

ОСНОВЫ ДВУХЭТАПНОГО ПОДХОДА К СИСТЕМАТИЧЕСКОМУ ПРОГНОЗУ ЗЕМЛЕТРЯСЕНИЙ

В. Г. Гитис¹, А. Б. Дерендяев^{*,1}

Институт проблем передачи информации им. А. А. Харкевича Российской академии наук (ИППИ РАН), г. Москва, Россия

* **Контакт:** Александр Борисович Дерендяев, wintsa@gmail.com

Систематический прогноз землетрясений производится регулярно с постоянным интервалом в заранее выбранной сейсмически однородной зоне. Результатом каждой итерации прогноза является карта зоны тревоги, в которой ожидаются эпицентры целевых землетрясений. В рассматриваемой технологии реализованы следующие новые положения: 1 – Решение считается успешным, если на интервале прогноза все эпицентры целевых землетрясений попали в зону тревоги. 2 – Технология оптимизирует вероятность успешного обнаружения эпицентров землетрясений в серии прогнозов и вероятность успешного прогноза на очередной итерации. 3 – Технология позволяет оценить вероятность успешного решения на очередном интервале прогноза. Рассмотрены примеры применения метода для прогноза землетрясений Камчатки, Калифорнии и островной части Японии.

Ключевые слова: систематический прогноз землетрясений, машинное обучение, метод минимальной области тревоги, временные ряды GPS.

Цитирование: Гитис, В. Г., А. Б. Дерендяев Основы двухэтапного подхода к систематическому прогнозу землетрясений // Russian Journal of Earth Sciences. — 2025. — Т. 25. — ES3010. — DOI: 10.2205/2025es000987 — EDN: MFJOOB

1. Введение

Идея о возможности прогнозирования землетрясений основывается на данных физического моделирования и фактических наблюдениях, которые свидетельствуют о том, что перед землетрясением происходят процессы, вызывающие аномальные изменения в геологической среде в районе очага [Завьялов, 2006; Соболев, 1993; Соболев и Пономарев, 2003; Kagan, 2013; King, 1986; Lighthill, 1996; Soloviev et al., 2014]. Однако успешность прогноза затруднена недостаточной изученностью моделей сейсмичности и неполнотой инструментальных данных о поведении сейсмического процесса. Некоторые авторы полагают гипотетичной саму возможность решения проблемы прогноза. В то же время никто не отрицает, что накопление новых данных, методов их обработки и знаний о физических процессах, ведущих к землетрясению, обеспечит прогресс в данной области.

В настоящее время появились новые системы сейсмологических, геодезических, электромагнитных и атмосферных наблюдений. Разрабатываются новые методы машинного обучения предсказанию редких аномальных явлений [Amei et al., 2012; Asim et al., 2018; Corbi et al., 2019; Kail et al., 2022; Kossobokov and Shebalin, 2003; Mignan and Broccardo, 2020; Panakkat and Adeli, 2007; Rhoades, 2013; Shebalin et al., 2014]. 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.

В статье рассматривается технология систематического прогноза землетрясений, которая направлена на оптимизацию двух критериев качества: вероятности успешного обнаружения всех эпицентров целевых землетрясений, происходящих на интервале

<https://elibrary.ru/MFJOOB>

Получено: 3 сентября 2024 г.
Принято: 15 января 2025 г.
Опубликовано: 21 мая 2025 г.



© 2025. Автор.

прогноза, и вероятности того, что на очередном интервале произойдут целевые землетрясения с эпицентрами в прогнозируемой зоне тревоги. Первый критерий отличается от общепринятого тем, что качество решения определяется не долей числа успешно обнаруженных эпицентров землетрясений от числа всех целевых эпицентров, а долей числа интервалов прогноза, на которых обнаружены все эпицентры землетрясений от числа всех интервалов с целевыми эпицентрами. Второй критерий важен для оценки эффективности практического использования результатов прогноза. Оптимизация этого критерия потребовала существенного развития подхода к систематическому прогнозу землетрясений. Основные элементы новой версии представлены в [разд. 2](#). В [разд. 3](#) и [разд. 4](#) представлены результаты моделирования прогноза землетрясений в регионах Камчатки, Калифорнии и Японии.

2. Систематический прогноз землетрясений

2.1. Качество прогноза

Систематический прогноз землетрясений дается регулярно и на постоянный временной интервал. Прогноз считается успешным, если все эпицентры землетрясений с целевыми магнитудами попадают в ограниченные по площади зоны тревоги.

Пусть число всех интервалов прогноза равняется N , число всех интервалов прогноза с эпицентрами в зоне анализа равняется M , а $M^* \leq M$ – число интервалов, на которых все эпицентры целевых землетрясений попали в ограниченные по размеру зоны тревоги.

Степень успешности прогноза землетрясений определяют следующие два показателя.

1. Вероятность обнаружения U , равная отношению:

$$U = \frac{M^*}{M}. \quad (1)$$

Чувствительность U позволяет оценить качество алгоритма обучения и информации, используемой для прогноза.

2. Вероятность P того, что на очередном интервале прогноз будет успешным:

$$P = \frac{M^*}{N}. \quad (2)$$

Показатель позволяет оценить эффективность практического использования результатов прогноза.

Оценка P зависит от количества интервалов, на которых прогнозируется зона тревоги. Известно, что сильные землетрясения происходят сравнительно редко. При большом количестве интервалов прогноза N и малом количестве интервалов прогноза с эпицентрами в зоне анализа M вероятность P оказывается слишком малой для принятия формальных практических решений.

Повышение качества прогноза возможно, если зоны тревоги рассчитывать не на каждом интервале прогноза, а только при условии, что на заданном интервале в зоне анализа ожидаются землетрясения. Для этого прогноз выполняется в два этапа, где первый этап заключается в определении интервалов тревоги, в которых ожидается появление эпицентров целевых землетрясений в зоне анализа. Второй этап заключается в вычислении на этом интервале тревоги зоны тревоги, в которой ожидаются все эпицентры землетрясений. Это позволяет оптимизировать оценку условной вероятности P .

2.2. Двухэтапный метод минимальной области тревоги

На первом этапе интервал тревоги вычислен успешно, если на нём произошли землетрясения с эпицентрами в зоне анализа. На втором этапе зона тревоги вычислена успешно, если в неё попали все эпицентры землетрясений, произошедшие на данном интервале тревоги.

Машинное обучение для прогноза выполняется на каждой итерации по всем ретроспективным данным: эпицентрам землетрясений и временных рядов космической геодезии. Все существующие данные преобразуются в сеточные поля признаков прогноза. Значениям полей в узлах сетки соответствуют компоненты векторов пространства признаков и атрибуты: номер временного среза t и пространственные координаты узла сетки (x, y) . По эпицентрам целевых землетрясений алгоритм обучения находит векторы пространства признаков, которые соответствуют узлам сетки, предшествующим эпицентрам землетрясениям и возможно являются предвестниками.

Модель метода минимальной области тревоги описана в работе [Gitis et al., 2021]. Пусть известен предвестник землетрясения $\mathbf{f}^{(n)}$. Векторы пространства признаков, покомпонентно большие или равные предвестнику также могут предшествовать аналогичным землетрясениям.

Выделенное множество векторов назовём ортантом $h^{(n)}$ с вершиной $\mathbf{f}^{(n)}$. Векторам ортанта $h^{(q)}$ соответствует множество $H^{(q)}$, состоящее для первого этапа прогноза из интервалов прогноза, а для второго этапа прогноза – из узлов координатной сетки.

Схемы обучения первого и второго этапов метода минимальной области тревоги принципиально совпадают.

Цель этапа 1: выбирать ограниченное число интервалов тревоги, которые содержат наибольшее число интервалов прогноза с целевыми землетрясениями в зоне анализа.

Цель этапа 2: вычислять на наибольшем числе интервалов тревоги ограниченные по площади зоны тревоги, которые содержат все целевые землетрясения интервала.

Алгоритм обучения метода минимальной области тревоги для двухэтапной схемы прогноза состоит из 9 операций.

1. *Определить предвестники целевых землетрясений.* Предвестниками целевого землетрясения являются векторы пространства признаков, соответствующие всем узлам сетки из цилиндра предвестника с центром основания в эпицентре землетрясения, радиусом R и образующей T .
2. *Оценить информативность предвестников.* Мера информативности предвестника оценивается по его ортанту.

Для простоты будем рассматривать алгоритм двухэтапного метода обучения для частного случая, когда мера информативности предвестника $\mathbf{f}^{(n)}$ определяется *объёмом тревоги* $\nu^{(n)}$ ортанта $h^{(n)}$. Мера информативности предвестника тем больше, чем меньше объём тревоги его ортанта. Векторам ортанта $h^{(n)}$ на первом этапе соответствует множество $H^{(n)}$ интервалов прогноза, а на втором этапе – множество узлов координатной сетки. Объём тревоги ортанта равен

$$\nu^{(n)} = |H^{(n)}|/|H|, \quad (3)$$

где: Этап 1. $|H|$ – мощность множества интервалов прогноза. Этап 2. $|H|$ – мощность множества всех узлов координатной сетки зоны анализа в координатах пространство-время.

