

Секция 4 «Моделирование геофизических процессов и полей»

ПРИМЕНЕНИЕ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ – МЕТОДА МОНТЕ-КАРЛО ПРИ АНАЛИЗЕ СЕЙСМОИОНОСФЕРНЫХ ЭФФЕКТОВ

Е.В. Липеровская

Институт физики Земли РАН

Ways of salvation of a number of model search problems of earthquake effect-“precursor” are suggested for different time series variations: time series with normal distribution law, time series with non-normal distribution law, time series with clearly defined periodicity. Different alternatives of series normalization are discussed. A precursor is detected by the method of epoch superposition. The validity of the discovered effect is evaluated by the method of background distribution modeling (Monte-Karlo method). A model problem of evaluation of part of an “earthquake”, having clear “effect-precursor”, is solved for different amplitudes of “seismo-impact”.

Введение

При исследовании эмпирических закономерностей каких-либо процессов неизбежно возникает необходимость математической обработки результатов измерений, в частности, если мы изучаем предвестники землетрясений, имея в распоряжении временные ряды каких-либо измеряемых величин. Следует отметить, что аномальные явления – предвестники, т.е. специфические явления, или вариации с очень большой амплитудой, которые перед землетрясениями наблюдаются, а без землетрясений не наблюдаются, встречаются редко. Поэтому исследователи ищут неспецифические вариации измеряемых величин, которые могли бы быть вызваны процессами подготовки землетрясений [1]. С процессами подготовки пытаются связывать вариации, которые чаще встречаются перед землетрясениями, чем в остальное, так называемое «фоновое» время. Что подразумевают, когда говорят «чаще встречаются перед землетрясениями»? За какое время перед землетрясениями? И перед всеми событиями, или только перед некоторыми?

Обсудим эти вопросы.

В настоящее время принято различать долгосрочные предвестники – за годы и десятки лет, среднесрочные – за месяцы, краткосрочные и оперативные – возникающие за время от нескольких недель до часов.

Последние исследования сейсмологических предвестников, проведенные М.В.Родкиным – сотрудником Геофизического Центра РАН в Москве – показали, что так называемые долгосрочные и среднесрочные предвестники скорее характеризуют состояние среды, в которой может произойти землетрясение, чем непосредственно процесс подготовки землетрясения, в частности? ожидаемую магнитуду и интервал времени между появлением предвестника и моментом землетрясения [2]. Соответственно, случае (в согласии с распространенным на Западе мнением о принципиальной непредсказуемости землетрясений) так называемые долгосрочные и среднесрочные предвестники на самом деле (строго говоря) предвестниками не являются. Поэтому тем большее значение приобретает изучение краткосрочных и оперативных предвестников землетрясений.

Здесь необходимо подчеркнуть, что **речь идет не о прогнозе землетрясений, а об изучении предвестников**. Полученные закономерности помогут построить адекватные физические модели сейсмического процесса и его связи с процессами на поверхности Земли, в атмосфере и ионосфере.

Итак, задача исследователя – достоверно выявить закономерности, возникающие перед землетрясениями.

Предположим, что у исследователя создалось впечатление, что за несколько дней перед землетрясением вариации некоторого измеряемого параметра возникают чаще, чем без землетрясений. Например, увеличивается (или уменьшается) значение измеряемой частоты, увеличивается вероятность наблюдения какого-то явления, изменяются спектральные характеристики изучаемого параметра. Соответственно возникает вопрос: связаны ли с подготовкой землетрясения наблюдаемые вариации, или они случайно возникли в это время. Или точнее, какова вероятность того, что наличие вариации перед землетрясениями не случайно. И здесь на помощь приходит статистический анализ.

До использования современной вычислительной техники вопросы статистического анализа решались в рамках аналитической статистики, представлявшей собой специальный и весьма развитый раздел математики.

В последние годы благодаря развитию вычислительной техники стало возможным проводить статистический анализ численно, даже не зная и не умея проводить соответствующие аналитические вычисления, путем статистического моделирования.

Такой численный анализ во-первых, не требует овладения аппаратом аналитической статистики, который достаточно специфичен и сложен. И, во-вторых, при правильном использовании численный подход более удобен, более гибок и может лучше отвечать условиям конкретной задачи, чем методы аналитической статистики.

При анализе данных наблюдений часто молчаливо используют закономерности, строго справедливые только для нормального закона распределения. Однако известно, что такое (нормальное, Гауссовское) распределение справедливо как предельный закон распределения при действии большого числа (реально 5-10) ограниченных по величине независимых факторов. Более часто встречающееся ограничение связано с тем, что в реальности, факторы часто оказываются зависимыми.

При использовании численного моделирования никаких исходных предположений о характере исследуемого распределения не делается.

В предлагаемых лекциях будут рассмотрены несколько модельных задач с использованием «виртуальных» экспериментальных данных и к этим данным применено численное моделирование.

Во всех предлагаемых случаях будет поставлен вопрос о достоверности наблюдаемого эффекта и решаться этот вопрос будет с использованием численных методов.

Перейдем к модельным примерам.

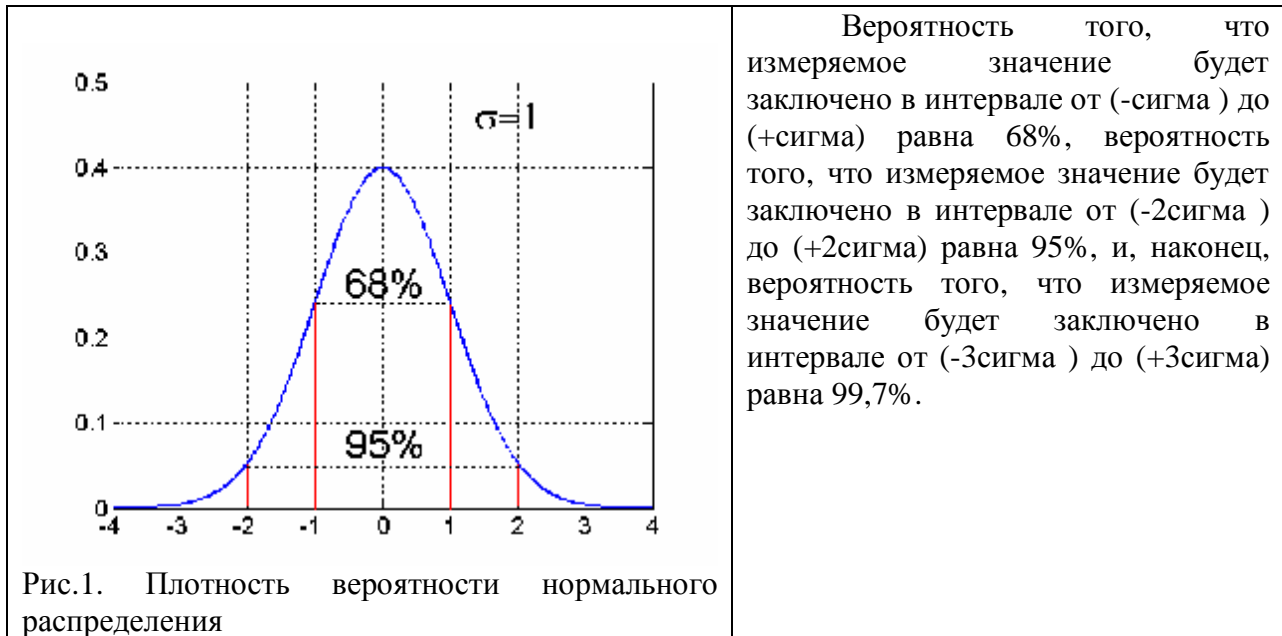
1. При наличии рядов данных ионосферных измерений для большого числа землетрясений (например, нескольких десятков) часто бывает удобно применить **метод наложения эпох**. Этот метод особенно эффективен, когда амплитуда сейсмогенной вариации меньше амплитуды шума.

В таких случаях выявить эффект визуально практически невозможно и благое пожелание ученых докомпьютерной эпохи «приведите пример аномалии» может оказаться является плохо выполнимым.

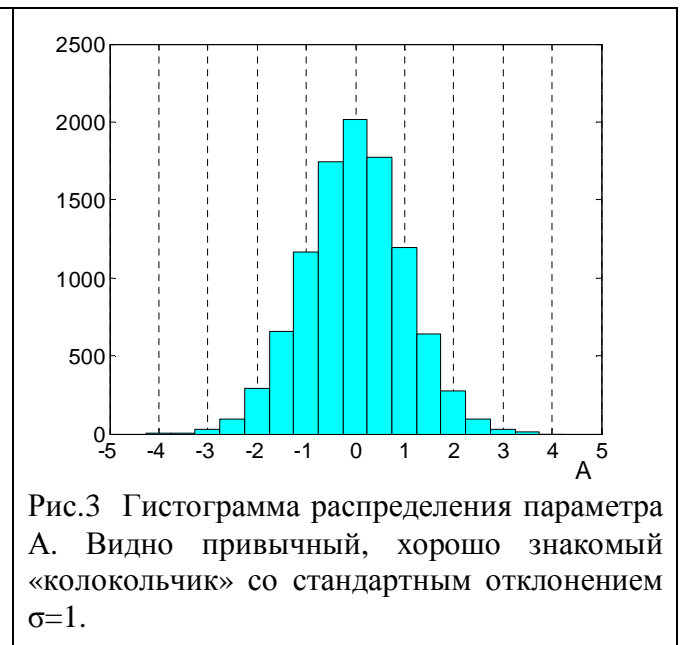
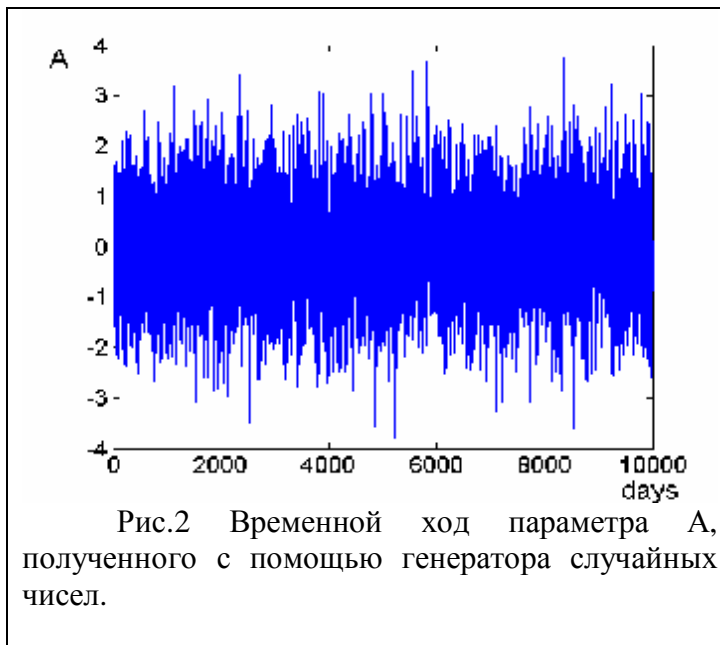
Приведем пример. Смоделируем временной ход гипотетического параметра A , подчиняющегося законам нормального распределения с $X=0$ и $\sigma=1$.

Пусть задан ход некоторого параметра A на отрезке времени в 10 тыс отсчетов (например, дней). Ход моделируется с помощью генератора случайных чисел – используется стандартная Матлабовская программа.

Напомним, что плотность вероятности нормального или гауссовского распределения выражается формулой $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-X)^2}{2\sigma^2}\right)$, вид этой функции представлен на Fig_0 для среднего значения $X=0$ и стандартного отклонения $\sigma=1$.



На следующих рисунках – временной ход и гистограмма распределения параметра А.



Будем везде рисовать ненормированные гистограммы, по оси ОУ – число случаев наблюдения. Это дает дополнительную информацию о числе случаев, и помогает оценивать достоверность.

Предположим, что за 10 000 дней произошли 300 «событий» – назовем их «реальными землетрясениями». Дни «реальных землетрясений» тоже определяем с помощью генератора случайных чисел.

