

Программа PPPeriod_M

выделения периодических компонент интенсивности сейсмического режима путем оценки эволюции приращения максимума логарифмической функции правдоподобия в скользящем событийном окне.

А.А. Любушин, доктор физ.-мат. наук
Институт физики Земли РАН им. О.Ю.Шмидта,
123995, Москва, Большая Грузинская, 10; факс: +007-499-2556040;
e-mail: lyubushin@yandex.ru
<http://AlexeyLyubushin.narod.ru/Index.htm>

Целью программы является вычисление приращений максимума логарифмической функции правдоподобия для модели интенсивности точечного процесса с заданным периодом по отношению к максимуму логарифмической функции правдоподобия для модели чисто пуассоновского процесса с постоянной интенсивностью (для нулевой гипотезы). В отличие от программы **PPPeriod.exe** здесь оценка производится только в скользящем временном окне, причем используется так называемое **событийное окно**, то есть временное окно переменной длины, содержащее одинаковое число событий. Вследствие использования событийного окна временные метки в выводном grd-файле являются безразмерными и равными числу событий, соответствующему правому концу событийного окна. По этой же причине значения периодов также являются безразмерными и измеряются в числе событий. Значения периодов сканируются в фиксированных пределах от минимального периода, равного 1, до максимального, равного числу событий в окне.

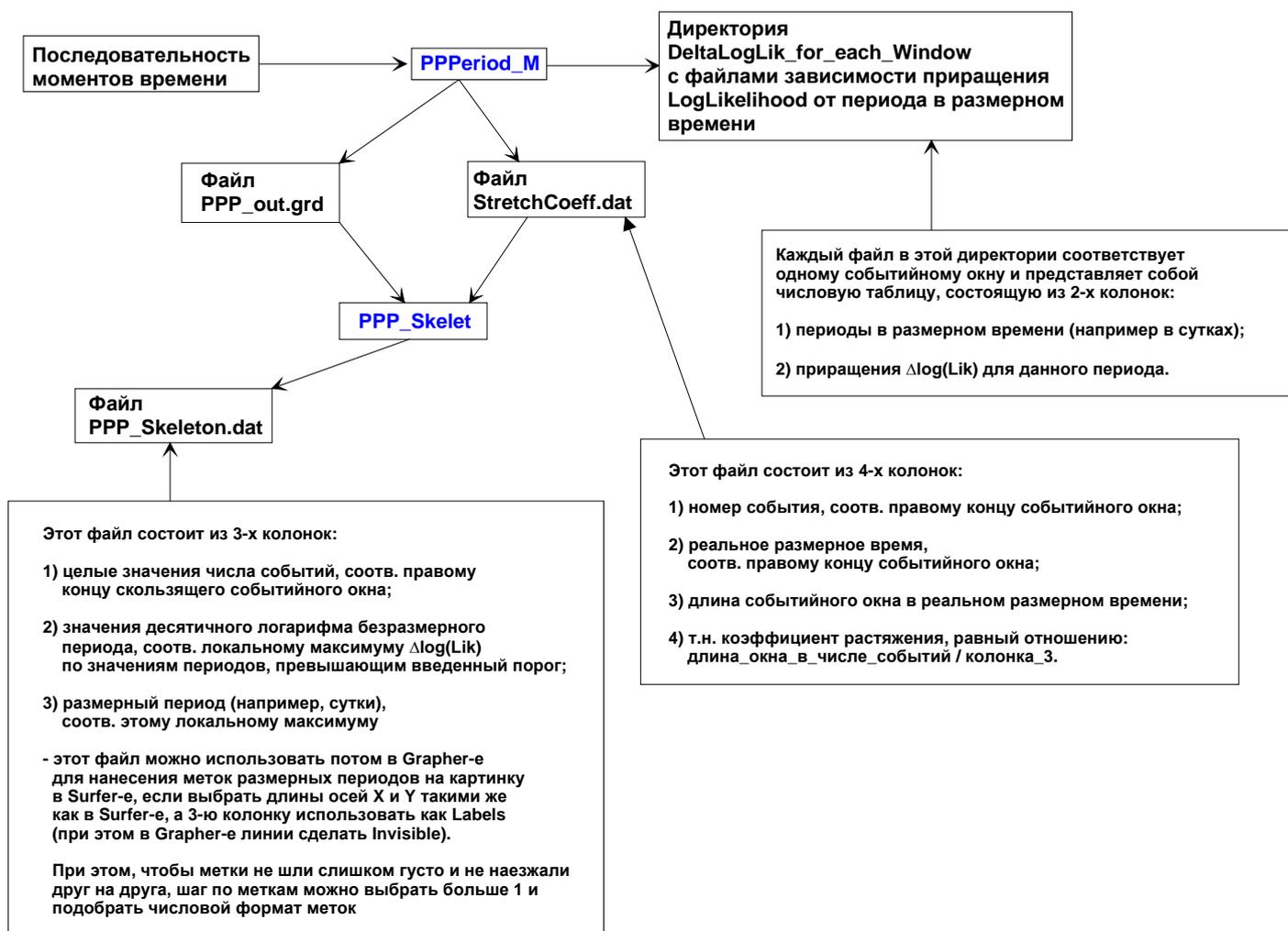
Входным файлом программе может служить любая символьная таблица чисел, в первой колонке которой стоят монотонно неубывающие числа, интерпретируемые как значения моментов времени. Прочие колонки входного файла, если таковые имеются, игнорируются.

После запуска программы необходимо ответить на следующие вопросы:

1. Имя входного файла. Файл должен находиться в той же директории, что и загрузочный модуль программы. Если файла нет или он имеет неправильную структуру (например, в первой колонке входной таблицы значения убывают для каких-то строк), то программа завершает работу с соответствующим сообщением.
2. Ввести значение числа пробных значений периодов, покрывающих сканируемый диапазон безразмерных периодов (от 1 до числа событий в окне) с шагом, равномерным в логарифмическом масштабе.
3. Задать объем событийного окна (число событий).
4. Задать смещение событийного окна (число событий).

Для связи безразмерных периодов и безразмерных временных меток, соответствующих правым концам событийных окон, их размерным значениям, предусмотрены, помимо основного выводного файла **PPP_out.grd**, дополнительные выводные файлы **StretchCoeff.dat** и выводимые в поддиректорию **"DeltaLogLik_for_each_Window"**. Для соответствия безразмерных периодов размерным необходимо воспользоваться вспомогательной программой **PPP_Skelet.exe**.

Ниже приведена блок-схема использования программ **PPPeriod.exe** и **PPP_Skelet.exe**, содержащая описание структуры всех выводных файлов.



Блок-схема применения программ и описание структуры всех выводных файлов.

Описание метода.

