Blewitt G. et al. Harnessing the GPS data explosion for interdisciplinary science // Eos. — 2018. — Т. 99.
44. Okada Y. et al. Recent progress of seismic observation networks in Japan — Hi-net, F-net, K-NET and KiK-net — // Earth, Planets and Space. — 2004. — Т. 56. — С. xv–xxviii.
45. Obara K. et al. A densely distributed high-sensitivity seismograph network in Japan: Hi-net by National Research Institute for Earth Science and Disaster Prevention // Review of scientific instruments. — 2005. — Т. 76. — №. 2.

Optimization of the approach to systematic earthquake forecasting

V.G. Gitis, A.B. Derendyaev

The systematic forecast of earthquakes is carried out regularly in a pre-selected analysis zone. At each iteration, new data on the seismic process is loaded, processed, transformed into grid-based spatiotemporal fields, machine learning is applied, and a forecast is provided with a constant time interval. The result is a map of the alert zone where epicenters of all target earthquakes are expected within the forecast interval. The minimum alert area method is used for the forecast. In the new version of the systematic earthquake forecast, the solution to the problem is divided into two stages. At the first stage, the algorithm identifies alarm intervals containing target earthquakes with epicenters in the analysis area. At the second stage, during alarm intervals, the algorithm predicts alarm zones containing all epicenters of target earthquakes. This allows to optimize the estimation of the probability of detecting epicenters of target earthquakes in a series of forecasts and the probability that all epicenters of target earthquakes will be within the predicted alert zone in a single forecast. An example of applying the method to the earthquake forecast in Kamchatka, California and the island part of Japan are considered.

KEYWORDS: systematic earthquake prediction, machine learning, minimum alarm area method, GPS time series, forecast quality