При работе с временными окрестностями землетрясений часто используют обозначения (0) день – сутки, в которые произошло землетрясение, (-1) день – сутки перед ним, (+1) день – сутки после него и т.д.

Моделируем воздействие процесса подготовки землетрясения на ход исследуемого параметра. Пусть для землетрясений с номерами 1,4,7,10, ... происходит увеличение исследуемого параметра A на (-1) день на 0.5σ : для землетрясений 2, 5, 8, 11 ... – увеличение параметра A на (-2) день на 0.5σ , для землетрясений 3,6,9,12, ... увеличение A на (-3) день на 0.5σ .

При такой небольшой «сейсмодобавке» на фоне случайного процесса исследователь этого не увидит!. Рассмотрим временные окрестности первых 6 землетрясений.

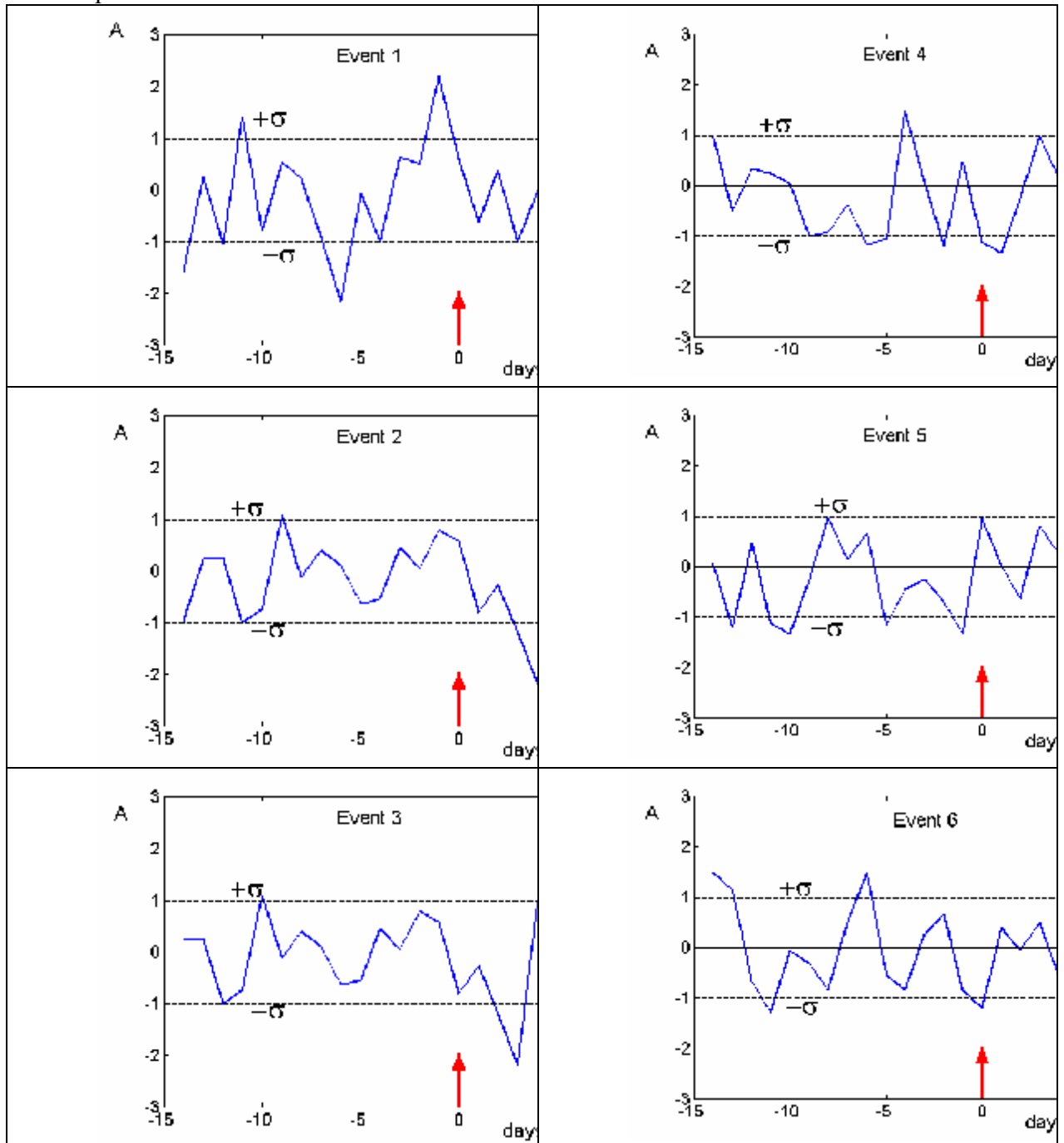


Рис.4 Временной ход параметра A в окрестности первых 6 землетрясений. Воздействие процессов подготовки «включено».

Для событий 1 и 4 смоделировано увеличение на (-1) день.

Для 1 события можно вроде бы видно увеличение на (-1) день, а для четвертого – нет.

Для событий 2 и 5 – увеличение на (-2) день – не наблюдается вовсе.

Для событий 3 и 6 увеличение на (-3) день также не наблюдается.

А для события 4 наблюдается всплеск на (-4) день, а мы его не моделировали! Всплеск получился случайно!

Тут было бы уместно сказать несколько слов о любви геофизиков к 2σ .

Вообще говоря, к 2σ , а тем более к величинам поменьше, нужно относиться очень осторожно. В статьях порой можно прочесть, что при исследовании какого-нибудь одного!!! сильного землетрясения обнаружили, что исследуемый параметр на (-3) день превысил, например, среднее на 1.7σ . При этом на рисунке показывают интервал от (-5) до (+1) дня, то есть показана неделя. Но ничего не говорится о законе распределения параметра – насколько он близок к нормальному.

Предположим даже, что он нормальный. Рассчитаем теоретическую вероятность того, что в один из рассмотренных дней параметр выйдет за интервал $\pm 1.7\sigma$.

Вероятность того, что в один день случайно окажется такое значение, равна 0.09. Соответственно, вероятность того, что значение не выйдет за границу $\pm 1.7\sigma$, равна 0.91.

Вероятность того, что ни в один из 7 дней (причем значения в соседние дни считаются независимыми, а для ионосферных параметров это не совсем так!) составит 0.91^7 , что составляет 0.52. Значит, вероятность того, что хотя бы одно значение за неделю выйдет за интервал $\pm 1.7\sigma$, составляет $1-0.52=0.49$.

Практически, половина. Причем исследователь заранее не знает, будет ли у него уменьшение или увеличение исследуемого параметра перед землетрясением, в одних работах по исследованию предвестников землетрясений провозглашается увеличение, в других уменьшение исследуемого параметра, причем одного и того же. И то и другое можно интерпретировать, и знак эффекта может зависеть от свойств землетрясения.

Самое интересное, что часто исследователь бывает прав, научная интуиция позволила правильно выявить возможный эффект, вызванный процессами подготовки землетрясений. И при статистическом исследовании подтверждается наличие эффектов именно такого знака и с такими характерными временами.

В частности, именно благодаря научной интуиции Бейкера и Девиса, заметивших отклик на сильнейшее Аляскинское землетрясение, сейчас существует целое направление – сейсмоионосферные исследования, это направление интенсивно развивается, по данной тематике опубликованы несколько сотен статей и несколько монографий.

Это с одной стороны. А с другой стороны, человеку свойственно искать закономерности даже там, где их нет, где процесс случайный. Биологами специально проводились эксперименты: испытуемым предлагали с одной стороны, зашумленные закономерности, а с другой – случайные последовательности. И в случайных последовательностях испытуемые закономерности также «успешно» находили. Такие исследования проводили в Москве, в Институте высшей нервной деятельности, и на эту тему еще в 1970 году был опубликован сборник статей, удостоенный Государственной премии [2]. Исследования в этом направлении продолжаются и сейчас.

А теперь рассмотрим поведение нашего параметра $A(i)$ относительно любимой геофизиками границы в 2 стандартных отклонения.

До того, как мы «включили» землетрясения и слегка испортили нормальное распределение, у нас 2.5% всех значений превышали 2σ , т.е. примерно 250 значений из 10 000 (для данного набора случайных чисел получилось 238 значений)

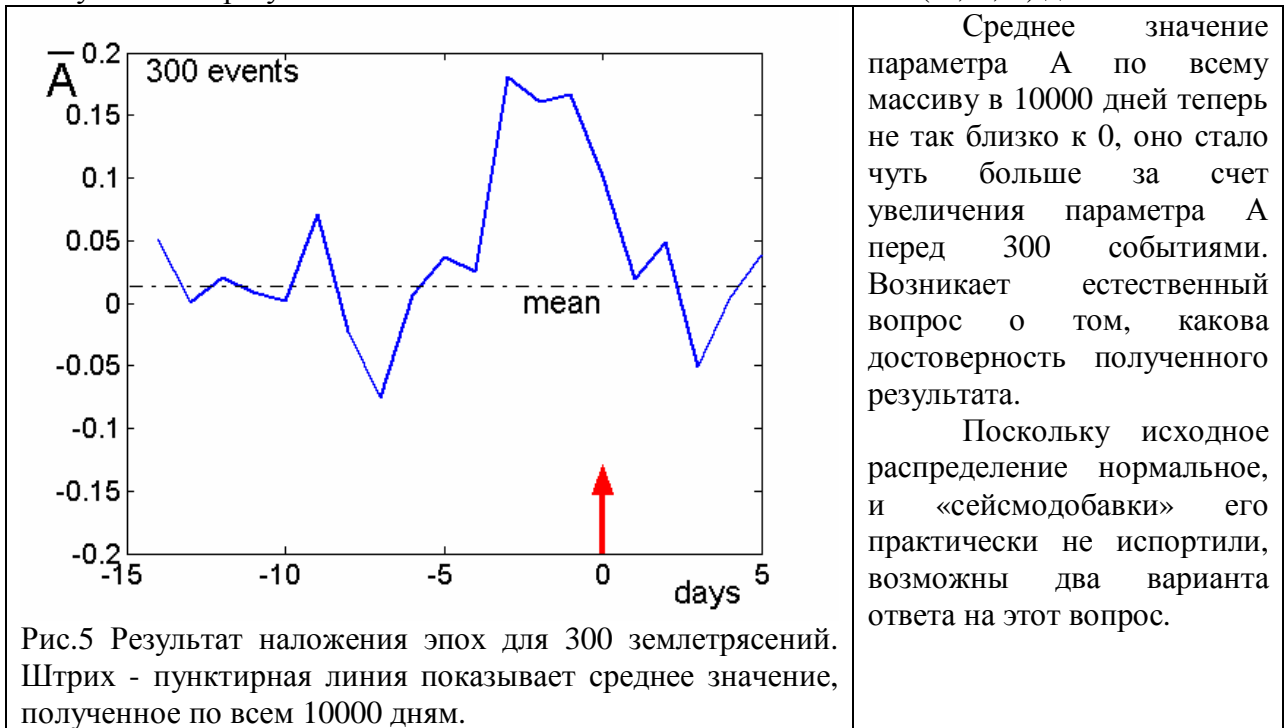
А теперь выясним: сколько значений $A(i)$ из 10000 превышает 2σ после «включения» действия процессов подготовки землетрясений?

Подсчитав, получим, что таких значений стало 252. Т.е. при «включении» воздействия процессов подготовки 300 землетрясений добавились **всего** 14 дней, когда амплитуда параметра А превысила 2 стандартных отклонения!

Напомним, что по нашему предположению процессы подготовки оказывают действие перед **всеми** землетрясениями, только в разные дни –от (-1) до(-3).