Рассматривается модель интенсивности сейсмического процесса в виде:

$$\lambda(t) = \mu \cdot (1 + a \cos(\omega t + \varphi)) \quad (1)$$

где частота ω , амплитуда a , $0 \leq a \leq 1$, фазовый угол φ , $\varphi \in [0, 2\pi]$ и множитель $\mu > 0$ (описывающий пуассоновскую часть интенсивности) являются параметрами модели. Таким образом, пуассоновская часть интенсивности модулируется гармоническим колебанием.

Зафиксируем какое-то значение частоты ω . Логарифмическая функция правдоподобия в этом случае для серии наблюдаемых событий равна:

$$\begin{aligned} \ln L(\mu, a, \varphi | \omega) &= \sum_{t_i} \ln(\lambda(t_i)) - \int_0^T \lambda(u) du = \\ &= N \ln(\mu) + \sum_{t_i} \ln(1 + a \cos(\omega t_i + \varphi)) - \mu T - \frac{\mu a}{\omega} [\sin(\omega T + \varphi) - \sin(\varphi)] \end{aligned} \quad (2)$$

Взяв максимум выражения (2) по отношению к параметру μ нетрудно найти что:

$$\mu = \hat{\mu}(a, \varphi | \omega) = \frac{N}{T + a(\sin(\omega T + \varphi) - \sin(\varphi)) / \omega} \quad (3)$$

Подставляя (3) в формулу (2) получаем

$$\ln(L(\hat{\mu}, a, \varphi | \omega)) = \sum_{t_i} \ln(1 + a \cos(\omega t_i + \varphi)) + N \cdot \ln(\hat{\mu}(a, \varphi | \omega)) - N \quad (4)$$

Следует заметить, что выражение $\hat{\mu}(a=0, \varphi | \omega) \equiv \hat{\mu}_0 = N/T$ является оценкой интенсивности процесса при условии, что он является однородным пуассоновским (чисто случайным).

Таким образом, приращение логарифмической функции правдоподобия вследствие рассмотрения более богатой, чем для чисто случайного потока событий, модели интенсивности с гармонической компонентой с заданной частотой ω равно:

$$\Delta \ln L(a, \varphi | \omega) = \sum_{t_i} \ln(1 + a \cos(\omega t_i + \varphi)) + N \cdot \ln(\hat{\mu}(a, \varphi | \omega) / \hat{\mu}_0) \quad (5)$$

Пусть

$$R(\omega) = \max_{a, \varphi} \Delta \ln L(a, \varphi | \omega), \quad 0 \leq a \leq 1, \varphi \in [0, 2\pi] \quad (6)$$

Функция (6) может рассматриваться как обобщение спектра для последовательности событий. График этой функции показывает насколько «более выгодна» периодическая модель интенсивности по сравнению с чисто случайной моделью. Максимальные значения функции (6) выделяют частоты, присутствующие в потоке событий.

Следующим очевидным обобщением метода является вычисление функции (6), используя наблюдаемые моменты времени не на всем интервале $(0, T]$, но внутри скользящего временного окна заданной длины L . Пусть τ – время правого конца скользящего временного окна. Тогда выражение (6) становится функцией от 2-х аргументов: $R(\omega, \tau | L)$, которая может быть визуализирована в виде 2-мерных карт или 3-мерных рельефов на плоскости аргументов (ω, τ) . Эта частотно-временная диаграмма позволяет исследовать динамику возникновения и развития периодических компонент внутри исследуемого потока событий.

До сих пор использовалась программа **PPPeriod**, в которой длина окна бралась фиксированной в физическом, размерном, времени. Ниже приводятся примеры применения той же оценки (6), но для скользящего событийного окна, то есть, окна, включающего в себя фиксированное число событий. Использование событийного окна одного и того же объема позволяет более корректно, по сравнению с обычным временным окном фиксированной длины в физическом времени (и, как следствие этого, с разным числом событий внутри каждого окна), проводить сравнение статистической значимости пиков приращения максимумов логарифмической функции правдоподобия для различных периодов и окон.

Пусть для одного и того же набора данных $X^{(N)}$, состоящего из N независимых наблюдений, рассматриваются 2 гипотезы:

- 1) гипотеза H_0 : $X^{(N)}$ распределена в соответствии с плотностью $p_0(X^{(N)} | \theta_0)$;
- 2) гипотеза H_1 : $X^{(N)}$ распределена в соответствии с плотностью $p_1(X^{(N)} | \theta_1)$.

где θ_0 и θ_1 - векторы неизвестных параметров, имеющих размерности m_0 и m_1 , причем гипотеза H_1 является более «богатой»: $m_1 > m_0$, а вектор параметров θ_1 полностью включает в себя компоненты вектора θ_0 . Рассмотрим разницу между логарифмами правдоподобий для этих двух гипотез, при условии, что для векторов параметров взяты их оценки метода максимального правдоподобия:

$$\Delta \ln L(X^{(N)}) = \ln \left(\max_{\theta_1} p_1(X^{(N)} | \theta_1) \right) - \ln \left(\max_{\theta_0} p_0(X^{(N)} | \theta_0) \right) \quad (7)$$

Очевидно, что $\Delta \ln L(X^{(N)}) \geq 0$. Согласно теореме Уилкса, при условии справедливости гипотезы H_0 величина (7) имеет асимптотическое распределение:

$$\Delta \ln L(X^{(N)}) \sim \frac{\chi_m^2}{2}, \quad m = m_1 - m_0, \quad N \rightarrow \infty \quad (8)$$

В нашем случае разница между числом параметров $m = 2$ и, следовательно, удвоенная величина (8) имеет асимптотическую плотность распределения χ_2^2 , равную $e^{-x/2}/2$, а сама величина (8) распределена асимптотически с экспоненциальной плотностью e^{-x} или

$$\Pr\{\Delta \ln(L) < x\} = 1 - e^{-x}, \quad N \rightarrow \infty \quad (9)$$

при условии, что анализируемая последовательность моментов времени распределена согласно пуассоновскому закону с постоянной интенсивностью. Ниже в примерах используется $N = 200$. Считая это число достаточно большим, получим, что для порога 4 вероятность случайного превышения равна 0.018. Таким образом, уровень значимости максимумов $\Delta \ln(L)$ для различных тестируемых периодов, превышающих уровень 4, можно принять за 98%.

ЛИТЕРАТУРА.

- Любушин А.А., В.Ф.Писаренко, В.В.Ружич, В.Ю.Буддо.* (1998) Выделение периодичностей в сейсмическом режиме. – Вулканология и сейсмология. N1. С.62-76.
- Любушин А.А.* (2007) «Анализ данных систем геофизического и экологического мониторинга». М.: Наука, 2007, 228с.
- Cox D.R., Lewis P.A.W.* (1966) The statistical analysis of series of events. London, Methuen (Русский перевод: *Кокс Д., П.Льюис* (1969) Статистический анализ последовательностей событий. М., Мир. 312с.)
- Rao C.R.* (1965) Linear statistical inference and its applications. John Wiley & Sons, Inc. N.Y., London, Sydney (Русский перевод: *Рао С.Р.* (1968) Линейные статистические методы и их применение. М., Наука. 548 с.).
- Wilks S.S.* (1962) Mathematical statistics. John Wiley & Sons, Inc. N.Y., London, Sydney. (Русский перевод: *Уилкс С.* (1967) Математическая статистика. М., Наука. 632 с.).