3. Упорядочить ортанты по мере их информативности $\nu^{(a)} \leq \nu^{(b)} \leq \nu^{(c)} \leq \dots \leq 1$.

Упорядоченность предвестников по мере информативности обеспечивает первым предвестникам наиболее значимое влияние на качество прогноза. В нашем случае это предвестники с малыми значениями объёма тревоги.

4. Присвоить всем векторам пространства признаков значение 1.
5. Выбрать предвестник с наивысшей мерой информативности (с наименьшим объёмом тревоги $\nu^{(a)}$).
6. Присвоить векторам ортанта $h^{(a)}$ значение $V(\mathbf{f}^{(a)})$, равное объёму области тревоги $\nu^{(a)}$.
7. Выбрать предвестник $\mathbf{f}^{(b)}$ с наибольшей мерой информативности среди оставшихся.
8. Присвоить множеству векторов $h^{(b)} \setminus h^{(a)}$ значение $V(\mathbf{f}^{(b)})$, равное объёму тревоги объединения ортантов $h^{(a)} \cup h^{(b)}$, и т.д.

9. Отобразить функцию объёма тревоги $V(\mathbf{f})$, на первом этапе на интервалы прогноза t , а на втором этапе на пространственные узлы координатной сетки (x, y) . При этом на первом этапе значение объёма тревоги для t -го временного среза $V(t)$ равно минимальному значению объёмов тревоги векторов, соответствующих узлам сетки этого среза $V(\mathbf{f}_{x,y,t})$:

$$V(t) = \min_{x,y} V(\mathbf{f}_{x,y,t}). \quad (4)$$

Функция объёма тревоги $V(\mathbf{f})$ вычисляется на всем материале обучения. Согласно алгоритму обучения, на первом этапе значение функции объёма тревоги $V(\mathbf{f}) = \mathbb{V}$ равно отношению числа интервалов прогноза со значениями объёма тревоги меньшими или равными $\mathbb{V}$ к числу всех интервалов прогноза. На втором этапе значение функции объёма тревоги $V(\mathbf{f}) = \mathbb{V}$ равно отношению числа узлов сетки со значениями объёма тревоги меньшими или равными $\mathbb{V}$ к числу всех узлов сетки зоны анализа в координатах пространство-время.

На каждой итерации для принятия решения о назначении интервала прогноза интервалом тревоги или решения о размере зоны тревоги выбирается пороговое значение объёма тревоги $\mathbb{V}$. Порог может выбираться заранее на основе качественных показателей прогноза или оцениваться на каждой итерации из условий минимума функции потерь. В частности, порог может вычисляться из условия минимума суммы вероятности ошибок пропуска хотя бы одного эпицентра землетрясения в зоне тревоги и величины объёма тревоги при обучении $\mathbb{V}$. Решение о прогнозе принимается по неравенству $V(\mathbf{f}) \leq \mathbb{V}$. Для первого этапа выполнение этого неравенства на временном срезе t определяет назначение интервала $(t, t + \Delta t)$ интервалом тревоги, в котором ожидаются эпицентры целевых землетрясений в зоне анализа. Для второго этапа выполнение неравенства на интервале тревоги $(t, t + \Delta t)$ определяет выбор узлов сетки зоны тревоги, в которой ожидаются все эпицентры целевых землетрясений.

Можно видеть, что алгоритм двухэтапного прогноза даёт результат, эквивалентный результату одноэтапного прогноза, у которого при подсчете вероятности P исключены интервалы, на которых не объявлена тревога.

Важно отметить, что вычисленной при обучении зоне тревоги можно дать текстовое объяснение. Действительно, алгоритм обучения формирует зоны тревоги из узлов сетки, которые соответствуют векторам из объединения ортантов. Принадлежность каждого узла сетки к зоне тревоги определяется конъюнкциями, выражающими принадлежность соответствующих векторов $\mathbf{f}$ ортантам $h^{(n)}$ с вершинами в точках предвестников $\mathbf{f}^{(n)}$

$$\mathbf{f} \in h^{(n)}, \quad \text{если} \quad \bigcap_{i=1}^I (f_i \geq f_i^{(n)}). \quad (5)$$

Далее принадлежность всех узлов сетки зоне тревоги можно объяснить с помощью импликации: «ЕСЛИ выполняются (конъюнкция 1) ИЛИ (конъюнкция 2) ИЛИ ..., ТО прогнозируется данная зона тревоги». Большинство узлов сетки зоны тревоги относится к одним и тем же ортантам. Поэтому число конъюнкций в объяснении невелико. Кроме того, объяснять зону тревоги можно с помощью представления списка аналогичных землетрясений, участвовавших в обучении. Аналогичными являются землетрясения с предвестниками, которые являются вершинами ортантов, содержащих векторы, соответствующие узлам сетки зоны тревоги.

3. Моделирование

При моделировании мы указываем время начала обучения, время начала тестирования и время конца тестирования. От времени начала обучения до начала тестирования выполняется обучение области тревоги для прогноза на первом интервале тестирования. На следующей итерации обучение выполняется заново от времени начала обучения до начала второго интервала тестирования. Эта процедура повторяется, и к моменту времени конца тестирования выполняется обучение по всем анализируемым данным

и даётся прогноз на интервал, следующий после конца тестирования. Таким образом, на каждой итерации выполняется обучение и последний срез поля тревоги определяет прогноз на следующий интервал.

При моделировании мы сравниваем результаты прогноза по одноэтапной и двухэтапной схемам. В одноэтапной схеме зона тревоги вычисляется перед каждым прогнозом [Gitis and Derendyaev, 2020]. В двухэтапной схеме сначала принимается решение, является ли следующий интервал прогноза интервалом тревоги. Затем зона тревоги рассчитывается только для интервала тревоги. Оценки объёма тревоги, полученные алгоритмом в процессе обучения, обозначим буквой V . Оценка вероятности успешного обнаружения прогнозируемых событий U равна отношению количества успешных прогнозов к количеству всех предсказываемых событий при условии, что объём тревоги при обучении V не превышает $\mathbb{V}$. Результаты обучения могут быть представлены зависимостью $U(V)$. Объём тревоги на основе данных тестирования обозначим как W . Отличие от V заключается в том, что V – это ожидаемый объём тревоги, определенный при обучении, а W – фактически наблюдаемый объём тревоги на основе тестирования.

Моделирование выполняется для регионов Камчатка, Калифорния и Япония. Во всех трех случаях мерой информативности ортанта является его объём тревоги на данных обучения. На каждом интервале прогноза порог объёма тревоги $\mathbb{V}$ выбирается из условия минимума суммы вероятности ошибок пропуска хотя бы одного эпицентра землетрясения в зоне тревоги, и объёма тревоги при обучении V .

3.1. Прогноз землетрясений

Камчатка. Прогноз землетрясений Камчатки выполняется по данным каталога землетрясений Камчатского филиала ФИЦ «Единая Геофизическая служба РАН», <http://sdis.emsd.ru/info/earthquakes/catalogue.php> [Чебров и др., 2013; Чеброва и др., 2020]. 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 и произошло $Q = 30$ целевых землетрясений с эпицентрами в зоне анализа. Зона анализа и тестовые эпицентры землетрясений показаны на рис. 1.

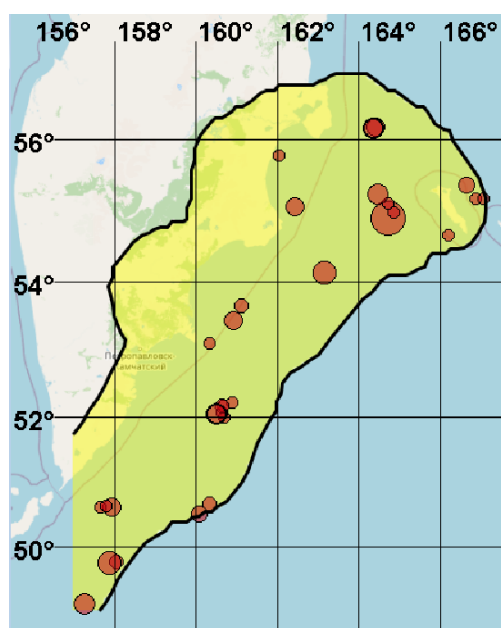


Рис. 1. Зона анализа и эпицентры тестовых землетрясений на интервале 08.01.2013 до 14.07.2023 с магнитудами $m \geq 6,0$ и глубинами эпицентров $H \leq 60$ км.