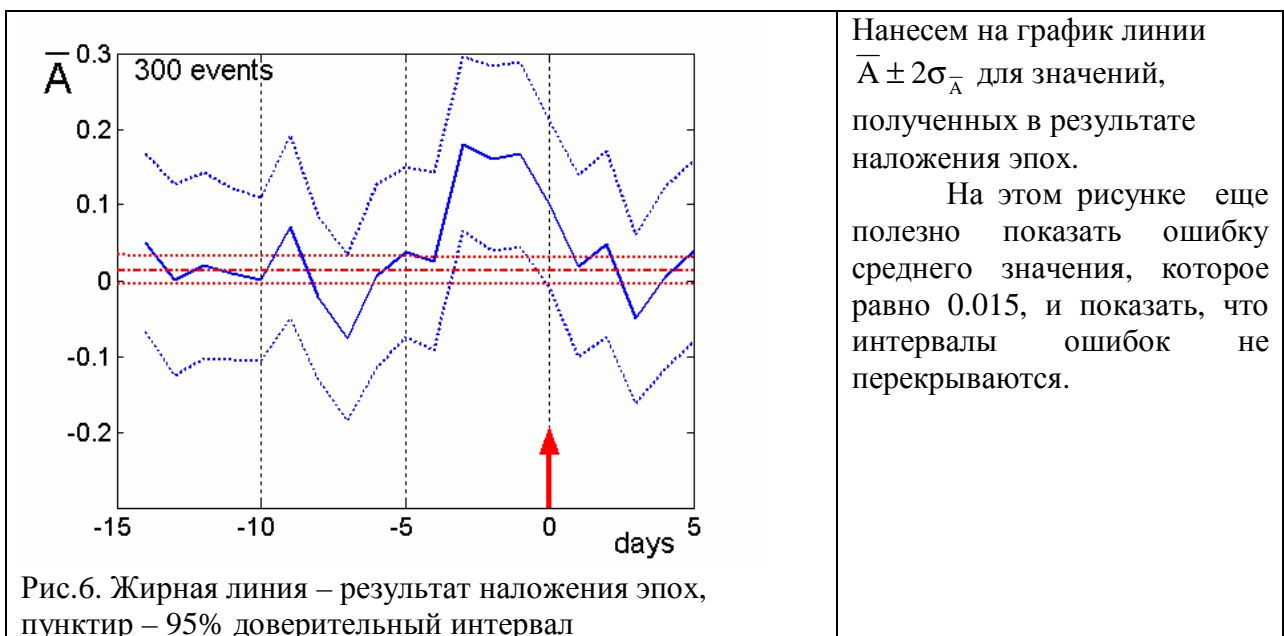
А теперь наложим эпохи – т.е. вычислим среднее значение параметра А за каждый день –за (-14), (-13)... и т.д. до(+5)) по всем 300 землетрясениям (см. рис...).

Теперь видно, что на (-3,-2,-1) дни наблюдается увеличение параметра А. Полученные в результате наложения эпох значения \bar{A} больше в (-3,-2,-1) дни.



В первом варианте оценки предполагается параметрический подход, можно показать ошибки, используя законы нормального распределения параметра А. Для оценки

ошибки используем стандартное отклонение среднего $\sigma_{\bar{A}} = \frac{\sigma}{\sqrt{N}}$



Для этого также используют стандартное отклонение среднего для 10000 чисел, оно равно стандартному отклонению, вычисленному по всей совокупности измерений σ/\sqrt{N} .

В данном случае $\sigma=1$, $N=10000$; получаем ошибку среднего значения на 95% доверительном уровне составляет $2*0.01=0.02$.

Во все дни, кроме (-1,-2,-3) интервалы ошибок перекрываются, но в «сейсмодни» интервалы ошибок не перекрываются. Значит, на 95% уровне значимости, или можно сказать, с вероятностью не меньшей 0.95 наблюдается повышение А в (-1,-2,-3) дни перед землетрясением.

Во втором варианте оценки предполагает численный подход.

А теперь попробуем получить тот же результат, не делая никаких предположений о распределении параметра А. Применим численные методы.

Метод Монте-Карло (методы Монте-Карло) — общее название группы численных методов, основанных на получении большого числа реализаций случайного процесса, который формируется таким образом, чтобы его вероятностные характеристики совпадали с аналогичными величинами решаемой задачи. Этот метод широко используется для решения задач в областях физики, математики, экономики, оптимизации, теории управления и др.

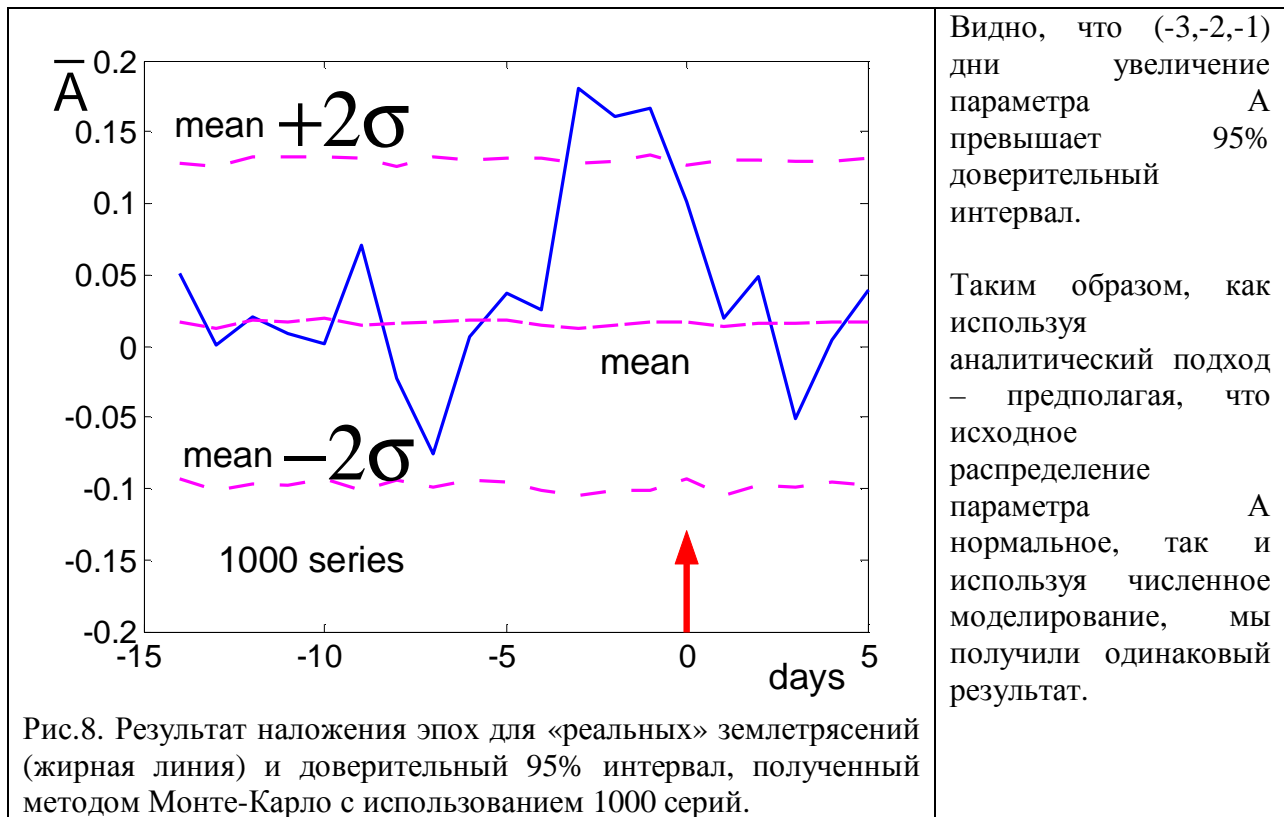
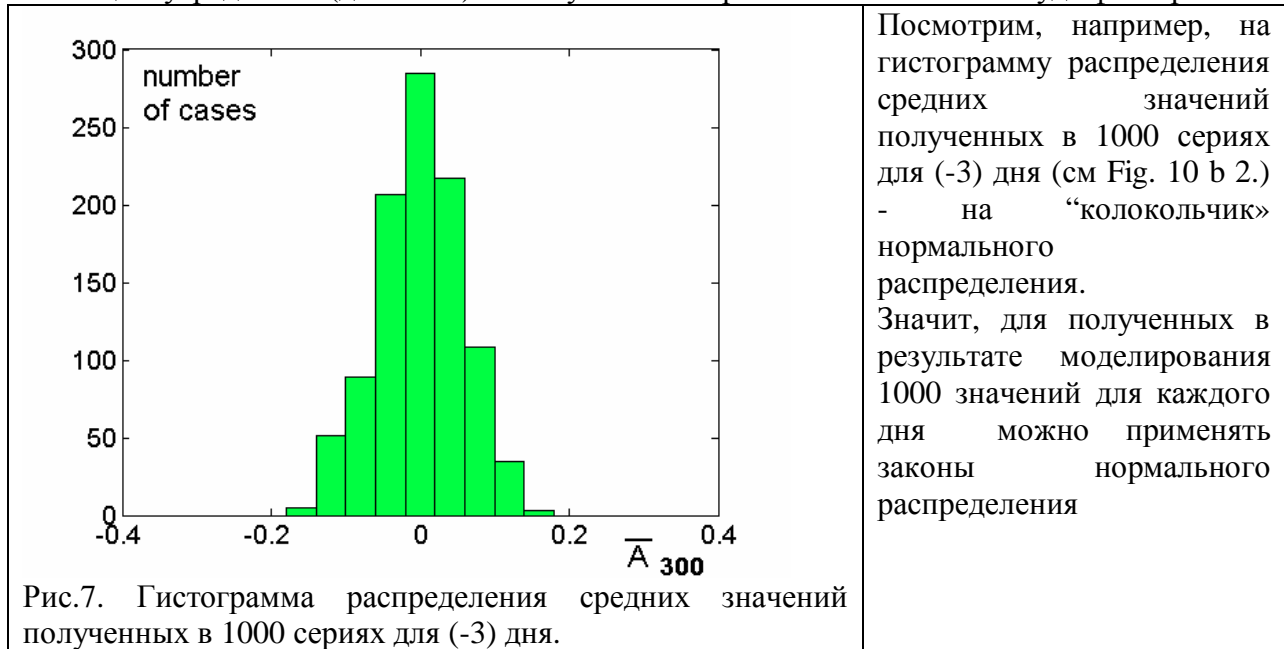
Идея была развита Станиславом Уламом, который на отдыхе задался вопросом, какова вероятность того, что пасьянс «сложится». Ему в голову пришла идея, что вместо того, чтобы использовать обычные для подобных задач соображения комбинаторики, можно просто поставить «эксперимент» большое число раз и, таким образом, подсчитав число удачных исходов, оценить их вероятность. Он же предложил использовать компьютеры для расчетов методом Монте-Карло.

Появление первых электронных компьютеров, которые могли с большой скоростью генерировать псевдослучайные числа, резко расширило круг задач, для решения которых стохастический подход оказался более эффективным, чем другие математические методы. После этого произошел прорыв и метод Монте-Карло стал применяться во многих задачах, однако его использование не всегда было оправдано из-за большого количества вычислений, необходимых для получения ответа с заданной точностью.

Годом рождения метода Монте-Карло формально считается 1949 год, когда в свет выходит статья Метрополиса и Улама «Метод Монте-Карло». Название метода происходит от названия города в княжестве Монако, широко известного своими многочисленными казино, поскольку именно рулетка является одним из самых широко известных генераторов случайных чисел. Станислав Улам пишет в своей автобиографии «Приключения математика», что название было предложено Метрополисом в честь его дяди, который был азартным игроком.

Пусть задана такая же последовательность, как в рассмотренном выше примере, тот же набор «реальных» землетрясений и смоделированный на (-1,-2,-3) дни «предвестниковый эффект». Для оценки достоверности полученного при наложении эпох эффекта воспользуемся методом моделирования фонового распределения – методом Монте-Карло. Создадим 1000 серий по 300 виртуальных «землетрясений» в каждой. Для каждого такого «виртуального» землетрясения выделим интервал длительностью 20 дней, от (-14) до (+1) дня для этого виртуального землетрясения, точно так же, как для реальных землетрясений, и наложим эпохи для виртуальных землетрясений.. Получим 1000 результатов наложения эпох – 1000 кривых, а для каждого дня от (-14) до (+1) получим 1000 значений.

Исходное распределение параметра A может быть каким угодно, а эти **полученные в результате моделирования и наложения эпох 1000 значений для каждого дня будут распределены в соответствии с нормальным законом**, потому что получены с помощью усреднения (действия) 300 случайных ограниченных по амплитуде факторов.