Примеры применения

Применим программу `PPPeriod_M`, в которой длина окна в числе событий = 200, смещение = 5, сканируются 200 периодов автоматически от 1. Далее приводятся примеры применения программы `PPPeriod_M` к анализу сейсмического процесса после сильных землетрясений, когда сильная изменчивость интенсивности является серьезным препятствием для применения оценок в скользящем окне постоянной временной длины.

Суматранское землетрясение

Рассмотрим последовательность из 2606 землетрясений с магнитудой ≥ 4.5 , происшедших в окружности радиуса 5 градусов с центром в эпицентре Суматранского землетрясения 24.12.2004 (3.32° С.Ш., 95.85° В.Д.) сразу после катастрофы в течение примерно 1130 суток, представленная на рис.1.

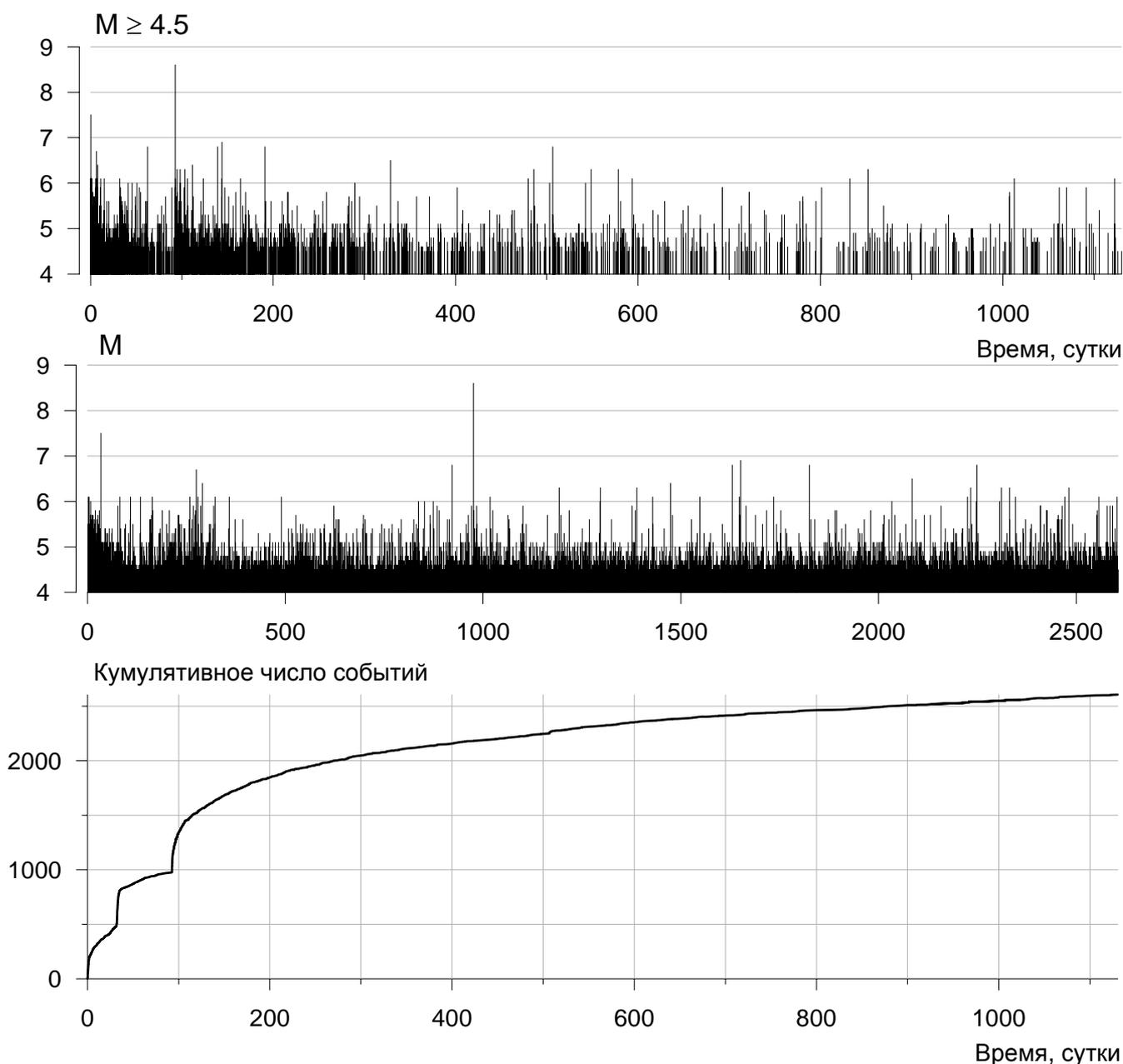
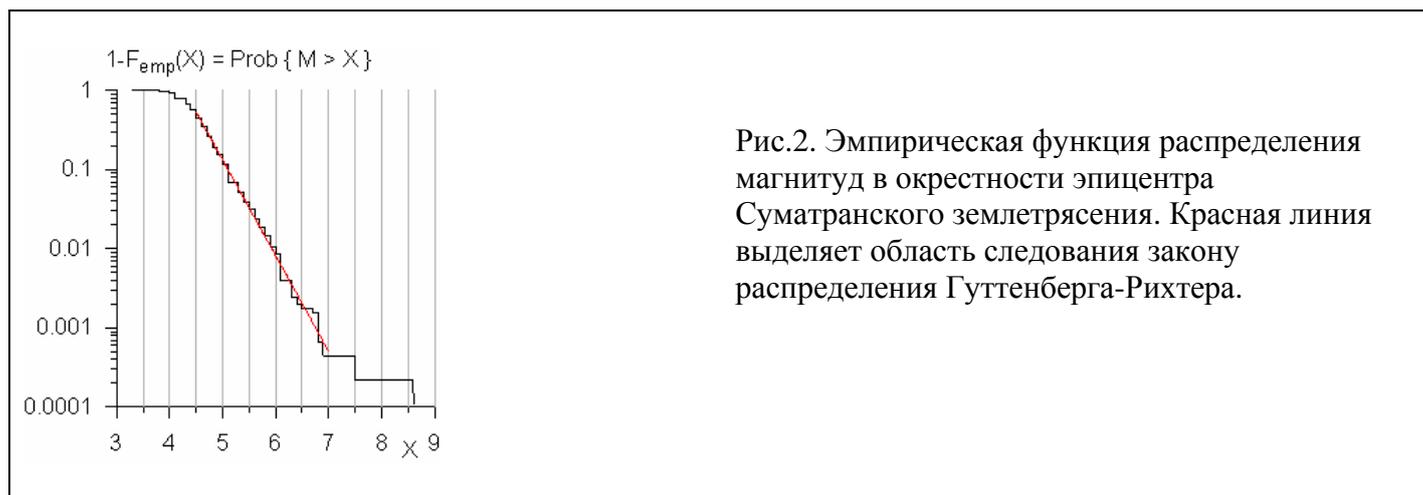


Рис.1.