Мы анализировали около 30 сеточных полей. Наиболее существенными оказались следующие поля.

S_1 – поле плотности эпицентров землетрясений, вычисленное методом гауссовского ядерного сглаживания с радиусом $R_0 = 50$ км и коэффициентом затухания $T_0 = 100$ суток.

S_2 – поле b -value, вычисленное аналогично S_1 с радиусом $R_0 = 100$ км и коэффициентом затухания $T_0 = 365$ суток.

S_3 – поле плотности эпицентров, вычисленное с помощью метода адаптивного весового сглаживания (AWS) по полю S_1 [Gitis and Derendyaev, 2023; Gitis et al., 2015, 2017].

S_4 – поле изменения S_3 во времени. Значения поля $s_4(t)$ в момент t равны 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,

S_5 – поле AWS b -value, вычисленное по полю S_2 с помощью метода AWS.

$S_6 = S_4 \cdot S_5$ – поле произведений значений полей S_4 и 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 $s_1 = 1$.

Одноэтапный прогноз. При прогнозе на каждом интервале тестирования вычислялись зоны тревоги, в которых ожидалось все землетрясения интервала. Для прогноза использовались значения поля S_4 , близкие к максимальным. Параметры цилиндра предвестника $R = 8$ км, $T = 91$ суток.

Наглядным представлением качества прогноза являются графики зависимостей оценок вероятности обнаружения интервалов прогноза, на которых все эпицентры целевых землетрясений попадают в прогнозную зону тревоги, от объёма тревоги при обучении $U(V)$ и от объёма тревоги при тестировании $U(W)$. Оси абсцисс графиков показывают долю средней площади зоны тревоги относительно площади зоны анализа, а по оси ординат можно определить соответствующие значениям V и W количества интервалов, которые содержат все эпицентры целевых землетрясений. Эти зависимости аналогичны ROC-кривым (receiver operating characteristic, рабочая характеристика приёмника) [Bradley, 1997; Kaplan and McFall, 1951]. В работах по теории прогноза землетрясений качество прогноза часто представляют с помощью диаграммы ошибок Молчана [Молчан, 1991; Molchan, 2010].

Графики зависимости $U(V)$ для одноэтапного прогноза показаны на рис. 2. На графиках можно видеть, что при выбранных порогах $V = 0,18$ и $W = 0,31$ на $M^* = 13$ интервалах из $M = 17$ в зону тревоги попали все эпицентры землетрясений. На рисунке можно видеть существенные отличия зависимостей $U(V)$ и $U(W)$ по величинам объёмов тревоги V и W . Возможной причиной этих отличий может являться значимое увеличение числа зарегистрированных землетрясений после 2010 г.

Двухэтапный прогноз. На первом этапе прогнозируются целевые землетрясения с эпицентрами в зоне анализа. Функция, прогнозирующая интервалы тревоги, определяется по значениям полей S_3 и S_5 , близким к максимальным. Параметрами цилиндра предвестника являются: радиус $R = 10$ км и образующая $T = 61$ суток.

На втором этапе на каждом интервале тревоги прогнозируются зоны тревоги со всеми эпицентрами целевых землетрясений. Для вычисления зон тревоги используются близкие к максимальным значения поля S_4 . Параметрами цилиндра предвестника являются радиус цилиндра $R = 8$ км и образующая $T = 91$ суток.

Калифорния. Для прогноза землетрясений Калифорнии анализировались два типа исходных данных: каталог землетрясений Национального центра данных (National Earthquake Information Center, NEIC) [Barnhart et al., 2019] на интервале 01.01.1995 – 23.09.2023 с магнитудами $m \geq 2,4$ и глубинами гипоцентров $H \leq 160$ и временные ряды суточных горизонтальных смещений поверхности земли на интервале

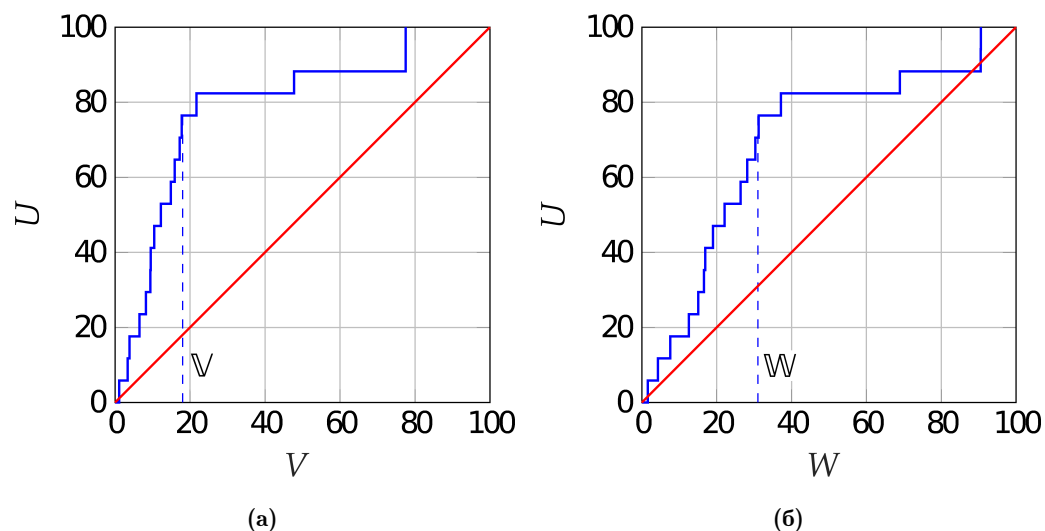


Рис. 2. Графики зависимостей $U(V)$ и $U(W)$ для региона Камчатки. Значения U , V и W даны в процентах.

01.01.2008 – 14.11.2023. Временные ряды получены из Nevada Geodetic Laboratory (NGL), <http://geodesy.unr.edu/about.php> [Blewitt et al., 2018]. Среднее расстояние между 1204 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. Тестирование выполняется от 08.01.2015 до 23.09.2023. Прогноз даётся с интервалом 30 суток. За время тестирования было сделано $N = 107$ прогнозов. Из них на $M = 10$ интервалах прогноза произошли $Q = 18$ землетрясений в зоне анализа. Зона анализа и эпицентры тестируемых землетрясений показаны на рис. 3.

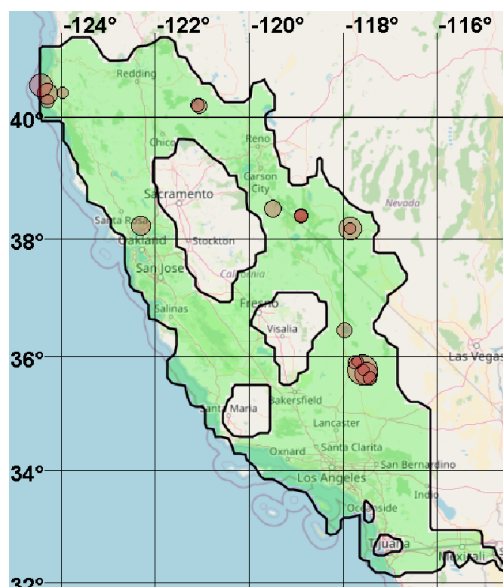


Рис. 3. Зона анализа и эпицентры тестовых землетрясений на интервале 08.01.2015 до 23.09.2023 с магнитудами $m \geq 5,3$ и глубинами эпицентров $H \leq 60$ км.

Мы анализировали около 40 сеточных полей, вычисленных по каталогу землетрясений и по временным рядам смещений поверхности Земли по данным GPS.

S_1 – поле плотности эпицентров землетрясений, вычисленное с помощью метода гауссовского ядерного сглаживания с радиусом $R_0 = 50$ км, временным интервалом $T_0 = 50$ суток.

S_2 – поле значений b -value, вычисленное аналогично полю S_1 с параметрами $R = 150$ км, $T = 365$ суток.

S_3 – поле изменения плотности эпицентров во времени. Значения поля равны отношениям разности средних значений поля S_1 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. Значения поля равны отношениям разности средних значений поля S_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 [Gitis et al., 2021].

F_1 – поле дивергенции скорости деформаций.

F_2 – сглаженное поле изменений дивергенции скорости деформаций. Значения поля равны сглаженным с радиусом $R = 20$ км отношениям разности средних значений дивергенции в двух последовательных интервалах $T_1 = T_2 = 360$ суток к стандартному отклонению этой разности.

F_3 – поле ротора скорости деформаций.

F_4 – сглаженное поле изменений ротора скорости деформаций. Значения поля аналогичны значениям поля F_2 .

F_5 – поле пространственных корреляций значений полей изменений дивергенции F_2 и изменений ротора и F_4 скорости деформаций в окне 150×150 км.