Далее зададимся вопросом: сколько серий нужно создавать?

Ниже на рис. Приведен пример, когда при моделировании использовано 100 серий,

Отметим, что кривые средних и 2 стандартных отклонений тем более гладкие, чем больше серий. Отсюда можно сделать практический вывод: нужно столько серий, чтобы смоделированные кривые стали гладкими и заметно не менялись, если число серий увеличить, например, вдвое.

Обратим внимание, что картинку, близкую к Рис.8 можно получить, если подсчитать среднее по всему массиву данных – прямая линия на рис.– и провести ± 2 стандартных отклонений среднего, которое можно оценить как среднеквадратичное отклонение, деленное на \sqrt{N} , где $N=300$ (см. рис.). Но последняя процедура правильная, только если исходное распределение – нормальное.

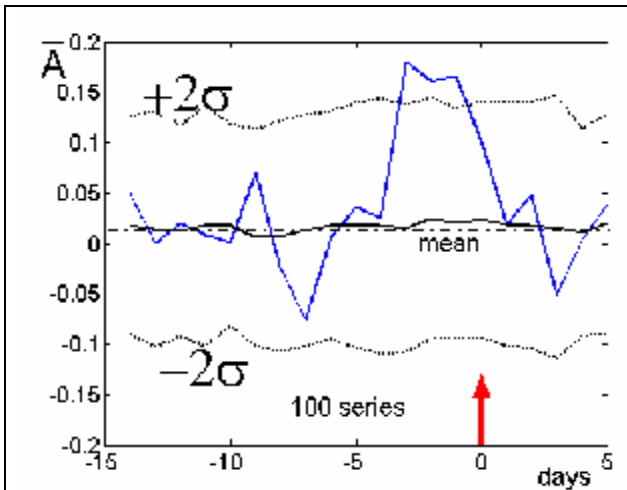


Рис.9. Результат моделирования, полученный с использованием 100 серий.

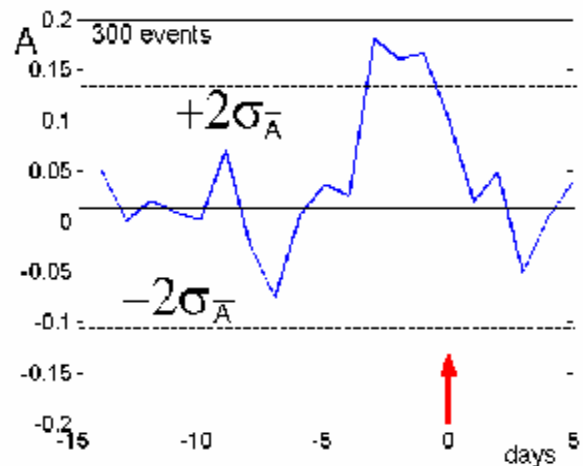


Рис. 10. Аналитическая оценка погрешности

2. Усложним модельную задачу.

Создадим последовательность элементов, распределение которых нельзя назвать нормальным. Например, рассмотрим последовательность чисел Вольфа – числа солнечных пятен.

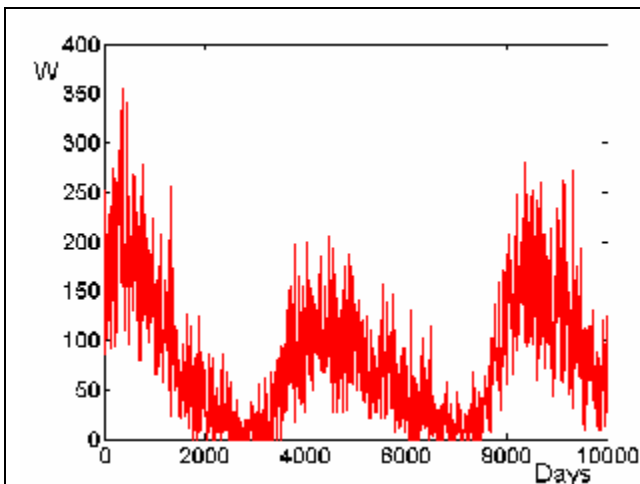


Рис. .11 На рис представлен ход W –числа Вольфа за 10 000 дней начиная с января 1957 г. – примерно за 30 лет.

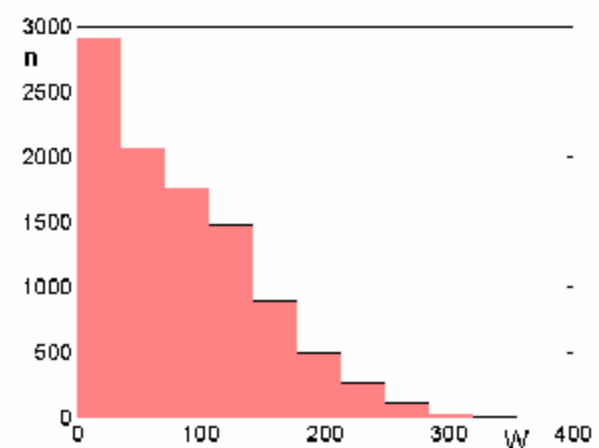


Рис. 12. Гистограмма распределения W.

Для такого распределения $\text{mean}=83$, $\text{median}=72$, $\text{std}=63$.

Теперь с помощью генератора случайных чисел выберем из этой совокупности чисел Вольфа 10 000 значений.. Конечно, получится, что некоторые значения взяты по несколько раз, а некоторые – ни разу, но в данном случае это не важно, наоборот, так и

надо, мы моделируем числовой ряд. Назовем эту последовательность смоделированных чисел последовательностью A .

В этом ряду есть нулевые значения, для удобства заменим их на единичные.

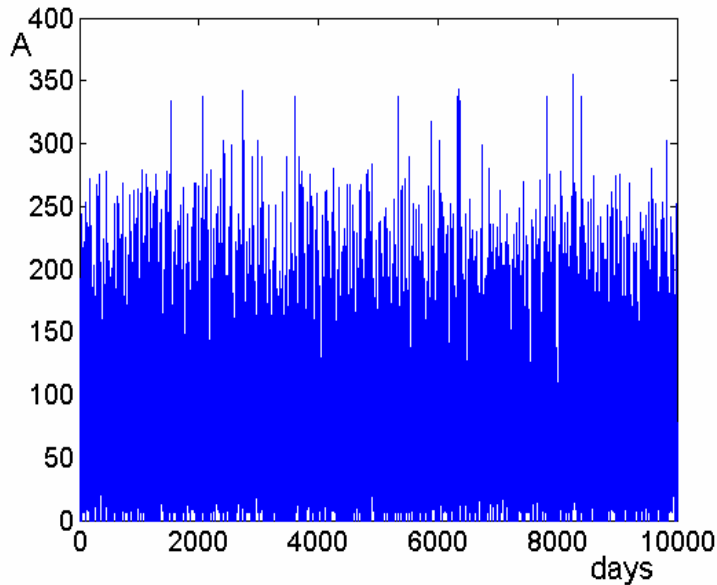


Рис. 13. Смоделированный числовой ряд A

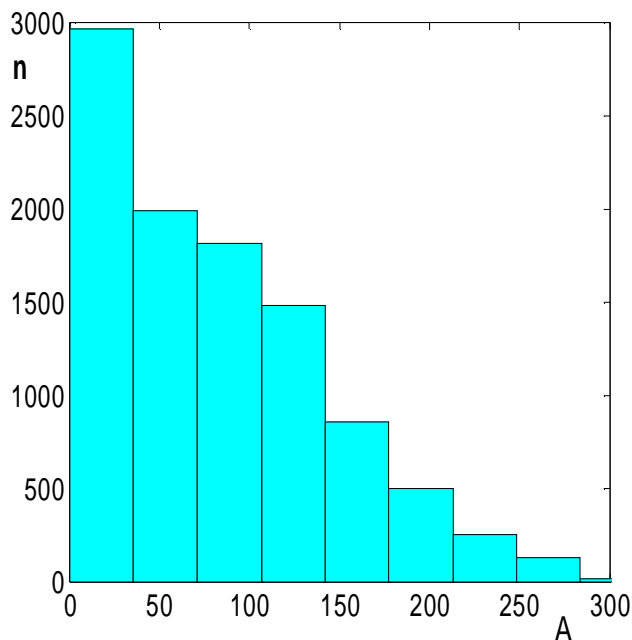
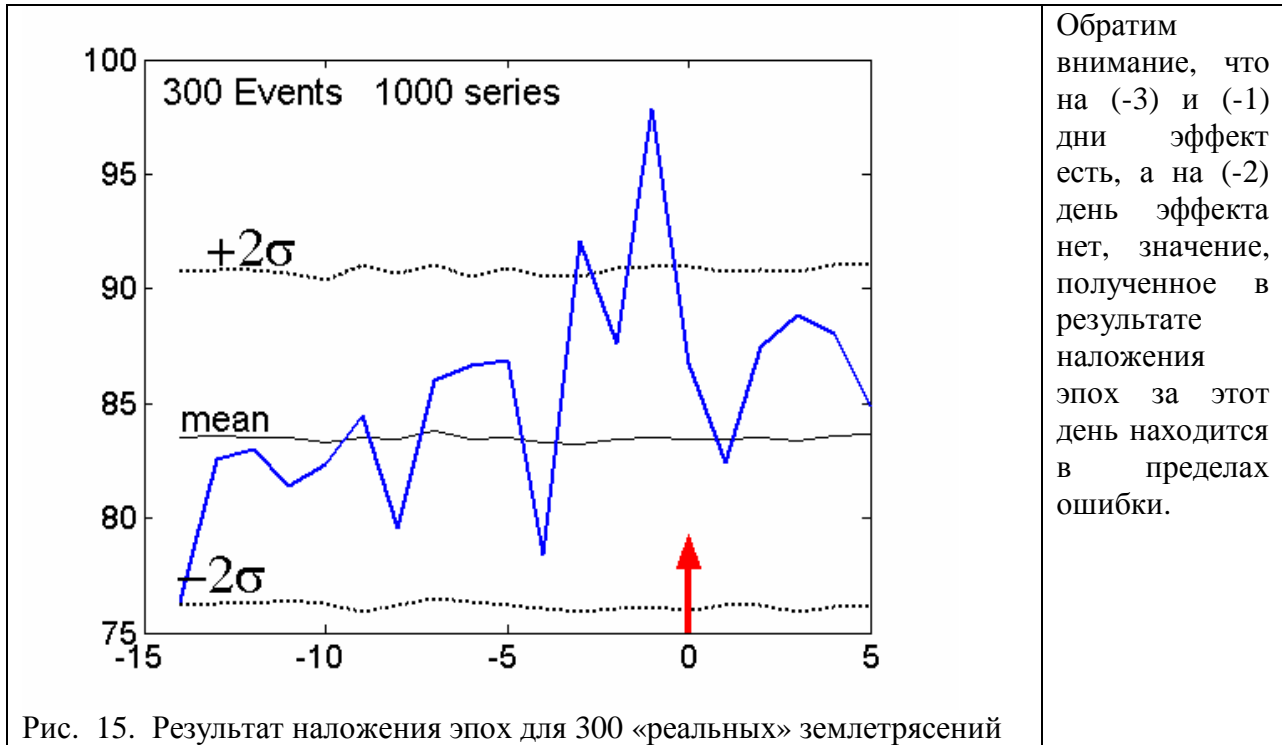


Рис..14 Гистограмма распределения этого числового ряда A

Среднее значение элементов этого ряда A $\text{mean}=83$, медиана $\text{median}=71$ стандартное отклонение $\text{std}=64$. Числа, естественно, близки к аналогичным для исходного ряда чисел Вольфа W .