Верхний график на рис.1 представляет собой последовательность магнитуд в зависимости от времени в сутках, прошедших от главного толчка, следующий график – то же самое, но в зависимости от прошедшего числа событий.

Выбор ограничения по магнитуде $M \geq 4.5$ связан с тем, что лишь начиная с 4.5 значения магнитуды в этом регионе являются статистически представительными и их распределение удовлетворяет экспоненциальному закону Гуттенберга-Рихтера. Это видно из графика на рис.2, где представлена эмпирическая функция распределения магнитуд:



Частотно-временная диаграмма эволюции $\Delta \ln(L)$ представлена на рис.3.

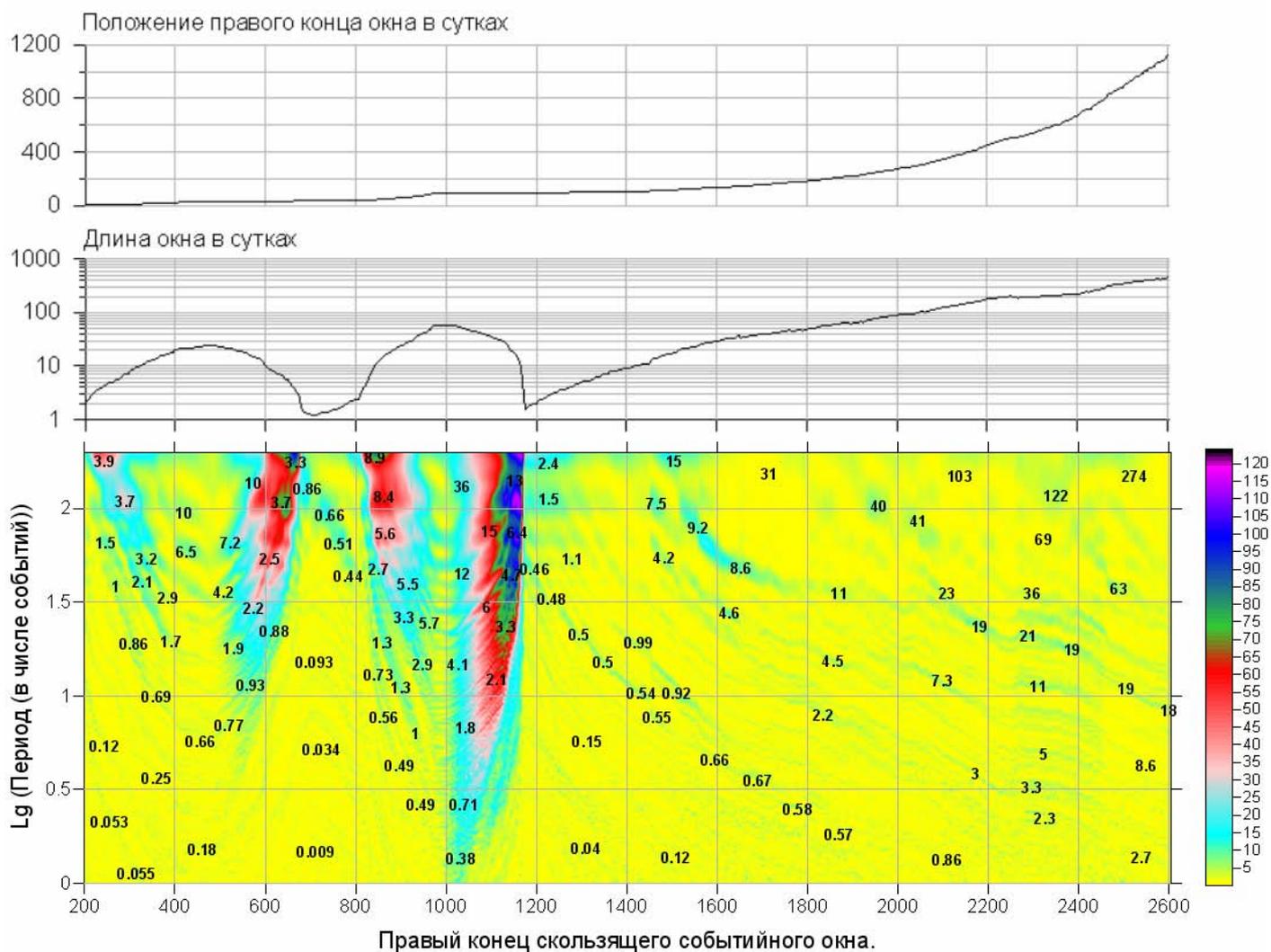


Рис.3. Черными цифровыми метками на частотно-временной диаграмме проставлены значения периодов в сутках.

Частотно-временная диаграмма эволюции $\Delta \ln(L)$ в скользящем событийном окне содержит непрерывные линии максимальных значений, вдоль которых физический период может быть как постоянным, так и изменяться, вследствие изменчивости длины временного окна. Поэтому далее эти диаграммы будем снабжать числовыми метками, дающими значения периода в сутках в данной точке диаграммы. В основном нас будут интересовать линии локальных максимальных значений («хребты» диаграммы).

Рассмотрим более детально фрагмент частотно-временной диаграммы для начального участка, до кумулятивного числа событий 500, представленный на рис.4.