Одноэтапный прогноз. При прогнозе на каждом интервале тестирования вычисляются зоны тревоги, в которых ожидаются все землетрясения интервала. Для прогноза использовались близкие к максимальным значения полей F_2 , S_1 и близкие к минимальным значения поля S_4 . Параметры цилиндра предвестника составляют $R = 8$ км, $T = 31$ суток.

На рис. 4 показаны графики зависимостей $U(V)$ и $U(W)$. На графиках можно видеть, что при порогах $V = 0,15$ и $W = 0,22$ на $M^* = 8$ интервалах из $M = 10$ в зону тревоги попали все эпицентры землетрясений.

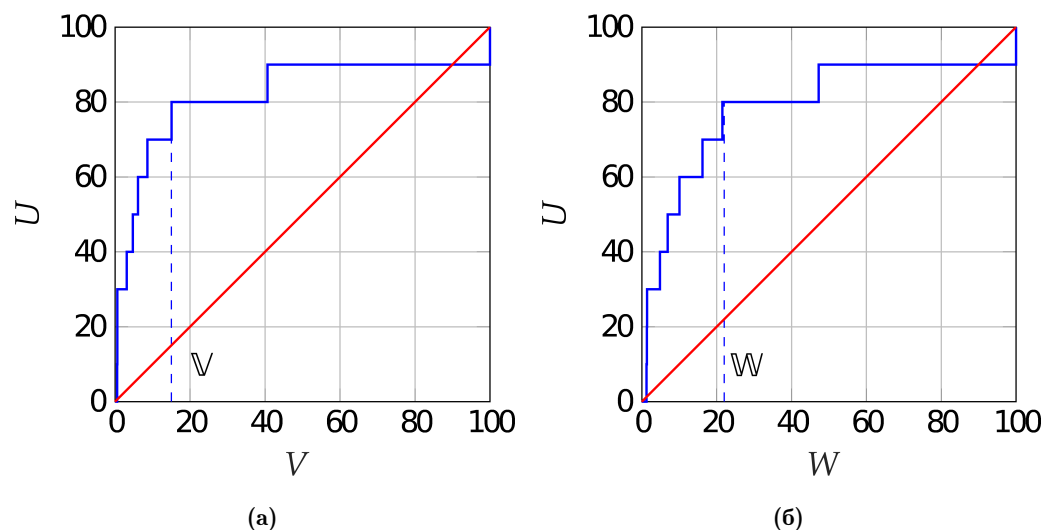


Рис. 4. Графики зависимостей $U(V)$ (слева) и $U(W)$ (справа) для региона Калифорнии. Значения U , V и W даны в процентах.

Двухэтапный прогноз. На первом этапе прогнозируются целевые землетрясения с эпицентрами в зоне анализа. Функция, прогнозирующая интервалы тревоги, определяется по значениям полей S_2 и F_2 , близким к максимальным. Параметрами цилиндра предвестника являются: радиус $R = 6$ км и образующая $T = 31$ суток.

На втором этапе зоны тревоги вычисляются только на интервалах тревоги. Как и в случае пространственно-временного прогноза мы использовали значения полей F_2 ,

S_1 , близкие к максимальным, и значения поля S_4 , близкие к минимальным. Параметры цилиндра предвестника составляют $R = 8$ км, $T = 61$ суток.

Островная часть 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 $m \geq 2,4$ и глубинами гипоцентров $H \leq 160$, взятыми из каталога Japan Meteorological Agency [Obara et al., 2005; Okada et al., 2014]. 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> [Blewitt et al., 2018]. Средние минимальные расстояния между 1229 приёмными станциями GPS составляют 12,8 км, стандартное отклонения этих значений составляет 5,4 км. Обучение начинается с 02.03.2011. Тестирование выполняется от 10.04.2015 до 28.04.2023 с интервалом 30 суток. За время тестирования на $M = 17$ интервалах из $N = 99$ интервалов прогноза произошли $Q = 20$ целевых землетрясений с эпицентрами в зоне анализа. Зона анализа и эпицентры тестируемых землетрясений показаны на рис. 5.

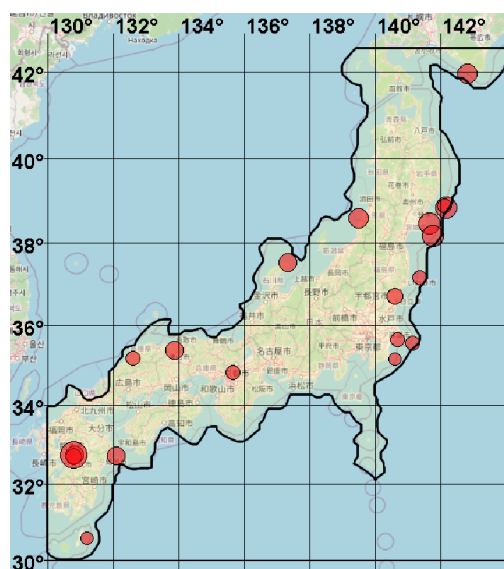


Рис. 5. Зона анализа и эпицентры тестовых землетрясений на интервале 10.04.2015 до 23.04.2023 с магнитудами $m \geq 6,0$ и глубинами эпицентров $H \leq 60$ км.

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. В отличие от Калифорнии данные GPS оказались менее информативными. Возможная причина этого состоит в более глубоких очагах землетрясений.

S_1 – поле плотности эпицентров землетрясений, вычисленное с помощью метода гауссовского ядерного сглаживания с радиусом $R_0 = 30$ км, временным интервалом $T_0 = 60$ суток.

S_2 – поле значений b -value, вычисленное аналогично полю S_1 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. 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.

Одноэтапный прогноз. При прогнозе на каждом интервале тестирования вычислялись зоны тревоги, в которых ожидалось все землетрясения интервала. Для прогноза использовались близкие к максимальным значения поля S_1 . Параметры цилиндра предвестника составляют $R = 8$ км, $T = 61$ суток.

На рис. 6 показаны графики зависимостей $U(V)$ и $U(W)$. На графиках можно видеть, что при порогах $V = 0,2$ и $W = 0,23$ на $M^* = 14$ интервалах из $M = 17$ в зону тревоги попали все эпицентры землетрясений.

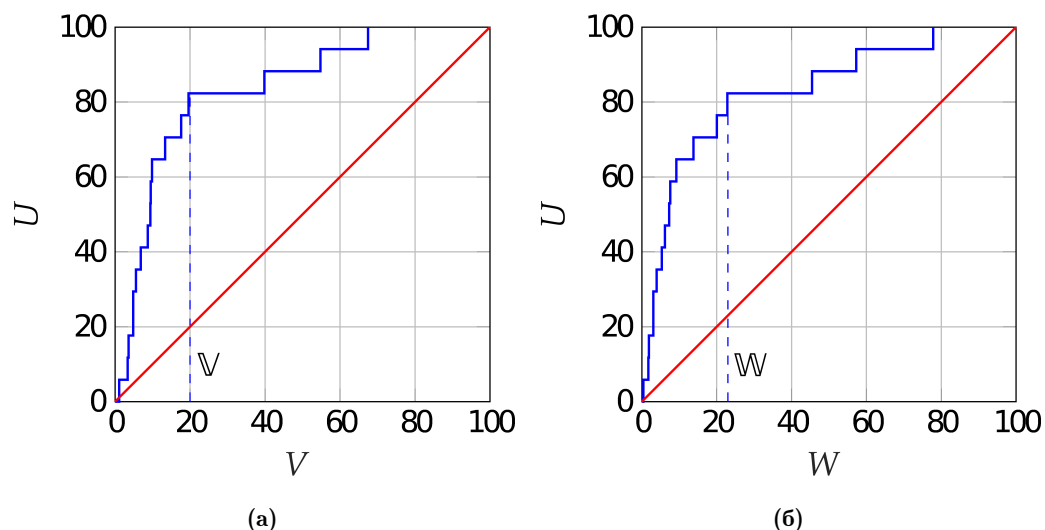


Рис. 6. Графики зависимостей $U(V)$ (слева) и $U(W)$ (справа) для региона Японии. Значения U , V и W даны в процентах.

Двухэтапный прогноз. На первом этапе прогнозируются целевые землетрясения с эпицентрами в зоне анализа. Функция, прогнозирующая интервалы тревоги, определяется по значениям близким к минимальным значениям поля S_2 и близким к максимальным значениям поля S_3 . Параметрами цилиндра предвестника являются: радиус $R = 16$ км и образующая $T = 61$. Тревога в узлах сетки продолжается $T = 61$.

На втором этапе 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.

3.2. Основные результаты моделирования

В одноэтапной схеме перед каждым прогнозом вычисляется зона тревоги. При двухэтапной схеме сначала принимается решение является ли очередной интервал прогноза интервалом тревоги, а затем на интервале тревоги вычисляется зона тревоги. Принятие решений осуществляется в результате сравнения вычисленного при обучении объёма тревоги V с порогом V .