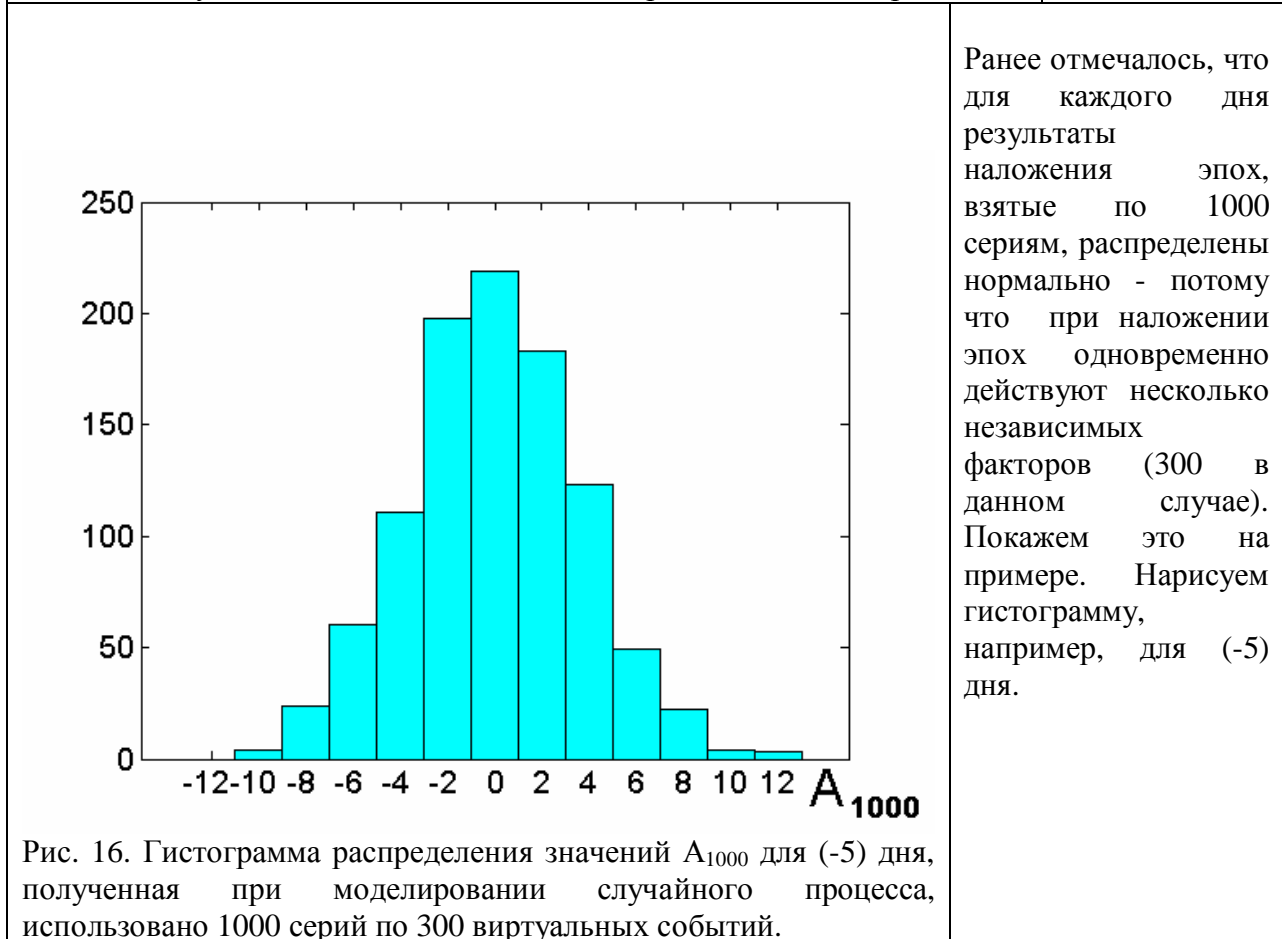
Выберем с помощью генератора случайных чисел 300 событий, которые назовем «реальными землетрясениями» и повторим процедуру – задаем увеличение A на 0.5 стандартного отклонения для землетрясений с номерами (1, 4, 7, ...), на (-1) день, для землетрясений с номерами 2, 5, 8, ... величину A увеличиваем на (-2) день, для землетрясений с номерами 3, 6, 9, ... величину A увеличиваем на (-3) день.

Проведем наложение эпох и моделирование случайного процесса, используя 1000 серий по 300 виртуальных событий.



Обратим внимание, что на (-3) и (-1) дни эффект есть, а на (-2) день эффекта нет, значение, полученное в результате наложения эпох за этот день находится в пределах ошибки.

Рис. 15. Результат наложения эпох для 300 «реальных» землетрясений



Ранее отмечалось, что для каждого дня результаты наложения эпох, взятые по 1000 сериям, распределены нормально - потому что при наложении эпох одновременно действуют несколько независимых факторов (300 в данном случае). Покажем это на примере. Нарисуем гистограмму, например, для (-5) дня.

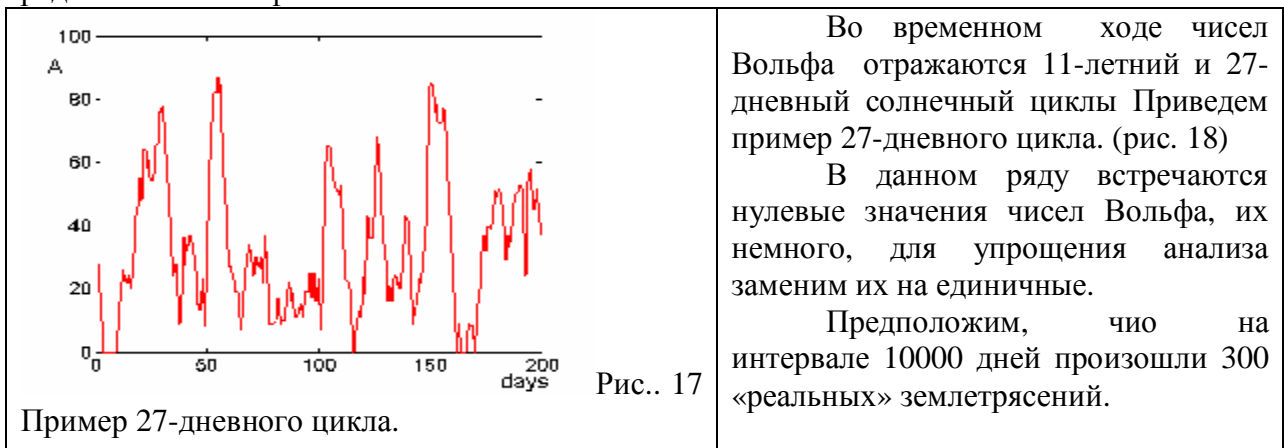
Рис. 16. Гистограмма распределения значений A_{1000} для (-5) дня, полученная при моделировании случайного процесса, использовано 1000 серий по 300 виртуальных событий.

Чтобы упростить рисунок, сначала вычислим среднее за этот день по 1000 сериям, и затем вычтем это среднее из каждого значения из 1000, чтобы сместить гистограмму и

центр поместить в начало координат для удобства восприятия. Получаем нормальный «колокольчик».

При анализе «сейсмоэффекта» на исходной выборке с нормальным распределением было достаточно «сейсмодобавки» величиной в $\frac{1}{2}$ стандартного отклонения, чтобы эффект с достоверностью выше 95% выделялся при наличии данных всего лишь за 100 землетрясений. При исходном «ненормальном» распределении амплитуда добавки, вызванной «подготовкой землетрясений» мала, чтобы выявить эффект, имея в распоряжении данные только для 100 событий.

3. Теперь рассмотрим модельную задачу, более приближенную к реальной ситуации. Воспользуемся для моделирования хода исследуемого параметра A настоящим рядом чисел Вольфа.



Предположим, что для половины землетрясений на (-1) день происходит увеличение величины A на 10%, а для половины ничего не меняется.

Теперь проанализируем четыре различных способа обработки и выявления «сейсмоэффекта».

1 способ. Если «в лоб» просто наложить эпохи, эффект статистически надежно выявить не удастся, поскольку разные годы 11-летнего солнечного цикла очень различаются амплитуды параметра A , а землетрясений немного, всего 300.



2 способ. Теперь попробуем наложить эпохи, используя тот же набор землетрясений, предварительно пронормировав накладываемые ряды на среднее значение по этому же интервалу. Далее оценим достоверность методом Монте-Карло.

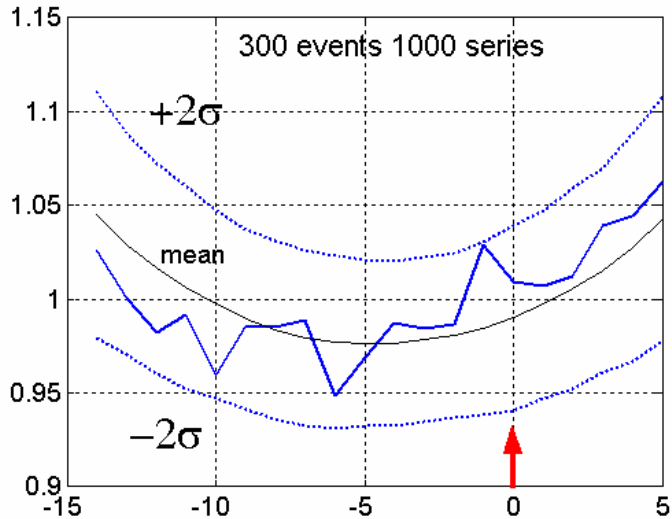


Рис.19 Наложение эпох для ряда с сильной периодичностью; значения нормированы на среднее за рассматриваемый интервал.

Из рис.19 видно, что получается некое увеличение на (-1) день получается. Но, кроме того, получаем тренд среднего значения и доверительного 95% интервала, вызванный 27-дневным солнечным циклом. В целом, и достоверность недостаточная, и сам метод плохой из-за периодичности исходного ряда..

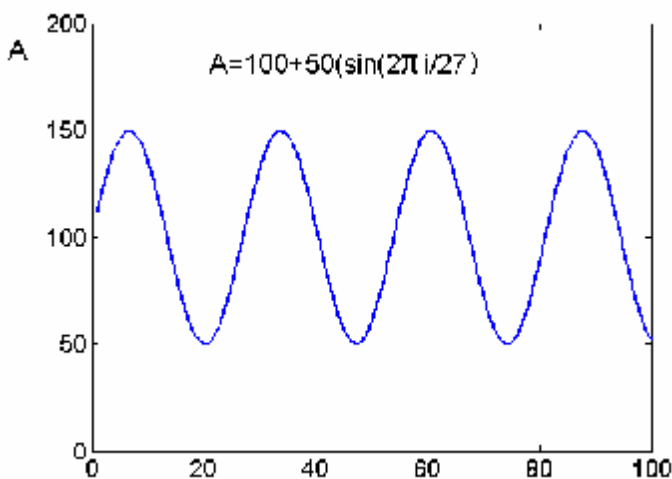


Рис.20 Модельная функция $A(i)=100+50*\sin(2\pi i/27)$

Кстати, с помощью моделирования можно убедиться, что именно 27-дневная периодичность создает такой тренд. Сделать это можно, например, рассмотрев периодическую модельную функцию $A(i)=100+50*\sin(2\pi i/27)$ – составленную из константы и периодической добавки.

Значения в каждом накладываемом интервале нормируются на среднее по этому интервалу, и когда мы берем интервалы в окрестности минимума, мы получаем большую кривизну. Из-за разницы в крутизне при наложении эпох центральная часть кривой опускается.