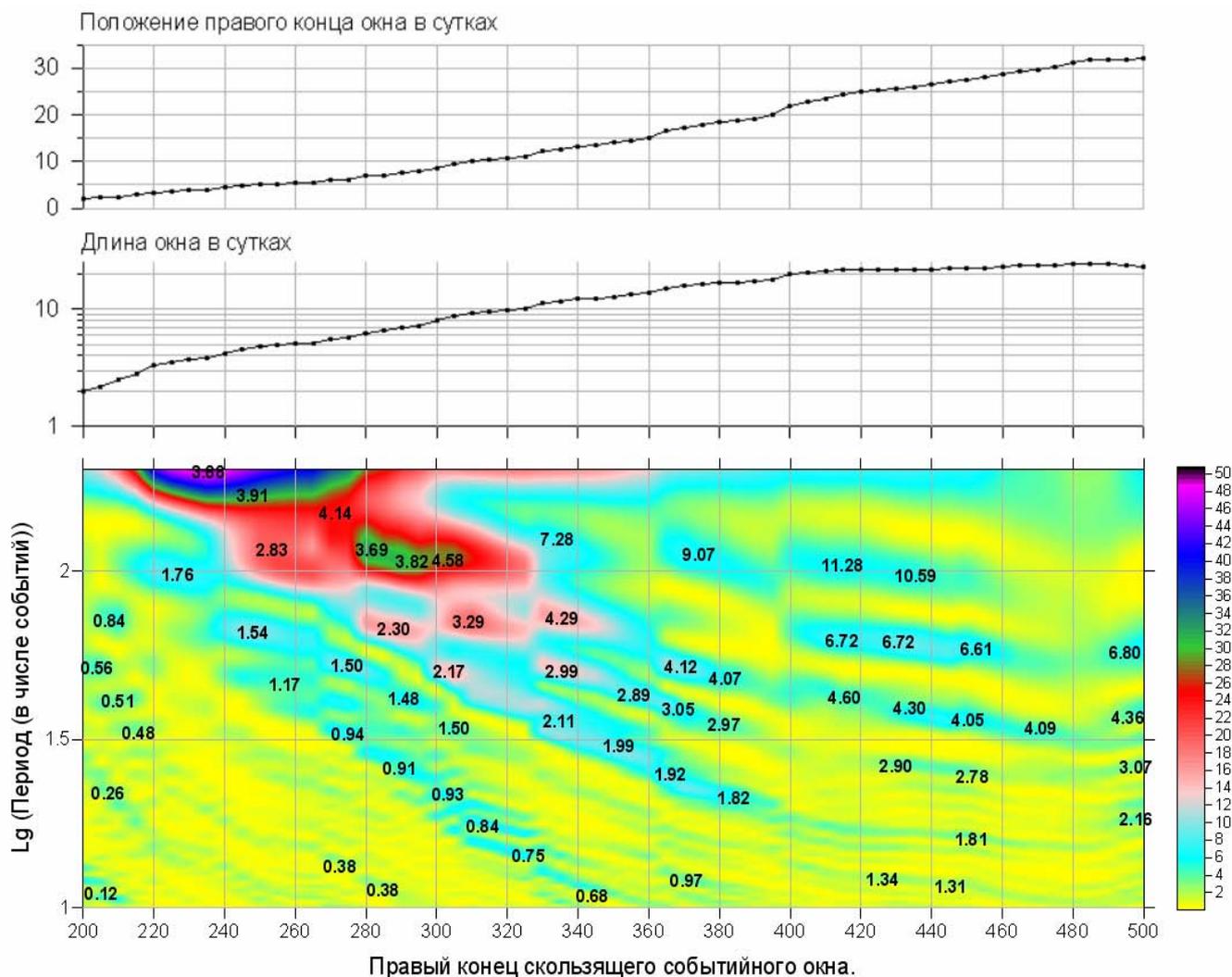


Рис.4. Черными цифровыми метками на частотно временной диаграмме проставлены значения периодов в сутках.

Видно, что максимум величин $\Delta \ln(L)$ для событийных окон с метками 210-270 приходится на периоды около 2.8-4.6 суток, причем этот максимум не расщепляется на отдельные гармоники в силу того, что физическая длина окна слишком коротка и поэтому разрешающая способность метода здесь недостаточна. Далее, по мере роста длины окна, метод начинает расщеплять гармоники и выделять колебания с отделенными значениями периодов. В частности, прослеживаются хребты значений с периодами от 1.8 до 0.7 суток, от 2.8 до 1.8 суток и т.д.

Частотно-временная диаграмма на рис.4 имеет вид начального «сгустка» периодов, который потом распадается на серию расходящихся лучей. Если округлить до целого числа среднее значение периодов вдоль каждого из лучей, до этим лучам можно присвоить имена: 1, 2, 3, 4.2, 6.6 и 10

Выделим теперь фрагмент для кумулятивных чисел событий от 500 до 800:

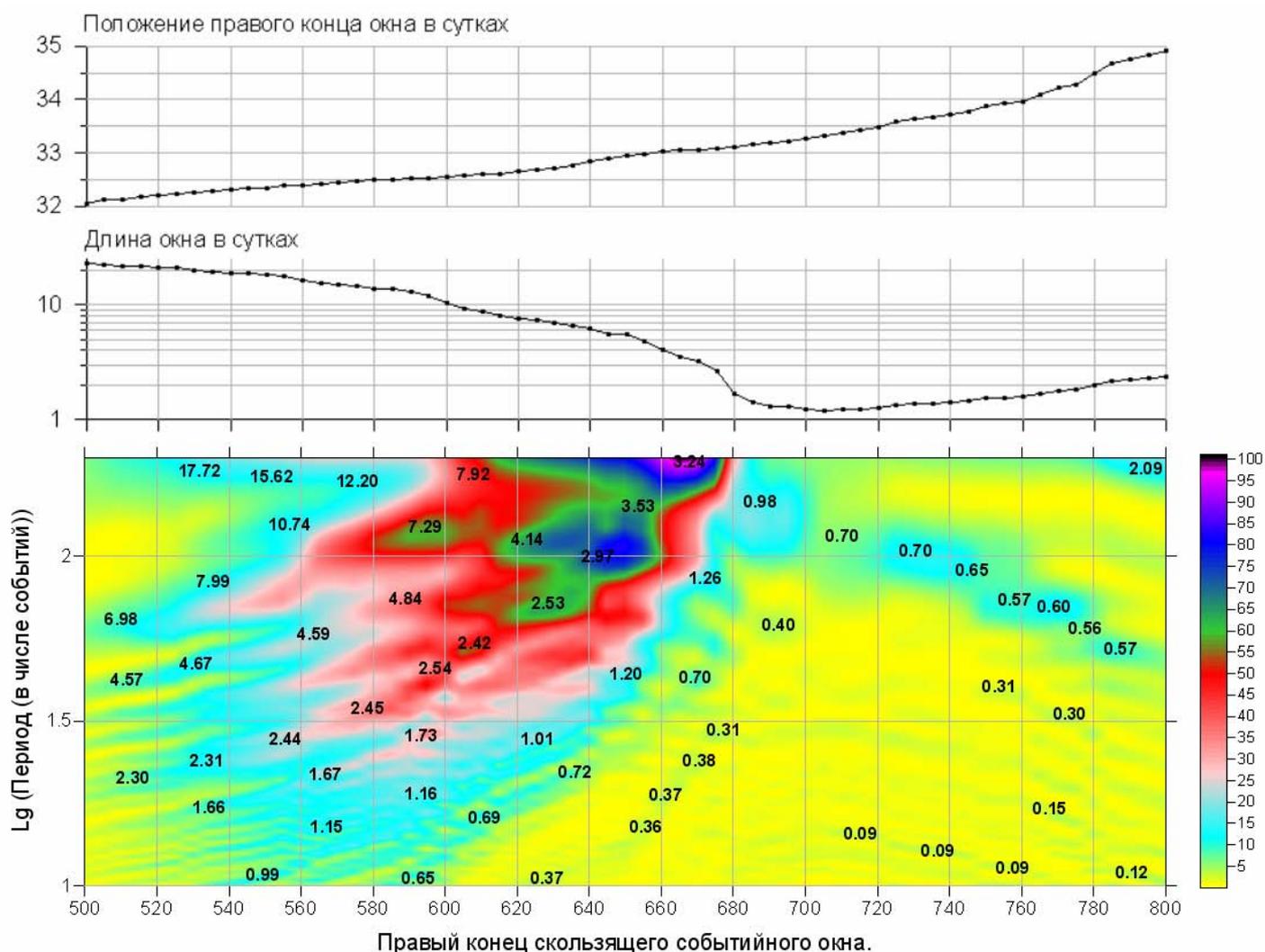


Рис.5. Черными цифровыми метками на частотно временной диаграмме проставлены значения периодов в сутках.