Для оценивания качества прогнозов мы используем следующие показатели.

Одноэтапный прогноз

V – пороговая величина объёма тревоги по узлам сетки при обучении.

W – пороговая величина объёма тревоги по узлам сетки при тестировании.

N – число всех интервалов при тестировании.

M – число всех интервалов с эпицентрами в зоне анализа при тестировании.

M^* – число интервалов со всеми эпицентрами в зоне тревоги при тестировании.

Оценка вероятности успешного обнаружения всех эпицентров землетрясений в зоне тревоги:

$$U = \frac{M^*}{M}. \quad (6)$$

Оценка вероятности того, что на очередном интервале прогноза все эпицентры целевых землетрясений окажутся в зоне тревоги:

$$P = \frac{M^*}{N} \quad (7)$$

Оценка вероятности того, что на очередном интервале прогноза нет эпицентров целевых землетрясений (ложная тревога):

$$P_{\text{л. т.}} = 1 - \frac{M^*}{N} \quad (8)$$

Оценка вероятности того, что на очередном интервале прогноза хотя бы один эпицентр находится вне зоны тревоги (пропуск цели):

$$P_{\text{п. ц.}} = \frac{M - M^*}{N} \quad (9)$$

Двухэтапный прогноз

$\mathbb{V}_2$ – пороговая величина объёма тревоги по узлам сетки на интервалах тревоги при обучении.

$\mathbb{W}_2$ – пороговая величина объёма тревоги по узлам сетки на интервалах тревоги при тестировании.

N_1 – число всех интервалов тревоги при тестировании.

M_1 – число всех интервалов тревоги с эпицентрами в зоне анализа при тестировании.

M_2 – число интервалов тревоги с эпицентрами в зоне тревоги при тестировании.

Оценка условной вероятности успешного обнаружения на интервале тревоги всех эпицентров землетрясений в зоне тревоги:

$$U_2 = \frac{M_2}{M_1}. \quad (10)$$

Оценка условной вероятности того, что на очередном интервале тревоги все эпицентры целевых землетрясений окажутся в зоне тревоги:

$$P_2 = \frac{M_2}{N_1} \quad (11)$$

Оценка вероятности того, что на очередном интервале тревоги нет эпицентров целевых землетрясений (ложная тревога):

$$P_{2 \text{ л. т.}} = 1 - \frac{M_1}{N_1} \quad (12)$$

Оценка вероятности того, что на очередном интервале прогноза хотя бы один эпицентр целевого землетрясения находится вне зоны тревоги (пропуск цели):

$$P_{2 \text{ п. ц.}} = \frac{M - M_2}{N} \quad (13)$$

Показатели качества прогноза землетрясений в регионах Камчатки, Калифорнии и Японии сведены в [табл. 1](#).

Из [табл. 1](#) видно, что для одноэтапной и двухэтапной схем прогноза оценки вероятности успешного обнаружения эпицентров всех целевых землетрясений в зонах тревоги являются удовлетворительными и составляют от 0,75 до 0,85. Оценки вероятности того, что на очередном интервале эпицентры всех целевых землетрясений попадут в зоны тревоги при одноэтапном прогнозе равны $P = 0,075 - 0,14$. Оценки условной вероятности

Таблица 1. Показатели качества прогноза

Метод	Показатель	Камчатка	Калифорния	Япония
Одноэтапный прогноз	Число интервалов прогноза при тестировании	128	107	99
	Число интервалов с эпицентрами в зоне анализа при тестировании	17	10	17
	Число интервалов со всеми эпицентрами в зоне тревоги при тестировании	13	8	14
	Оценка вероятности успешного обнаружения всех эпицентров в зоне тревоги (6)	0,76	0,8	0,82
	Оценка вероятности того, что на очередном интервале все эпицентры окажутся в зоне тревоги (7)	0,1	0,075	0,14
	Оценка вероятности того, что на очередном интервале в зоне тревоги нет эпицентров, ложная тревога (8)	0,9	0,925	0,86
	Оценка вероятности того, что на очередном интервале хотя бы один эпицентр вне зоны тревоги, пропуск цели (9)	0,03	0,019	0,03
Двухэтапный прогноз	Число интервалов тревоги при тестировании	50	30	36
	Число интервалов тревоги со всеми эпицентрами при тестировании	13	8	14
	Число интервалов тревоги со всеми эпицентрами в зоне тревоги при тестировании	11	6	11
	Оценка условной вероятности успешного обнаружения на интервале тревоги всех эпицентров в зоне тревоги (10)	0,85	0,75	0,85
	Оценка условной вероятности того, что на очередном интервале тревоги все эпицентры окажутся в зоне тревоги (11)	0,22	0,2	0,31
	Оценка вероятности того, что на очередном интервале тревоги нет эпицентров, ложная тревога (12)	0,74	0,73	0,61
	Оценка вероятности того, что на очередном интервале хотя бы один эпицентр вне зоны тревоги, пропуск цели (13)	0,05	0,04	0,06

сти того, что на очередном интервале тревоги эпицентры всех целевых землетрясений попадут в зоны тревоги при двухэтапном прогнозе равны $P_2 = 0,2 - 0,31$.

4. Дискуссия

Оценивание вероятности успешного прогноза землетрясений. На первом этапе перед очередным прогнозом землетрясений по результатам обучения известны зависимости вероятности обнаружения интервалов тревоги от объёмов тревоги $U_1(V_1)$ и $U_1(W_1)$.

Пусть эти зависимости совпадают с графиками рис. 7, полученными при моделировании прогноза землетрясений на Камчатке. Допустим, что значение прогнозируемого объёма тревоги на очередной интервал прогноза $V_1 = 0,1$. После этого алгоритм метода минимальной области тревоги принимает решение, что очередной интервал прогноза является интервалом тревоги, и все интервалы прогноза со значениями $V_1 \leq \bar{V} = 0,1$ являются интервалами тревоги. По зависимостям $U_1(V_1)$ и $U_1(W_1)$, показанным на графиках рис. 7, получаем, что порогу по объёму тревоги $V_1 = 0,1$ соответствуют значения $M_1 = 9$ и $W_1 = 0,19$. Отсюда согласно определению объёма тревоги получаем, что число интервалов тревоги на тестовых данных

$$N_1 = W_1 N = 0,19 \cdot 128 = 24$$

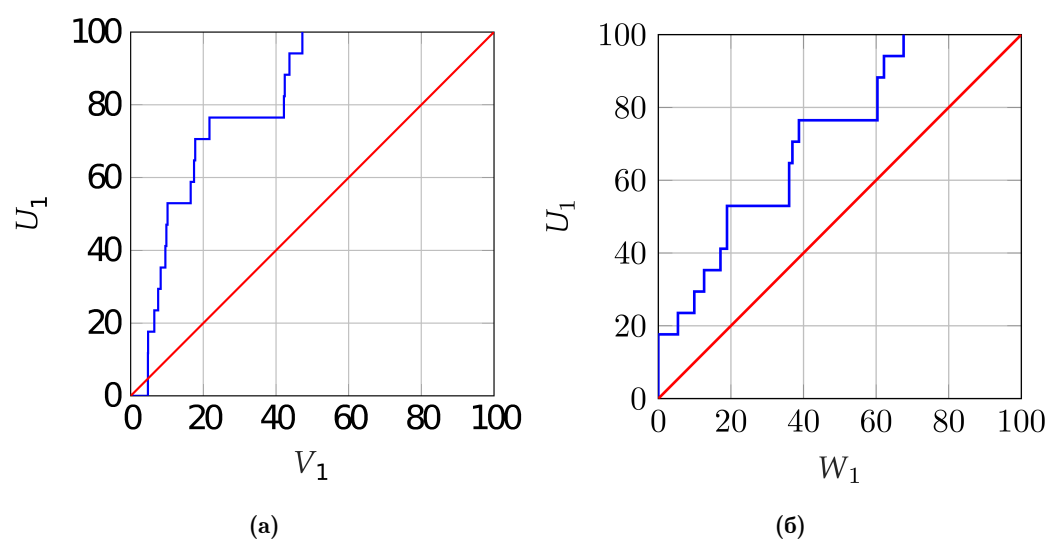


Рис. 7. Графики зависимостей $U_1(V_1)$ (слева) и $U_1(W_1)$ (справа) для региона Камчатки. Значения U_1 , V_1 и W_1 даны в процентах.