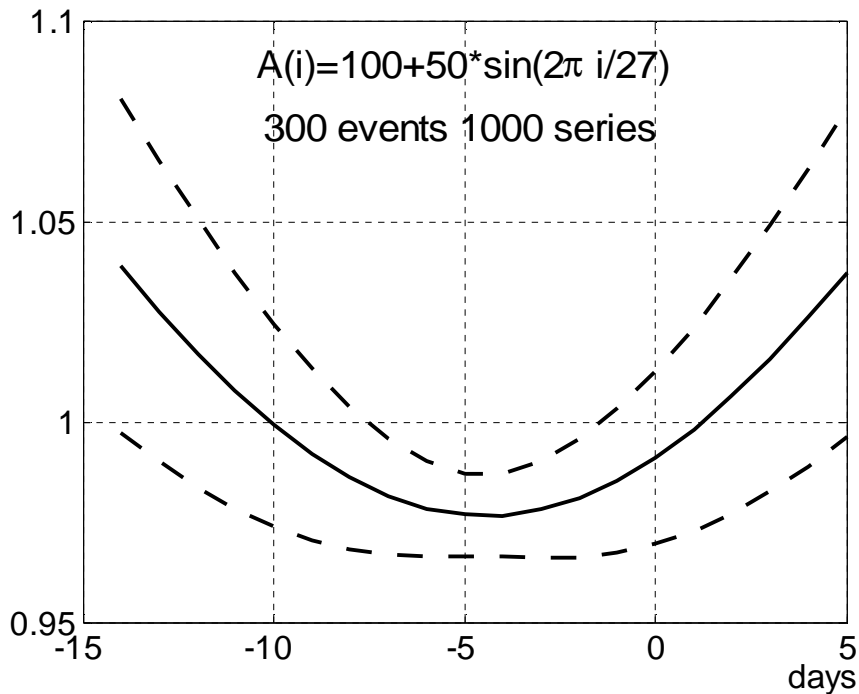


Рис. 21 Результат наложения эпох для модельной функции, используется, как обычно, 300 интервалов по 20 дней в 1000 сериях. Получается тренд такого же типа.

Часто исследователь рад-радехонек, что у него есть данные за интервалы в окрестности землетрясений, а что находится за этими интервалами – это фон, и он как бы несущественен. А в результате можно совершить методическую ошибку и получить «эффект уменьшения исследуемого параметра» в середине интервала. Без моделирования фонового распределения с помощью случайного процесса такую методическую ошибку не выявить.

Опять вернемся к рис. 19. На рис. 19 кривая наложения эпох для землетрясений почти касается пунктирной кривой, показывающей 95% доверительный интервал. Результат недостоверный на 95% уровне значимости и метод нельзя признать удачным.

Теперь пронормируем последовательность $A(i)=100+50\sin(2\pi i/27)$ на *среднее по скользящему интервалу 20 дней*.

Видно, что крутизна кривой больше в окрестности минимумов, чем в окрестности максимумов. Для этого на рис. Рис.22 для фрагмента кривой выбран более крупный масштаб по оси ОХ.

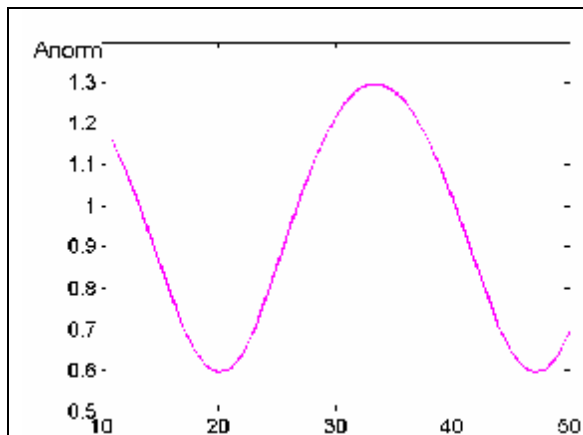


Рис.22 последовательность $A(i)=100+50\sin(2\pi i/27)$ на среднее по скользящему интервалу 20 дней.

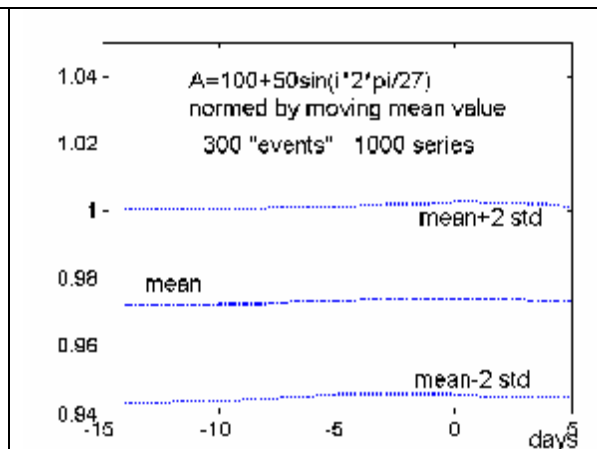


Рис.23. Результат наложения эпох для 300 интервалов по 20 дней для нормированной функции Anorm..

Но если сейчас промоделировать наложение эпох, никакой вогнутости не получится! Получится прямая! По-видимому, происходит это потому, что на краях при наложении эпох «участвуют» значения, непосредственно не входящие в исследуемый интервал.

Обратим внимание на то, что среднее значение – отнюдь не единица.

3 способ. Теперь попробуем наложить эпохи, пронормировав наш ряд A , соответствующий ряду реальных солнечных пятен на скользящее среднее за $(-10, +9)$ дни. Набор «реальных землетрясений» все те же.

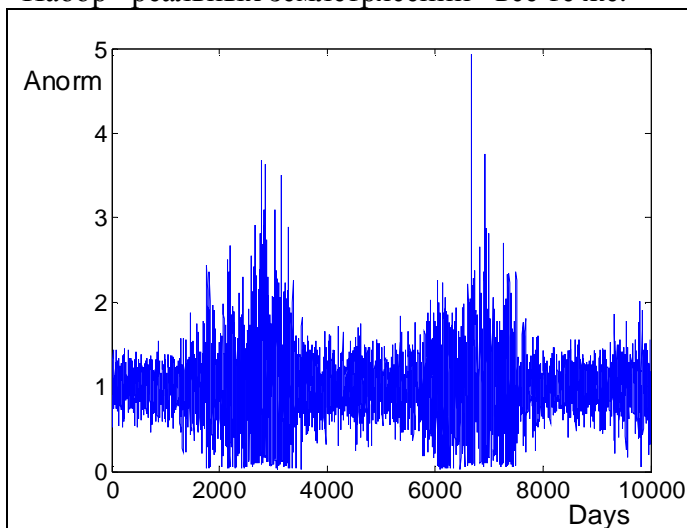


Рис.24 Anorm-нормированная на скользящее среднее по интервалу 20 дней последовательность A

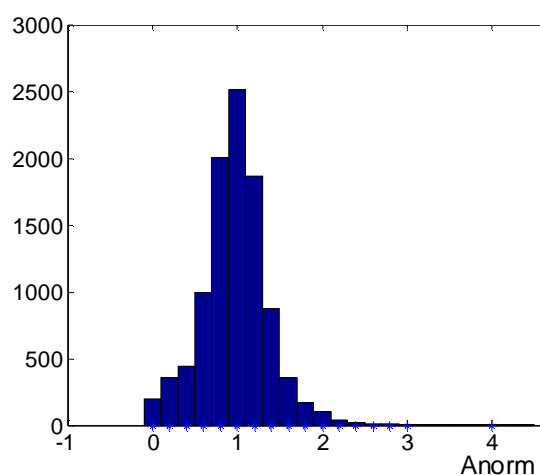


Рис. 25 Гистограмма ряда Anorm

Распределение Anorm близко к нормальному – потому что элементы получены в результате действия (в данном случае деления на среднее) нескольких ограниченных по величине факторов. Конечно, эти факторы не являются независимыми – значения солнечных пятен в соседние дни достаточно зависимы, но и, соответственно, распределение не совсем нормальное.

Наложим эпохи для последовательности Anorm..

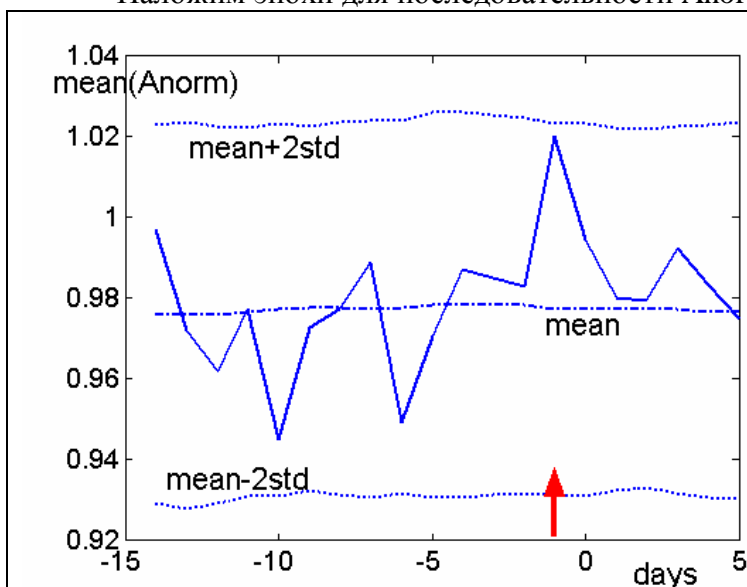


Рис. 26. Результат наложения эпох для последовательности Anorm. и моделирования фонового распределения

На рис. 26 кривая наложения эпох для землетрясений почти касается пунктирной кривой, показывающей $\text{mean} + 2 \text{std}$ (среднее + 2 стандартных отклонения), но в целом результат недостоверный на 95% уровне значимости.

4 способ. Теперь проведем исключение периодичности и нормирование данных на среднеквадратичное отклонение. Далее применим метода наложения эпох и определим достоверность выявленного эффекта с использованием метода Монте –Карло.

Для моделирования используем тот же исходный ряд данных, и те же «реальные землетрясения». Так же как в предыдущих трех случаях предполагаем, что для половины землетрясений на (-1) день происходит увеличение величины A на 10%, $A_{seism}(i)=A(i)*1.1$, а для половины ничего не происходит.

Попробуем избавиться от 11-летнего и 27-дневного циклов.

Часто для исключения правильной периодичности используют, например, быстрое преобразование Фурье, т.е. раскладывают в ряд Фурье, далее проводят обратное преобразование Фурье для некоторого диапазона частот, восстанавливая часть функции, вычитают восстановленную функцию из исходной. Мы так поступать не будем, а будем из текущего значения вычитать скользящее среднее, учитывая перечисленные ниже соображения.:

1.Спектр анализируемого ряда значений «розовый», и 27 дневная гармоника сравнительно слабая.

Ниже приведены примеры: розовый спектр последовательности чисел Вольфа (рис. 27) и для сравнения, спектр белого шума (рис. 28).

2. Периодичность анализируемого ряда неправильная. Солнце делает оборот за 27 земных дней, а экстремумы числа солнечных пятен приходятся на разные дни 27-дневного цикла.

3. Вычесть скользящее среднее технически проще.