Судя по быстрому уменьшению длины окна на среднем графике рис.6, на этом участке возник рой землетрясений средней силы (рис.1). Из этой диаграммы видно, что статистически значимое колебание интенсивности с периодам от 3 до 10 суток возникают в окне 510, а далее эти спектральные пики расширяются и для окон 620-680 сливаются в сплошной максимум. Примечательная особенность поведения сейсмичности для этих окон заключается в том, что, начиная с окна 690, формируется периодичность с мигрирующим периодом от 1 до 0.56 суток.

Выделим далее фрагмент для кумулятивных чисел событий от 800 до 1000:

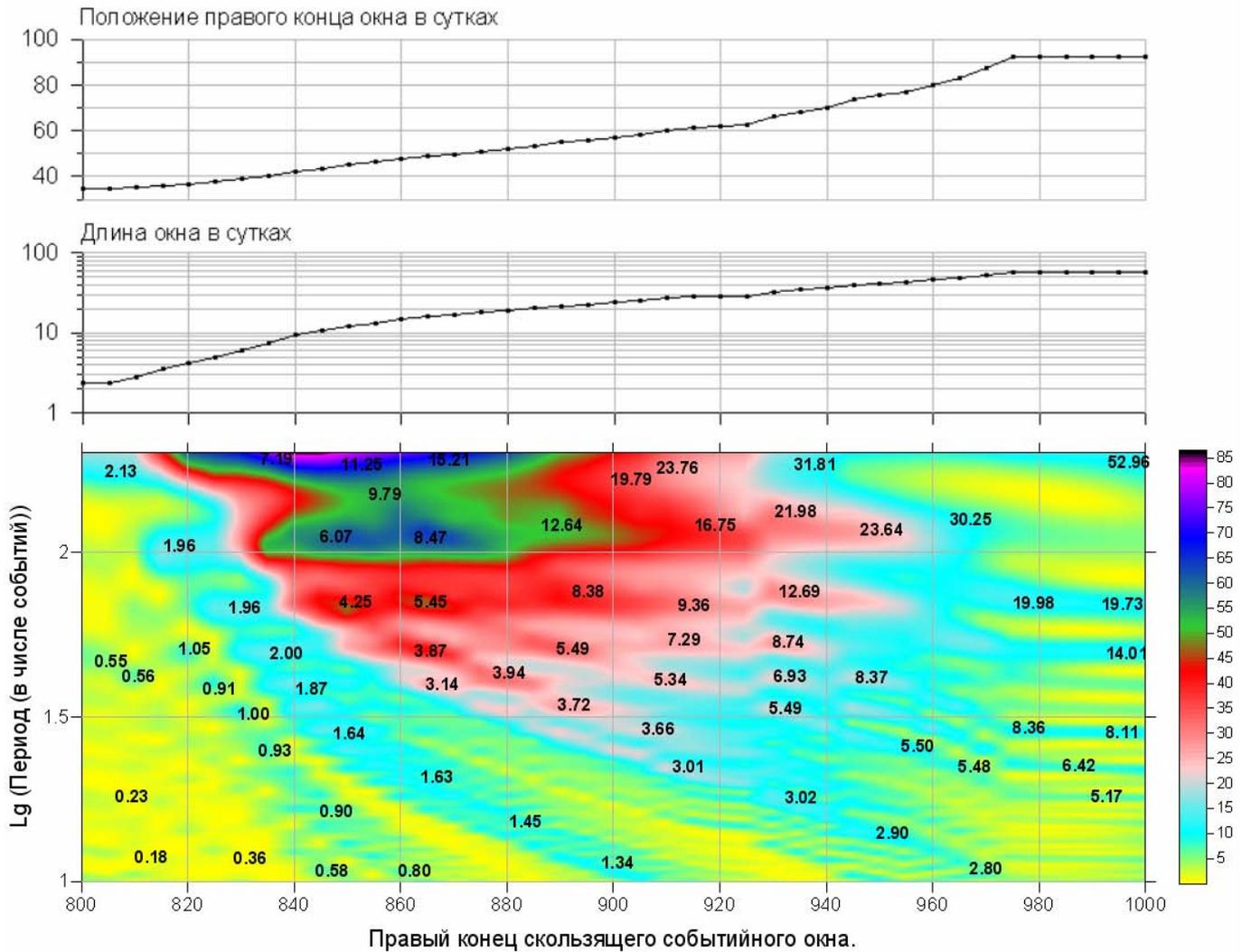


Рис.6. Черными цифровыми метками на частотно временной диаграмме проставлены значения периодов в сутках.

На этом фрагменте длина временного окна быстро растет по мере его выхода из роя, а затем возникает всплеск периодических компонент интенсивности. Далее длина окна выходит на практически постоянное значение, причем одновременно с этим постоянным становится и положение правого конца временного окна. Это говорит о том, что окно снова вошло в область высокой интенсивности. Этот всплеск интенсивности является началом афтершоковой активности после сильного землетрясения с магнитудой 8.6, события за номером 976. Отсюда следует, что рост периодических компонент интенсивности на фоне роста длины временного окна, представленный на рис.6, может рассматриваться как предвестник этого сильного землетрясения.

Из меток величин размерного периода, представленных на рис.6, следует, что очень малая длина окна в самом начале афтершоков приводит к потере разрешающей способности спектрального анализа и к тому, что образуется один сплошной спектральный максимум на условно низких частотах (окна 800-910). Далее, по мере роста длины окна, в этом сплошном спектральном максимуме начинает проглядываться тонкая структура в виде отдельных локальных экстремумов приращения максимума функции правдоподобия. В частности, выделяется период 12-15 суток (окна 910-925) и 25-30 суток (окна 940-980).

Начиная с окна 1000 (рис.3), наблюдается рост низкочастотных периодических компонент, что является проявлением афтершоков сильного события 18.03.2005, $M=8.6$, (2.09° С,Ш., 97.11° В.Д.),

номер события – 976. Как уже отмечалось, всплеск периодических компонент интенсивности, представленный ранее на рис.6, может рассматриваться как проявление форшоковой активности перед этим сильным событием. Исследуем более детально сейсмический процесс во временной окрестности этого землетрясения.