Рассматриваемый интервал прогноза является интервалом тревоги. Поэтому для него и для всех интервалов тревоги, для которых объём тревоги $V_1 \leq \bar{V} = 0,1$, алгоритм выполняет второй этап прогноза: вычисление на интервалах тревоги зон тревоги. Графики зависимостей $U_2(V_2)$ и $U_2(W_2)$ показаны на рис. 8. Из графика $U_2(W_2)$ можно видеть, что зоны тревоги, которые содержат все эпицентры землетрясений, алгоритм построил на $M_2 = 8$ интервалах тревоги. Следовательно, если интервалы тревоги выбираются при пороге объёма тревоги $\bar{V} = 0,1$, то оценка вероятности того, что все эпицентры целевых землетрясений на очередном интервале тревоги окажутся в зоне тревоги (успешный прогноз землетрясений) равна

$$P_5 = \frac{M_2}{N_1} = \frac{8}{24} = 0,33.$$

Проблема малых значений вероятности успеха очередного прогноза землетрясений. Приведённые в разд. 3.2 результаты моделирования показывают, что оценки условной вероятности P_2 того, что на очередном интервале тревоги все эпицентры целевых землетрясений окажутся в зоне тревоги, увеличились по сравнению с оценками вероятности P для одноэтапного прогноза в 2–2,5 раза. Однако их значения малы для формального принятия практических решений о времени и месте наступления сейсмической опасности.

Малые значения вероятности успеха очередного прогноза целевых землетрясений частично обусловлены самой постановкой задачи. В задачах классификации обычно

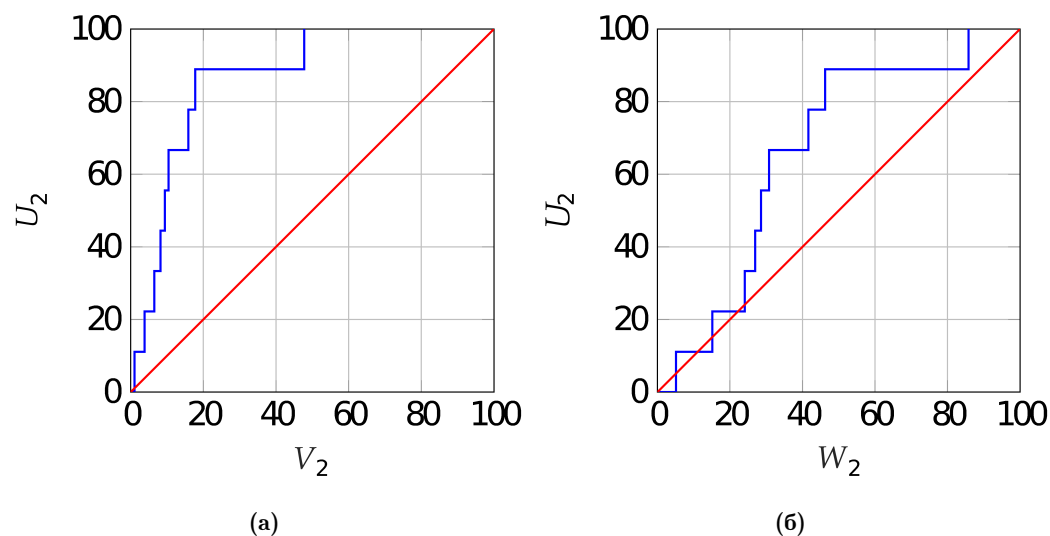


Рис. 8. Графики зависимостей $U_2(V_2)$ (слева) и $U_2(W_2)$ (справа) для региона Камчатки. Значения U_2 , V_2 и W_2 даны в процентах.

существуют объективные причины, по которым объекты одного класса являются сходными между собой, а объекты разных классов 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 качественных различий в физике сейсмического процесса, исходные данные не содержат физических свойств, которые определяют межклассовые различия для землетрясений с близкими магнитудами. Это обстоятельство в значительной степени осложняет задачу прогноза.

Вторая причина связана с прогнозом целевых землетрясений с эпицентрами в зоне анализа на первом этапе. Условная вероятность того, что на очередном интервале тревоги все эпицентры целевых землетрясений попадут в зоны тревоги, оценивается по формуле (10). Можно видеть, что успешное выделение интервалов тревоги имеет место, если число интервалов тревоги N_1 и число интервалов тревоги с эпицентрами в зоне тревоги M_2 приближаются к M . Сложность этой задачи состоит в том, что аномальные значения сейсмических параметров (предвестники землетрясений) часто проявляются за несколько интервалов прогноза до целевого землетрясения и эти аномалии заканчиваются через несколько интервалов после землетрясения. Поэтому алгоритм выделяет «лишние» интервалы тревоги.

Возможности практического использования технологии систематического прогноза землетрясений. Перед вычислением реального прогноза землетрясений должно выполняться машинное обучение на интервале 20–30 лет. Статистические показатели позволяют использовать результаты полученного компьютерного прогноза как сигнал к проведению более детальных исследований. Для выполнения этой схемы прогноза разработана сетевая платформа мониторинга сейсмогенных процессов, систематического прогноза землетрясений и проведения исследований для поддержки принятия решений по прогнозу сейсмической опасности [Gitis and Derendyaev, 2023]. Платформа поддерживает два уровня анализа данных:

Уровень А. Мониторинг сейсмогенных процессов и систематический прогноз землетрясений в автоматическом режиме. В разд. 2 показано, каждая прогнозируемая зона тревоги может сопровождаться тестовым объяснением. Оно может состоять из логического имплицативного высказывания и списка предшествующих целевых землетрясений, предвестники которых имели значения полей, аналогичные значениям узлов сетки зоны тревоги

Уровень Б. Детальное исследование гипотез по прогнозу землетрясений и идентификации аномальных геологических зон, которые специалист может сформулировать на уровне А.

Прототип платформы доступен по адресу <https://gis.iitp.ru/prognosis-gps/>. Он реализует одноэтапную схему прогноза.

Заключение

Рассматриваемая в статье технология систематического прогноза землетрясений реализует следующие три новых качества.

В публикациях вероятность успешного прогноза землетрясений обычно оценивается как отношение числа обнаруженных эпицентров целевых землетрясений к числу всех эпицентров целевых землетрясений. В рассмотренной технологии вероятность успешного прогноза землетрясений U оценивается как отношение числа интервалов прогноза на которых обнаружены все эпицентры целевых землетрясений к числу всех интервалов прогноза с эпицентрами целевых землетрясений в зоне анализа.

В публикациях качество прогноза обычно определяется по оценке вероятности успешного обнаружения землетрясений в серии прогнозов. Этой оценки недостаточно для практического использования результатов прогноза. В рассмотренной технологии добавлена вторая оценка качества прогноза: вероятность того, что на очередном интервале прогноза все эпицентры целевых землетрясений попадут в вычисленную зону тревоги.

Рассмотренная технология позволяет оценить вероятность того, что все эпицентры целевых землетрясений на очередном интервале прогноза попадут в вычисленную зону тревоги.

Развитие подхода и методов систематического прогноза землетрясений возможно по следующим трем направлениям: 1 – Дополнение типов исходных данных. 2 – Разработка новых методов обработки исходных данных. 3 – Разработка новых методов машинного обучения прогнозу редких событий.

Благодарности. Работа частично поддержана Государственным заданием № 0061-2022-0003 по теме «Исследование и разработка методов и сетевых технологий анализа больших пространственно-временных данных с целью междисциплинарного анализа и прогнозирования природных и социально-экономических процессов». Работа выполнена с использованием данных Камчатского филиала ФИЦ «Единая Геофизическая служба РАН», National Earthquake Information Center (NEIC), Japan Meteorological Agency (JMA) и Nevada Geodetic Laboratory (NGL).

Список литературы

- Завьялов А. Д. Среднесрочный прогноз землетрясений: основы, методика, реализация. — Москва : Наука, 2006. — EDN: [QKFPZD](#).
- Молчан Г. М. Оптимальные стратегии в прогнозе землетрясений. Современные методы интерпретации сейсмологических данных // Вычислительная сейсмология. — 1991. — Т. 24. — С. 3–18.
- Соболев Г. А. Основы прогноза землетрясений. — Москва : МАИК «Наука/Интерпериодика», 1993. — 313 с. — EDN: [TGSIGN](#).
- Соболев Г. А., Пономарев А. В. Физика землетрясений и предвестники. — Москва : МАИК «Наука/Интерпериодика», 2003. — 270 с. — EDN: [RVEBFL](#).
- Чебров В. Н., Дроздин Д. В., Кугаенко Ю. А. и др. Система детальных сейсмологических наблюдений на Камчатке в 2011 г. // Вулканология и сейсмология. — 2013. — № 1. — С. 18–40. — DOI: [10.7868/s0203030613010021](https://doi.org/10.7868/s0203030613010021).
- Чеброва А. Ю., Чемаев А. С., Матвеев Е. А. и др. Единая информационная система сейсмологических данных в Камчатском филиале ФИЦ ЕГС РАН: Принципы организации, основные элементы, ключевые функции // Геофизические исследования. — 2020. — Т. 21, № 3. — С. 66–91. — DOI: [10.21455/gr2020.3-5](https://doi.org/10.21455/gr2020.3-5).
- Amei 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. — Vol. 2, no. 7.