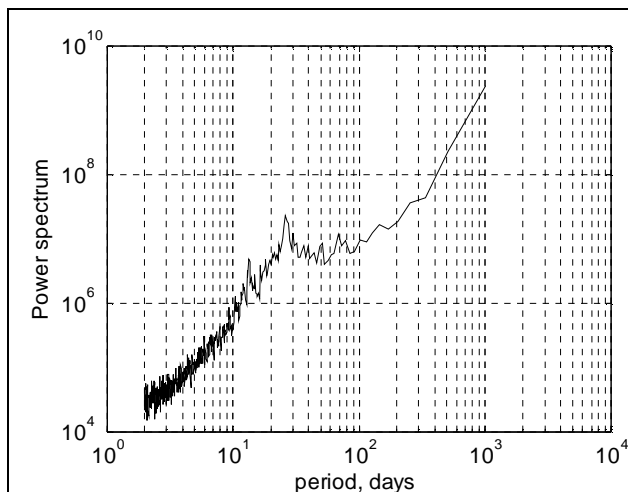


Рис.27 Розовый спектр индекса Вольфа

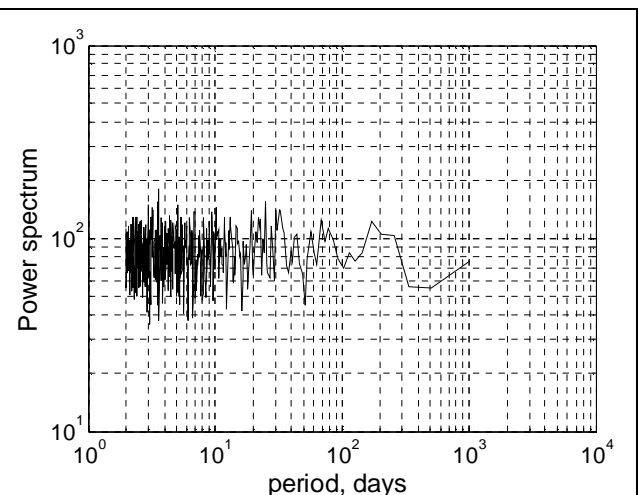


Рис 28 Белый спектр последовательности случайных чисел

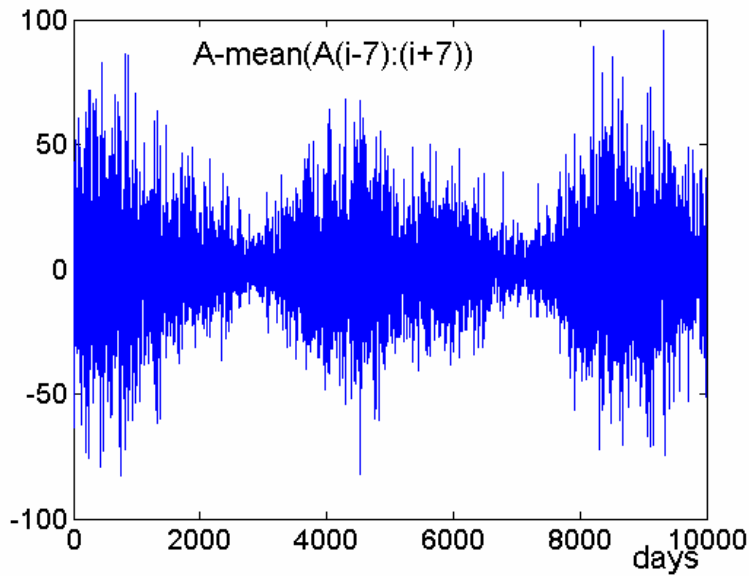


Рис. 29. Последовательность значений параметра A - $A_{\text{усредненное по 15 дням}}$

После вычитания скользящего среднего, получим последовательность, изображенную на рис. 23. Интервал усреднения возьмем меньше 27 дней – например, 15 дней. Это означает, что ищется эффект с характерным временем, много меньше 15 дней – от одного до 4 дней, т.е. примерно до четверти интервала усреднения. Конечно, амплитуда вариации краткосрочного эффекта уменьшится при вычитании среднего,

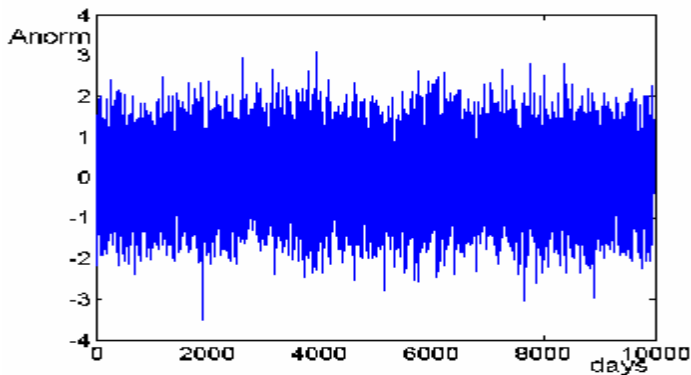


Рис.30. A_{norm} –разница текущего значения и сглаженного по скользящему интервалу параметра A , нормированная на скользящее среднеквадратичное отклонение по 15 дням.

усредненного по сравнительно короткому интервалу, но при таком усреднении хотя бы исчезнет 27-дневная зависимость. Но зависимость от 11 летнего солнечного цикла ещё остается. Одна из возможностей её исключения- следующая. Пронормируем получившийся ряд, поделив каждый элемент на скользящее среднеквадратичное отклонение по тому же интервалу. В результате зависимость от 11-летнего солнечного цикла практически исключена.

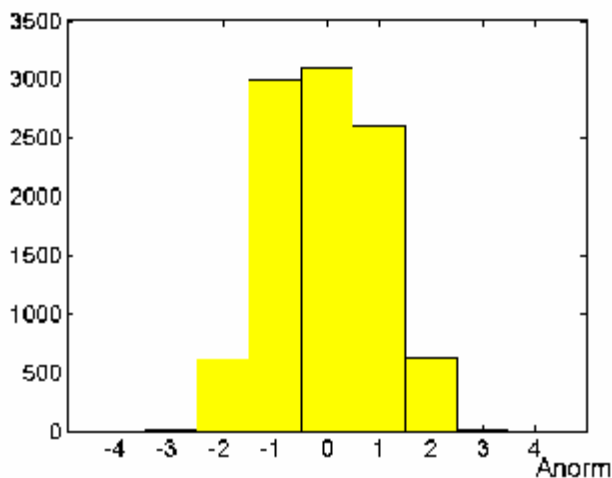


Рис. 31 Гистограмма полученного параметра A_{norm} .

Построим гистограмму распределения полученной величины A_{norm} .

Параметр имеет распределение, близкое к нормальному – поскольку значения A_{norm} получены при наличии воздействия (вычитание среднего и деление на среднеквадратичное) нескольких ограниченных по амплитуде и отчасти независимых факторов. Поэтому и распределение приближается к нормальному.

Такой метод можно применить и для исключения сезонной зависимости исследуемого параметра, если она имеет место.

Далее наложим эпохи для интервалов (-14, +5) во временной окрестности землетрясений и проведем моделирование фонового распределения методом Монте-Карло. Как и раньше, создадим 1000 серий по 300 виртуальных землетрясений в каждой, наложим эпохи временных интервалов (-14, +5) дней в окрестности виртуальных землетрясений. Как уже отмечалось, 1000 значений $\bar{A}_{\text{ном}}(i)$ для каждого дня i , полученные в 1000 сериях, распределены по нормальному закону. Для каждого дня вычисляем среднее значение и доверительный 95% интервал.



Рис.32 Результат наложения эпох для параметра $A_{\text{ном}}$. Пунктир – 95% доверительный интервал, полученный методом Монте-Карло.

Зададим вопрос: для скольких событий параметр $A_{\text{ном}}(-1)$ больше, и для скольких событий меньше медианы (или среднего)? Из чего складывается значение 0.17, -результат наложения эпох в (-1) день, которая с запасом превышает значение, равное среднему + 2 среднеквадратичных отклонения?

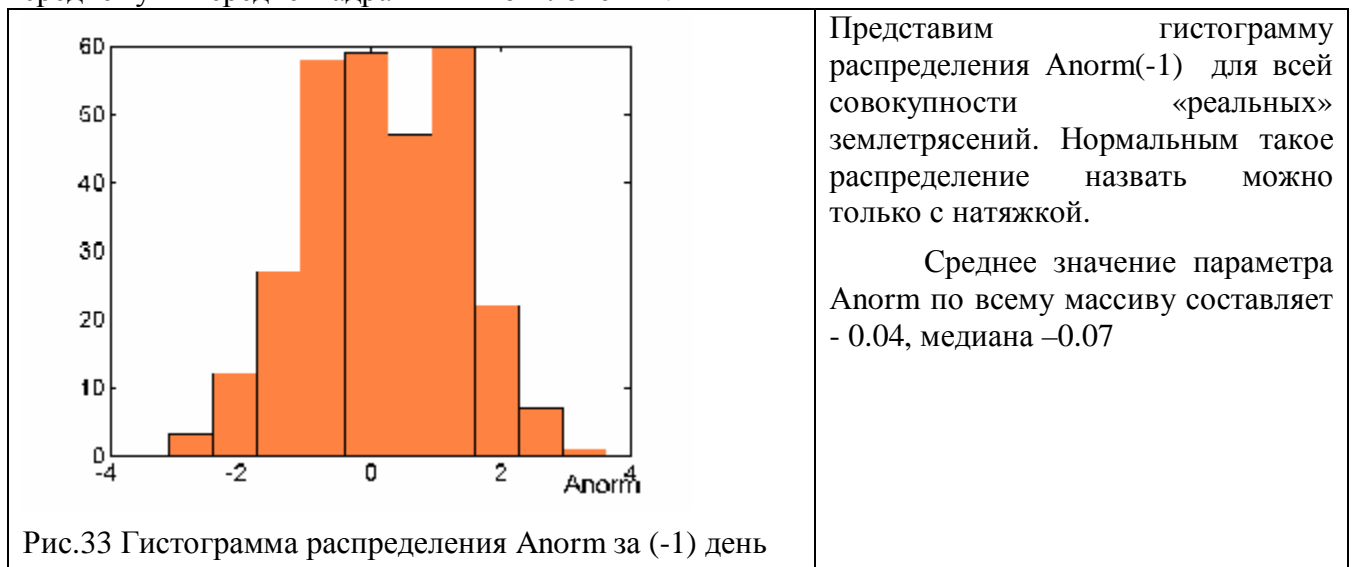


Рис.33 Гистограмма распределения $A_{\text{ном}}$ за (-1) день

Проанализировав распределение параметра A_{norm} за (-1) день для 300 «реальных землетрясений», получаем, что для 170 событий в (-1) день A_{norm} больше медианы, а для 130 A_{norm} меньше медианы. Другими словами, для 57% землетрясений наблюдается увеличение, для 43%-уменьшение величины A_{norm} . Ясно, что уменьшение и увеличение равновероятны, если нет никаких целенаправленных воздействий, (вызванных процессами подготовки землетрясений, например) если мы сравниваем величину A_{norm} с медианой, вычисленной по всей большой выборке в 10000 дней. Здесь же, однако, по выборке, состоящей из 300 значений за (-1) день разница составляет $57-43=14\%$, что соответствует 40 землетрясениям.

При моделировании мы полагали, что для каждого второго события (или для 150 землетрясений) наблюдается увеличение измеряемого параметра A на 10%. А в результате расчетов получили, что как бы только 40 землетрясений приводят к наблюдаемому эффекту.

Исследователю интересно знать, как географически распределены землетрясения, процессы подготовки которых вызывают предвестниковый эффект, какими свойствами (магнитуда, глубина, возможно тектонические особенности) они обладают. Словом, хотелось бы знать, какие землетрясения вызывают эффект, а какие нет. И далее от обработки данных и моделирования обратиться к изучению причин возникновения предвестников – к изучению возможных геофизических процессов в системе литосфера-атмосфера-ионосфера.

Представим графически массив значений $A_{norm}(-1)$ для наших «реальных» землетрясений.