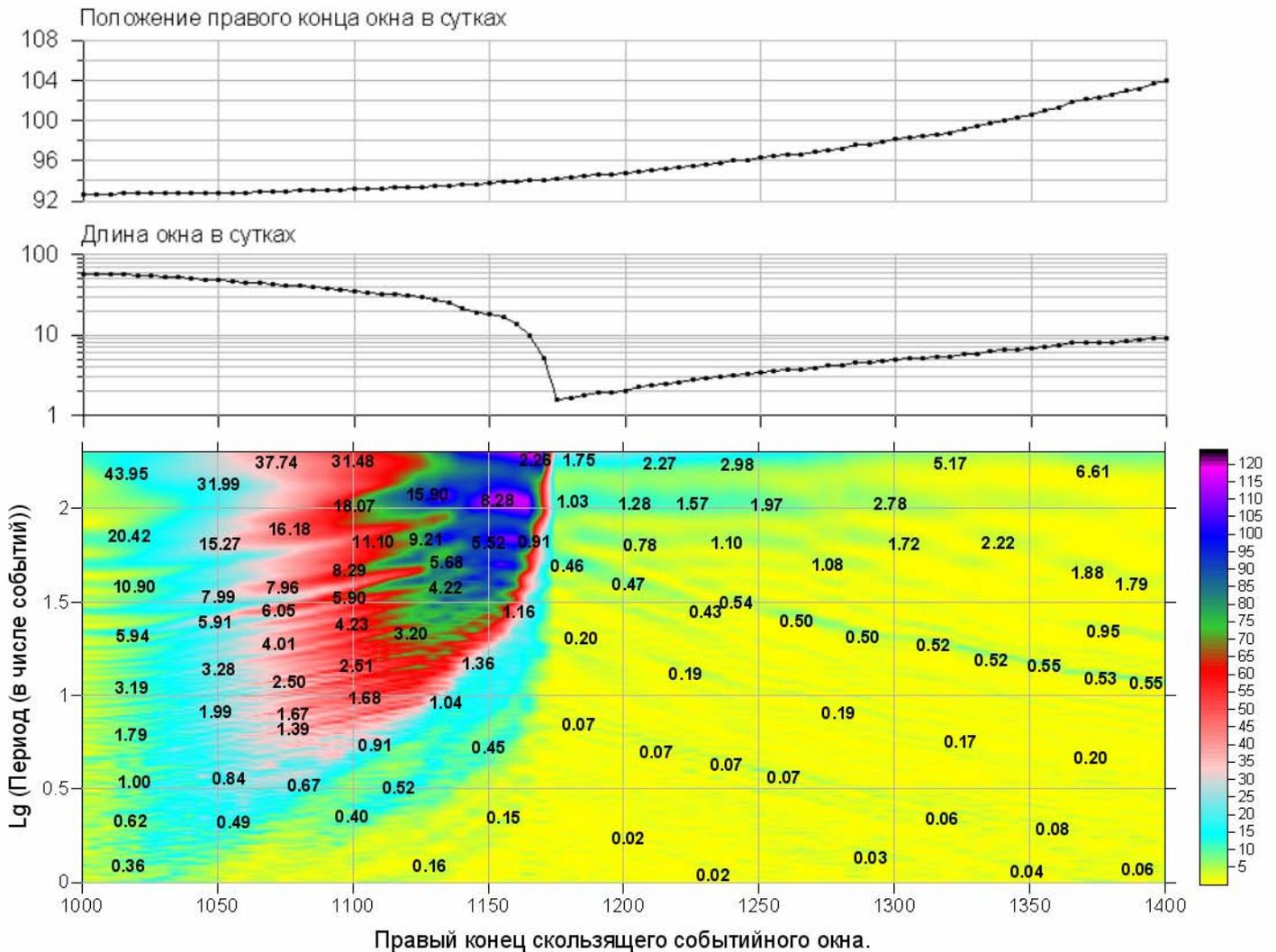


Рис.7. Черными цифровыми метками на частотно временной диаграмме проставлены значения периодов в сутках.

Из частотно-временной диаграммы на рис.7 видно, что, начиная с окна 1050, в афтершоках возникают низкочастотные периодические составляющие, которые исчезают после окна 1200. Прежде всего следует отметить, что по мере продвижения окна вглубь афтершоков, но при условии, что начальная часть окна все еще занята событиями перед сильным землетрясением, приводит к тому, что оценка приращения $\Delta \ln(L)$ приобретает сильные низкочастотные максимумы (окна 1050-1150). После того, как событийное окно уже большей частью или даже целиком состоит из афтершоков, величины $\Delta \ln(L)$ образуют сплошной спектральный максимум с периодами порядка длины окна (окна 1160-1240). Однако в них начинает выделяться тонкая структура с периодами от 0.5 до 1.5 суток (окна 1180-1240). Далее, после того как окно выходит из области очень интенсивных афтершоков (окна 1260-1400), опять становятся заметными значимые монохроматические узкие спектральные пики на периодах, близких к 0.5 суток, то есть к приливному периоду 12 часов.

Наконец, рассмотрим финальную часть выборки, для кумулятивных чисел событий начиная с 1400, которая характеризуется более спокойным, «фоновым» поведением (рис.8):