- Asim K. M., Idris A., Iqbal T., et al. Earthquake prediction model using support vector regressor and hybrid neural networks // PLOS ONE. — 2018. — Vol. 13, no. 7. — DOI: [10.1371/journal.pone.0199004](https://doi.org/10.1371/journal.pone.0199004).
- Barnhart W. D., Hayes G. P., Wald D. J. Global Earthquake Response with Imaging Geodesy: Recent Examples from the USGS NEIC // Remote Sensing. — 2019. — Vol. 11, no. 11. — P. 1357. — DOI: [10.3390/rs11111357](https://doi.org/10.3390/rs11111357).
- Blewitt G., Hammond W. C., Kreemer C. Harnessing the GPS Data Explosion for Interdisciplinary Science // Eos. — 2018. — Vol. 99. — P. 485. — DOI: [10.1029/2018eo104623](https://doi.org/10.1029/2018eo104623).
- Bradley A. P. The use of the area under the ROC curve in the evaluation of machine learning algorithms // Pattern Recognition. — 1997. — Vol. 30, no. 7. — P. 1145–1159. — DOI: [10.1016/s0031-3203\(96\)00142-2](https://doi.org/10.1016/s0031-3203(96)00142-2).
- Corbi F., Sandri L., Bedford J., et al. Machine Learning Can Predict the Timing and Size of Analog Earthquakes // Geophysical Research Letters. — 2019. — Vol. 46, no. 3. — P. 1303–1311. — DOI: [10.1029/2018gl081251](https://doi.org/10.1029/2018gl081251).
- Gitis V., Derendyaev A. The Method of the Minimum Area of Alarm for Earthquake Magnitude Prediction // Frontiers in Earth Science. — 2020. — Vol. 11. — DOI: [10.3389/feart.2020.585317](https://doi.org/10.3389/feart.2020.585317).
- Gitis V., Derendyaev A. A Technology for Seismogenic Process Monitoring and Systematic Earthquake Forecasting // Remote Sensing. — 2023. — Vol. 15, no. 8. — P. 2171. — DOI: [10.3390/rs15082171](https://doi.org/10.3390/rs15082171).
- Gitis V., Derendyaev A., Petrov K. Analyzing the Performance of GPS Data for Earthquake Prediction // Remote Sensing. — 2021. — Vol. 13, no. 9. — P. 1842. — DOI: [10.3390/rs13091842](https://doi.org/10.3390/rs13091842).
- Gitis V. G., Derendyaev A. B., Pirogov S. A., et al. Adaptive estimation of seismic parameter fields from earthquake catalogs // Journal of Communications Technology and Electronics. — 2015. — Vol. 60, no. 12. — P. 1459–1465. — DOI: [10.1134/s1064226915120098](https://doi.org/10.1134/s1064226915120098).
- Gitis V. G., Derendyaev A. B., Pirogov S. A., 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. — P. 1–8. — DOI: [10.1145/3102254.3102269](https://doi.org/10.1145/3102254.3102269).
- Kagan Y. Y. Earthquakes: Models, Statistics, Testable Forecasts. — Wiley, 2013. — DOI: [10.1002/9781118637913](https://doi.org/10.1002/9781118637913).
- Kail R., Burnaev E., Zaytsev A. Recurrent Convolutional Neural Networks Help to Predict Location of Earthquakes // IEEE Geoscience and Remote Sensing Letters. — 2022. — Vol. 19. — P. 1–5. — DOI: [10.1109/lgrs.2021.3107998](https://doi.org/10.1109/lgrs.2021.3107998).
- Kaplan S. M., McFall R. W. The Statistical Properties of Noise Applied to Radar Range Performance // Proceedings of the IRE. — 1951. — Vol. 39, no. 1. — P. 56–60. — DOI: [10.1109/jrproc.1951.230422](https://doi.org/10.1109/jrproc.1951.230422).
- King C. Y. Gas geochemistry applied to earthquake prediction: An overview // Journal of Geophysical Research: Solid Earth. — 1986. — Vol. 91, B12. — P. 12269–12281. — DOI: [10.1029/jb091ib12p12269](https://doi.org/10.1029/jb091ib12p12269).
- Kossobokov V., Shebalin P. Earthquake Prediction // Nonlinear Dynamics of the Lithosphere and Earthquake Prediction. — Berlin : Springer Berlin Heidelberg, 2003. — P. 141–207. — DOI: [10.1007/978-3-662-05298-3_4](https://doi.org/10.1007/978-3-662-05298-3_4).
- Lighthill J. A Critical Review of VAN: Earthquake Prediction from Seismic Electrical Signals. — Singapore : World Scientific, 1996. — DOI: [10.1142/3006](https://doi.org/10.1142/3006).
- Mignan A., Broccardo M. Neural Network Applications in Earthquake Prediction (1994-2019): Meta-Analytic and Statistical Insights on Their Limitations // Seismological Research Letters. — 2020. — Vol. 91, no. 4. — P. 2330–2342. — DOI: [10.1785/0220200021](https://doi.org/10.1785/0220200021).
- Molchan G. Space-Time Earthquake Prediction: The Error Diagrams // Pure and Applied Geophysics. — 2010. — Vol. 167, no. 8/9. — P. 907–917. — DOI: [10.1007/s00024-010-0087-z](https://doi.org/10.1007/s00024-010-0087-z).
- Obara K., Kasahara K., Hori S., 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. — Vol. 76, no. 2. — DOI: [10.1063/1.1854197](https://doi.org/10.1063/1.1854197).
- Okada Y., Kasahara K., Hori S., et al. Recent progress of seismic observation networks in Japan -Hi-net, F-net, K-NET and KiK-net- // Earth, Planets and Space. — 2014. — Vol. 56, no. 8. — P. 15–28. — DOI: [10.1186/BF03353076](https://doi.org/10.1186/BF03353076).
- Panakkat A., Adeli H. Neural network models for earthquake magnitude prediction using multiple seismicity indicators // International Journal of Neural Systems. — 2007. — Vol. 17, no. 01. — P. 13–33. — DOI: [10.1142/s0129065707000890](https://doi.org/10.1142/s0129065707000890).
- Rhoades D. A. Mixture Models for Improved Earthquake Forecasting with Short-to-Medium Time Horizons // Bulletin of the Seismological Society of America. — 2013. — Vol. 103, no. 4. — P. 2203–2215. — DOI: [10.1785/0120120233](https://doi.org/10.1785/0120120233).
- Shebalin P. N., Narteau C., Zecher J. D., et al. Combining earthquake forecasts using differential probability gains // Earth, Planets and Space. — 2014. — Vol. 66, no. 1. — P. 1–14. — DOI: [10.1186/1880-5981-66-37](https://doi.org/10.1186/1880-5981-66-37).
- Soloviev A. A., Gvishiani A. D., Gorshkov A. I., et al. Recognition of earthquake-prone areas: Methodology and analysis of the results // Izvestiya, Physics of the Solid Earth. — 2014. — Vol. 50, no. 2. — P. 151–168. — DOI: [10.1134/s1069351314020116](https://doi.org/10.1134/s1069351314020116).

THE FUNDAMENTALS OF A TWO-STAGE APPROACH TO SYSTEMATIC EARTHQUAKE PREDICTION

V. G. Gitis¹, A. B. Derendyaev^{**1}

Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute), Moscow, Russia

****Correspondence to:** Alexander Derendyaev, wintsa@gmail.com

A systematic earthquake prediction is performed regularly at fixed intervals within a preselected seismically homogeneous zone. The result of each prediction iteration is a map highlighting the alarm zones, where the epicenters of target earthquakes are expected. The proposed methodology introduces the following innovations: 1 – A prediction is considered successful if all epicenters of the target earthquakes during the forecast interval fall within the alarm zone. 2 – The methodology optimizes both the probability of successfully detecting earthquake epicenters across a series of forecasts and the success rate of predictions in each individual iteration. 3 – The methodology enables the estimation of the probability of success for the next forecast interval. Examples of the method's application are demonstrated for predicting earthquakes in Kamchatka, California, and the island region of Japan.

Keywords: systematic earthquake prediction, machine learning, the method of the minimum area of alarm, GPS time series.

Received: 3 September 2024

Accepted: 15 January 2025

Published: 21 May 2025



© 2025. The Authors.