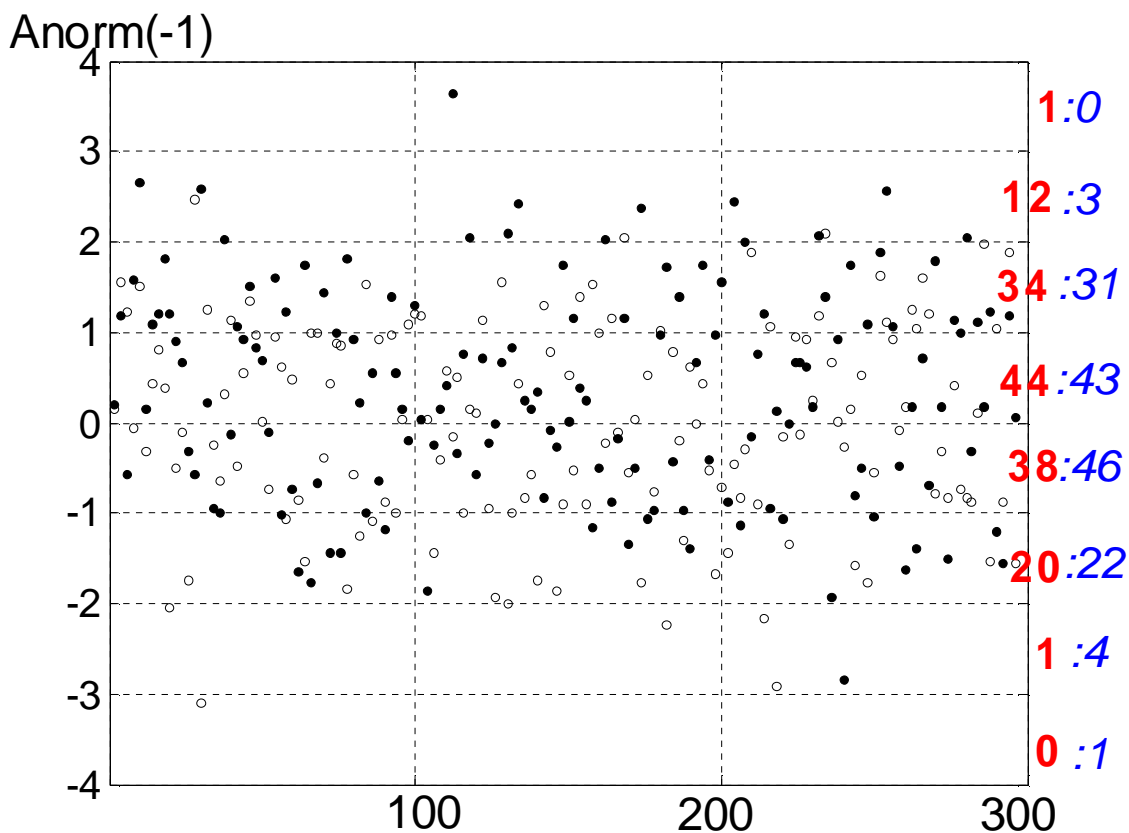


Рис.34 Темные точки показывают $A_{norm}(-1)$ для землетрясений, процессы подготовки которых «воздействуют» на ход параметра A . Пустые кружочки показывают $A_{norm}(-1)$ для землетрясений, для которых воздействия процессов подготовки не было, так как мы

его не моделировали. Прямым шрифтом показаны числа землетрясений в соответствующем диапазоне A_{norm} , для которых моделировался предвестниковый эффект. Наклонным шрифтом показаны числа землетрясений, для которых эффект не моделировался. По оси OX отложены номера событий.

В целом массив черных точек смещен вверх относительно массива синих звездочек.

Можно визуальнo оценить, какие землетрясения как бы дают предвестниковый эффект, а какие как бы не дают. Получается, что если отобрать землетрясения с большими амплитудами $A_{norm}(-1)$, то это будут скорее всего те землетрясения, которые вызвали увеличение $A_{norm}(-1)$. Значение $A_{norm}(-1) > 3$ встречается один раз; землетрясение отмечено красной звездочкой, процессы подготовки перед этим землетрясением воздействовали на исследуемый параметр А.

Если проанализировать землетрясения с меньшими амплитудами $A_{norm}(-1)$, то получаем, что при амплитудах $2 < A_{norm}(-1) < 3$ воздействие процессов подготовки имело место перед 12 событиями из 15.

А при меньших амплитудах $A_{norm}(-1)$, при $0 < A_{norm}(-1) < 2$ получаются примерно равные числа землетрясений, которые дали предвестниковый эффект, и которые его не дали. Нельзя сделать вывод, случайная ли вариация наблюдается в (-1) день, или скорее всего эта вариация вызвана процессами подготовки землетрясения.

В результате, однако, получилось, что с разумной вероятностью (например, с вероятностью выше 80%) можно выделить лишь 16 землетрясений, процессы подготовки которых вызывают эффект на (-1) день. Имеет смысл посмотреть, где эти землетрясения расположены и какими свойствами обладают.

Конечно, приведенные примеры предварительные. Следовало бы повторить компьютерное моделирование например, 1000 раз, и посмотреть уточненные цифры. Но ясно, что при принятой модели воздействия процессов подготовки на изучаемый параметр при выявлении событий, продуцировавших предвестниковый эффект, нет смысла рассматривать события с $A_{norm}(-1) < 2$.

Итак, мы получили, что если процессы подготовки землетрясений воздействуют в каждом втором случае, увеличивая измеряемый параметр на 10%, а всего анализируются 300 землетрясений, то мы можем отследить примерно полтора десятка землетрясений, т.е. 5%, которые с высокой степенью вероятности дали предвестниковый эффект. Более точно эту вероятность можно оценить с помощью моделирования случайного процесса методом Монте-Карло.

А если увеличение измеряемого параметра происходит для всех землетрясений на 10%, а не только для половины событий, что увидит исследователь? Промоделируем такую ситуацию. Создадим 100 серий, в каждой серии на нашей исходной последовательности А «включим действие» 100 виртуальных землетрясений, и пусть в (-1) день для каждого землетрясения $A(-1)$ увеличивается на 10%. Далее произведем вычитание скользящего среднего, пронормируем последовательность, получим значения A_{norm} . Значения $A_{norm}(-1)$ для 100 виртуальных землетрясений запоминаем для каждой серии. В следующих сериях процесс повторяем, используя исходную последовательность А.

По всем 100 массивам A_{norm} из 10 000 элементов каждый вычислим среднее = -0.04, медиана = -0.07, среднеквадратичное отклонение = 1.01. Среднее не равно нулю, потому что наша исходная последовательность чисел Вольфа была несимметричная, и мы получили после нормировки распределение близкое к нормальному, но не совсем нормальное.

Последовательности A_{norm} , полученные в различных сериях, немного различаются, потому что различаются дни виртуальных землетрясений, и, следовательно, значения A_{norm} в (-1) дни для этих виртуальных землетрясений.

Смоделированных $A_{norm}(-1)$ дней будет $100 \times 100 = 10000$. По совокупности таких дней среднее значение равно 0.35., т.е. оно значительно выше среднего значения по всем массивам, которое мы назовем «фоновым» средним значением.

Далее выясним, какая доля $A_{norm}(-1)$ в «сейсмоактивные» (-1) дни превышает «фоновое» среднее значение. В результате получается, что в 61% случаев $A_{norm}(-1)$ больше среднего значения, и, следовательно в 39% случаев $A_{norm}(-1)$ меньше среднего значения, или, иначе говоря, наблюдается приблизительно соотношение 3:2. Примерно такого порядка соотношения и получаются при исследовании сейсмоионосферных эффектов [3].

При той модели, которая здесь применялась – увеличении амплитуды исходного измеряемого параметра на 10 % и при числе землетрясений, равном 100, при наложении эпох получается, что амплитуда эффекта достигает 4 стандартных отклонений. На Рис. 35 представлен результат наложения эпох.

Но величина стандартного отклонения, которую мы получаем с помощью моделирования, зависит от числа событий – она обратно пропорциональна корню из числа событий, если событий много. Так что если событий будет примерно 50, а не 100, стандартное отклонение увеличится в 1.4 раза, и тогда амплитуда нашего эффекта будет превышать 2 стандартных отклонения, но будет меньше 3, что вполне реально в практике исследований.

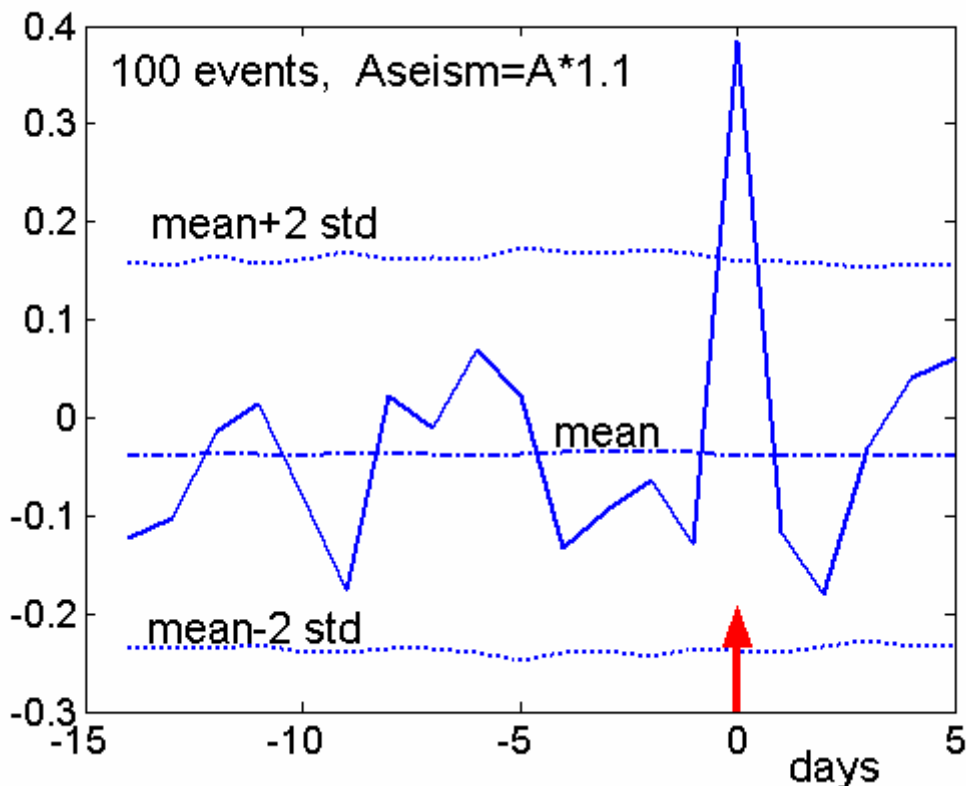


Рис. 35 Результат наложения эпох при $A_{seism}=A*1.1$ в день проявления сейсмозффекта –(-1) день

Поэтому исследователю не следует удивляться, если у него получилось, что только для 2/3 землетрясений вроде бы наблюдается изменение изучаемого параметра в

«нужную сторону», а для $1/3$ – не наблюдается. Нужно иметь в виду, что сравнительно слабое воздействие процессов подготовки землетрясений накладывается на случайные вариации исследуемого параметра.

До сих пор рассматривались модельные задачи на временных рядах без пропусков. Но при реальных исследованиях данные всегда с пропусками. Эффект, тем не менее, выявить можно, несмотря на пропуски, если данных много. Пропуски, уменьшая число данных, влияют на точность.

При применении моделирования фонового распределения методом Монте-Карло следует точно следовать алгоритму, примененному для выявления эффекта при наложении эпох. Например, если за рассматриваемый интервал произошли 200 землетрясений, а часть данных пришлось отбросить – например, данные в возмущенные дни – и для исследования осталось 60% данных, то при наложении эпох для реальных землетрясений для каждого дня будем располагать не 200 значениями, а приблизительно 120.

При создании серии с виртуальными землетрясениями следует проверять, совпадают ли числа данных для одного дня для данного набора виртуальных землетрясений с числом данных для реальных землетрясений. Желательно не допускать различия, большего 10-15%. Если же серия виртуальных землетрясений случайно оказалась столь неудачной, что число значений, используемых при наложении эпох, значительно отличается от аналогичного числа для реальных землетрясений, то полученные в такой серии значения не следует использовать при вычислении среднего и среднеквадратичного отклонения.

Выводы

1. Метод наложения эпох представляет собой мощный метод, позволяющий выделять слабый эффект на фоне шума. Достоверность выделения эффекта будет более высокой, если перед применением наложения эпох данные будут «очищены» от шумовых вариаций в не интересующем исследователя диапазоне характерных времен и «очищены» от вариаций, не интересующего исследователя происхождения.
2. Оценку достоверности результата проще всего провести методом моделирования фонового распределения – методом Монте-Карло.
3. В ряде случаев метод моделирования фонового распределения позволяет избежать методических ошибок, утрируя их.
4. Метод моделирования позволяет сделать некоторые заключения об амплитуде предполагаемого эффекта, поскольку подбирая различные амплитуды, мы как бы решаем обратную задачу.

Список литературы

- Липеровский В.А., Похотелов О.А., Шалимов С.Л.* Ионосферные предвестники землетрясений. – М.: Наука. 1992. – 304 с.
- Симонов П.В.* К вопросу о нарастающей тенденции синтеза наук // Методологические вопросы физиологии высшей нервной деятельности. – Л.: Наука. 1970. – 300с.
- Rishbeth H.* Ionoquakes: earthquake precursors in the ionosphere? // *Eos*. 2006. V.87, N32. P. 316-317.
- Pulinets S.A., Boyarchuk K.A.* Ionospheric precursors of earthquakes. Berlin: Springer, 2004. – 215 p.