Citation: Gitis, V. G., A. B. Derendyaev (2025), The Fundamentals of a Two-Stage Approach to Systematic Earthquake Prediction, *Russian Journal of Earth Sciences*, 25, ES3010, <https://doi.org/10.2205/2025ES000987>, EDN: MFJOOB

References

- Amei 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. — Vol. 2, no. 7.
- Asim K. M., Idris A., Iqbal T., et al. Earthquake prediction model using support vector regressor and hybrid neural networks // *PLOS ONE*. — 2018. — Vol. 13, no. 7. — DOI: [10.1371/journal.pone.0199004](https://doi.org/10.1371/journal.pone.0199004).
- Barnhart W. D., Hayes G. P., Wald D. J. Global Earthquake Response with Imaging Geodesy: Recent Examples from the USGS NEIC // *Remote Sensing*. — 2019. — Vol. 11, no. 11. — P. 1357. — DOI: [10.3390/rs11111357](https://doi.org/10.3390/rs11111357).
- Blewitt G., Hammond W. C., Kreemer C. Harnessing the GPS Data Explosion for Interdisciplinary Science // *Eos*. — 2018. — Vol. 99. — P. 485. — DOI: [10.1029/2018eo104623](https://doi.org/10.1029/2018eo104623).
- Bradley A. P. The use of the area under the ROC curve in the evaluation of machine learning algorithms // *Pattern Recognition*. — 1997. — Vol. 30, no. 7. — P. 1145–1159. — DOI: [10.1016/s0031-3203\(96\)00142-2](https://doi.org/10.1016/s0031-3203(96)00142-2).
- Chebrov V. N., Droznin D. V., Kugaenko Yu. A., et al. The system of detailed seismological observations in Kamchatka in 2011 // *Journal of Volcanology and Seismology*. — 2013. — Vol. 7, no. 1. — P. 16–36. — DOI: [10.1134/S0742046313010028](https://doi.org/10.1134/S0742046313010028).
- Chebrova A. Yu., Chemarev A. S., Matveenko E. A., et al. Seismological data information system in Kamchatka branch of GS RAS: Organization principles, main elements and key functions // *Geophysical Research*. — 2020. — Vol. 21, no. 3. — P. 66–91. — DOI: [10.21455/gr2020.3-5](https://doi.org/10.21455/gr2020.3-5). — (In Russian).
- Corbi F., Sandri L., Bedford J., et al. Machine Learning Can Predict the Timing and Size of Analog Earthquakes // *Geophysical Research Letters*. — 2019. — Vol. 46, no. 3. — P. 1303–1311. — DOI: [10.1029/2018gl081251](https://doi.org/10.1029/2018gl081251).
- Gitis V., Derendyaev A. The Method of the Minimum Area of Alarm for Earthquake Magnitude Prediction // *Frontiers in Earth Science*. — 2020. — Vol. 11. — DOI: [10.3389/feart.2020.585317](https://doi.org/10.3389/feart.2020.585317).
- Gitis V., Derendyaev A. A Technology for Seismogenic Process Monitoring and Systematic Earthquake Forecasting // *Remote Sensing*. — 2023. — Vol. 15, no. 8. — P. 2171. — DOI: [10.3390/rs15082171](https://doi.org/10.3390/rs15082171).

- Gitis V., Derendyaev A., Petrov K. Analyzing the Performance of GPS Data for Earthquake Prediction // Remote Sensing. — 2021. — Vol. 13, no. 9. — P. 1842. — DOI: [10.3390/rs13091842](https://doi.org/10.3390/rs13091842).
- Gitis V., Derendyaev A. B., Pirogov S. A., et al. Adaptive estimation of seismic parameter fields from earthquake catalogs // Journal of Communications Technology and Electronics. — 2015. — Vol. 60, no. 12. — P. 1459–1465. — DOI: [10.1134/s1064226915120098](https://doi.org/10.1134/s1064226915120098).
- Gitis V., Derendyaev A. B., Pirogov S. A., 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. — P. 1–8. — DOI: [10.1145/3102254.3102269](https://doi.org/10.1145/3102254.3102269).
- Kagan Y. Y. Earthquakes: Models, Statistics, Testable Forecasts. — Wiley, 2013. — DOI: [10.1002/9781118637913](https://doi.org/10.1002/9781118637913).
- Kail R., Burnaev E., Zaytsev A. Recurrent Convolutional Neural Networks Help to Predict Location of Earthquakes // IEEE Geoscience and Remote Sensing Letters. — 2022. — Vol. 19. — P. 1–5. — DOI: [10.1109/lgrs.2021.3107998](https://doi.org/10.1109/lgrs.2021.3107998).
- Kaplan S. M., McFall R. W. The Statistical Properties of Noise Applied to Radar Range Performance // Proceedings of the IRE. — 1951. — Vol. 39, no. 1. — P. 56–60. — DOI: [10.1109/jrproc.1951.230422](https://doi.org/10.1109/jrproc.1951.230422).
- King C. Y. Gas geochemistry applied to earthquake prediction: An overview // Journal of Geophysical Research: Solid Earth. — 1986. — Vol. 91, B12. — P. 12269–12281. — DOI: [10.1029/jb091ib12p12269](https://doi.org/10.1029/jb091ib12p12269).
- Kossobokov V., Shebalin P. Earthquake Prediction // Nonlinear Dynamics of the Lithosphere and Earthquake Prediction. — Berlin : Springer Berlin Heidelberg, 2003. — P. 141–207. — DOI: [10.1007/978-3-662-05298-3_4](https://doi.org/10.1007/978-3-662-05298-3_4).
- Lighthill J. A Critical Review of VAN: Earthquake Prediction from Seismic Electrical Signals. — Singapore : World Scientific, 1996. — DOI: [10.1142/3006](https://doi.org/10.1142/3006).
- Mignan A., Broccardo M. Neural Network Applications in Earthquake Prediction (1994-2019): Meta-Analytic and Statistical Insights on Their Limitations // Seismological Research Letters. — 2020. — Vol. 91, no. 4. — P. 2330–2342. — DOI: [10.1785/0220200021](https://doi.org/10.1785/0220200021).
- Molchan G. Space-Time Earthquake Prediction: The Error Diagrams // Pure and Applied Geophysics. — 2010. — Vol. 167, no. 8/9. — P. 907–917. — DOI: [10.1007/s00024-010-0087-z](https://doi.org/10.1007/s00024-010-0087-z).
- Molchan G. M. Optimal Strategies in Earthquake Forecasting. Modern Methods of Interpreting Seismological Data // Computational Seismology. — 1991. — Vol. 24. — P. 3–18. — (In Russian).
- Obara K., Kasahara K., Hori S., 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. — Vol. 76, no. 2. — DOI: [10.1063/1.1854197](https://doi.org/10.1063/1.1854197).
- Okada Y., Kasahara K., Hori S., et al. Recent progress of seismic observation networks in Japan -Hi-net, F-net, K-NET and KiK-net- // Earth, Planets and Space. — 2014. — Vol. 56, no. 8. — P. 15–28. — DOI: [10.1186/BF03353076](https://doi.org/10.1186/BF03353076).
- Panakkat A., Adeli H. Neural network models for earthquake magnitude prediction using multiple seismicity indicators // International Journal of Neural Systems. — 2007. — Vol. 17, no. 01. — P. 13–33. — DOI: [10.1142/s0129065707000890](https://doi.org/10.1142/s0129065707000890).
- Rhoades D. A. Mixture Models for Improved Earthquake Forecasting with Short-to-Medium Time Horizons // Bulletin of the Seismological Society of America. — 2013. — Vol. 103, no. 4. — P. 2203–2215. — DOI: [10.1785/0120120233](https://doi.org/10.1785/0120120233).
- Shebalin P. N., Narteau C., Zechar J. D., et al. Combining earthquake forecasts using differential probability gains // Earth, Planets and Space. — 2014. — Vol. 66, no. 1. — P. 1–14. — DOI: [10.1186/1880-5981-66-37](https://doi.org/10.1186/1880-5981-66-37).
- Sobolev G. A. Fundamentals of Earthquake Forecasting. — Moscow : MAIK “Nauka/Interperiodica”, 1993. — P. 313. — EDN: [TGSIGH](https://www.edn.ru/10.1142/3006) ; (in Russian).
- Sobolev G. A., Ponomarev A. V. Physics of earthquakes and precursors. — Moscow : MAIK “Nauka/Interperiodica”, 2003. — P. 270. — EDN: [RVEBFL](https://www.edn.ru/10.1142/3006) ; (in Russian).
- Soloviev A. A., Gvishiani A. D., Gorshkov A. I., et al. Recognition of earthquake-prone areas: Methodology and analysis of the results // Izvestiya, Physics of the Solid Earth. — 2014. — Vol. 50, no. 2. — P. 151–168. — DOI: [10.1134/s1069351314020116](https://doi.org/10.1134/s1069351314020116).
- Zavyalov A. D. Medium-term earthquake forecast: fundamentals, methodology, implementation. — Moscow : Nauka, 2006. — EDN: [QKFPZD](https://www.edn.ru/10.1142/3006) ; (in Russian).