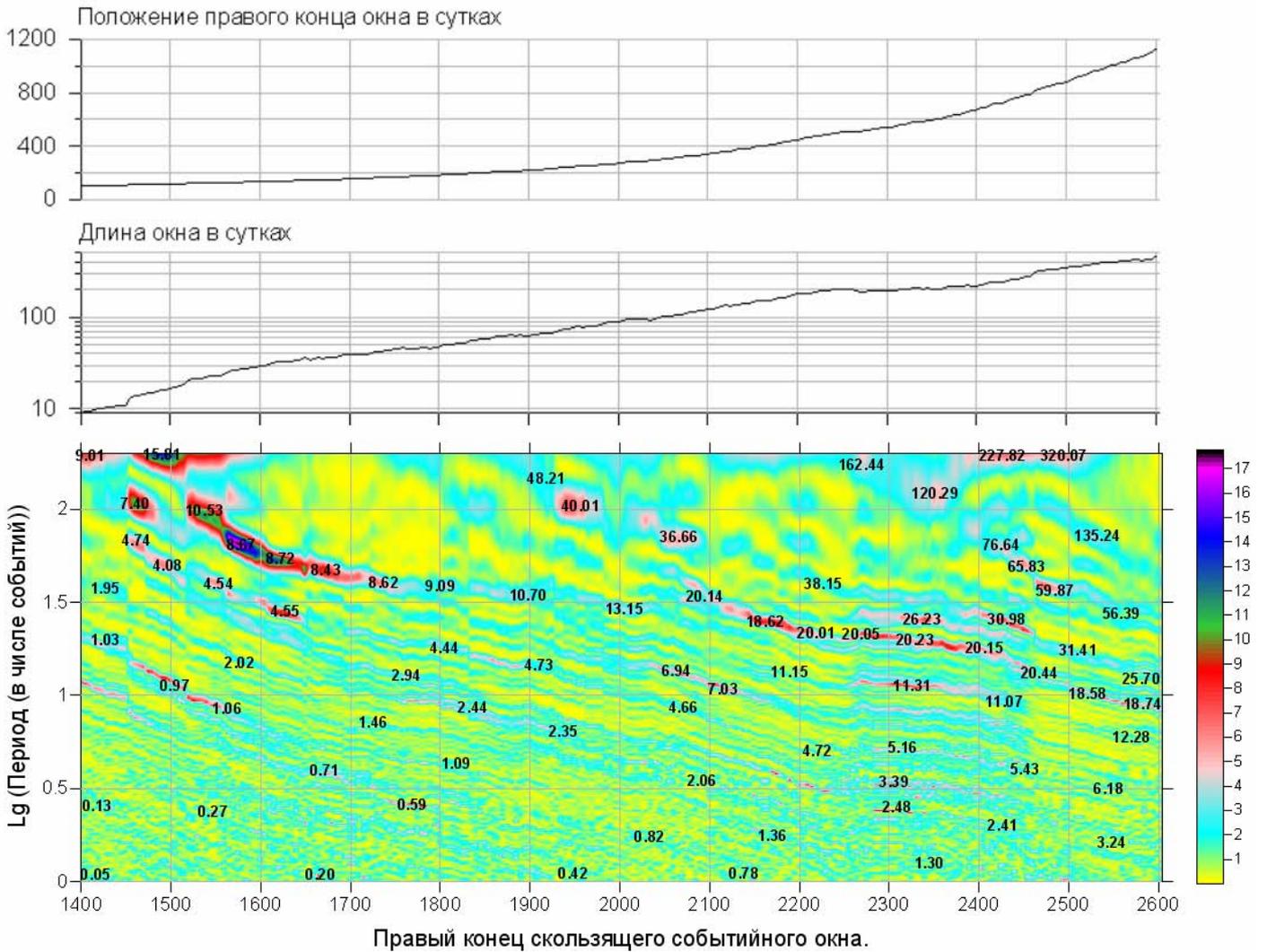


Рис.8. Черными цифровыми метками на частотно временной диаграмме проставлены значения периодов в сутках.

Из частотно-временной диаграммы видно, что есть значимые всплески периодических компонент для периодов от 8 до 15 суток и от 18 до 21 суток, образующие непрерывные линии максимумов. Полусуточная гармоника, отчетливо выделяемая на предыдущей диаграмме на рис.7, здесь уже является незначимой.

Землетрясение на Хоккайдо

Рассмотрим последовательность из 1135 времен событий (сутки) с магнитудой ≥ 4.0 в окружности радиуса 5 градусов с центром в эпицентре землетрясения около о. Хоккайдо 25.09.2003, $M=8.3$, (41.81° С.Ш., 143.91° В.Д.) и глубиной эпицентра не более 100 км – рис.9. Время в сутках отсчитывается сразу после момента главного толчка.

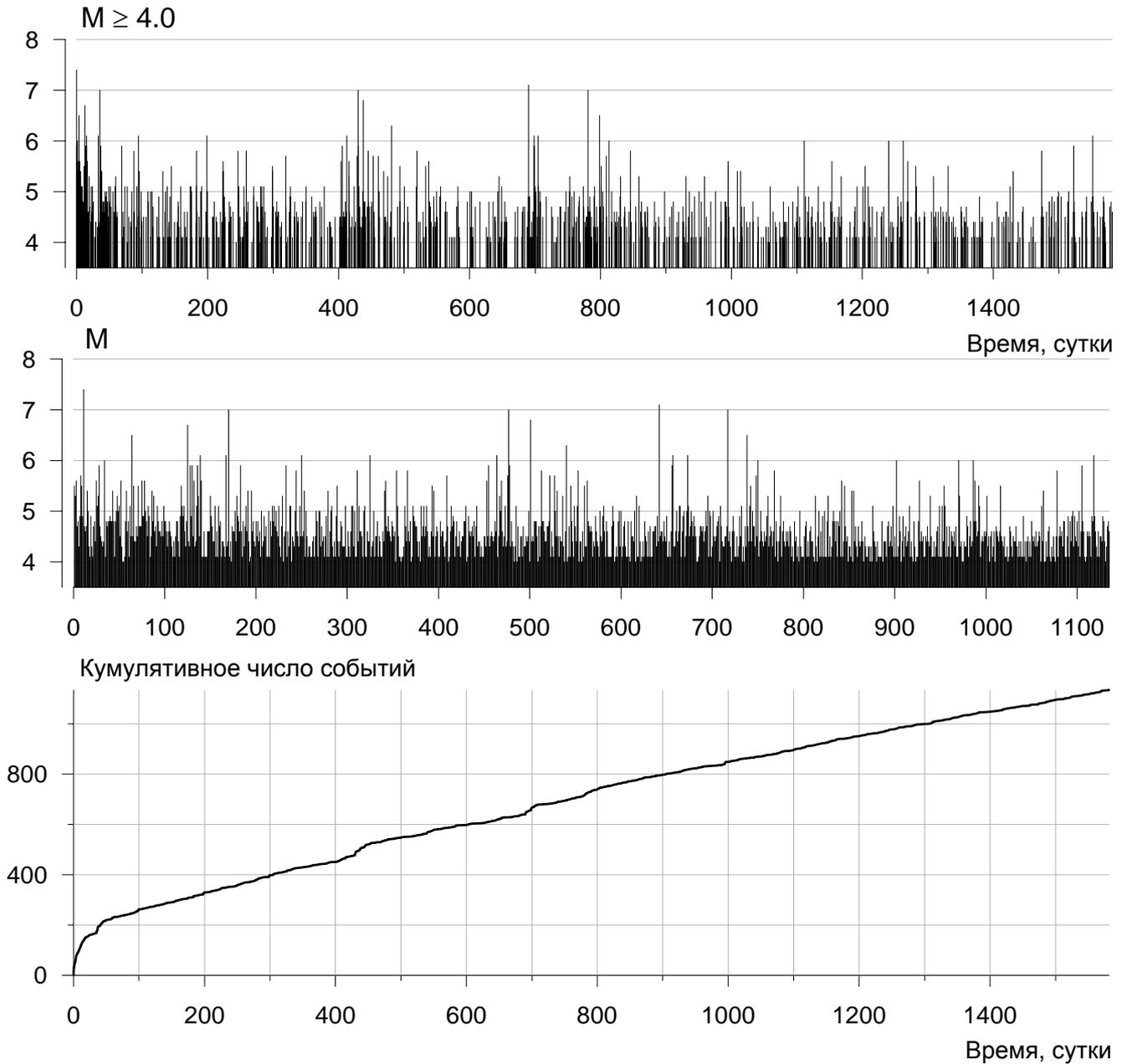
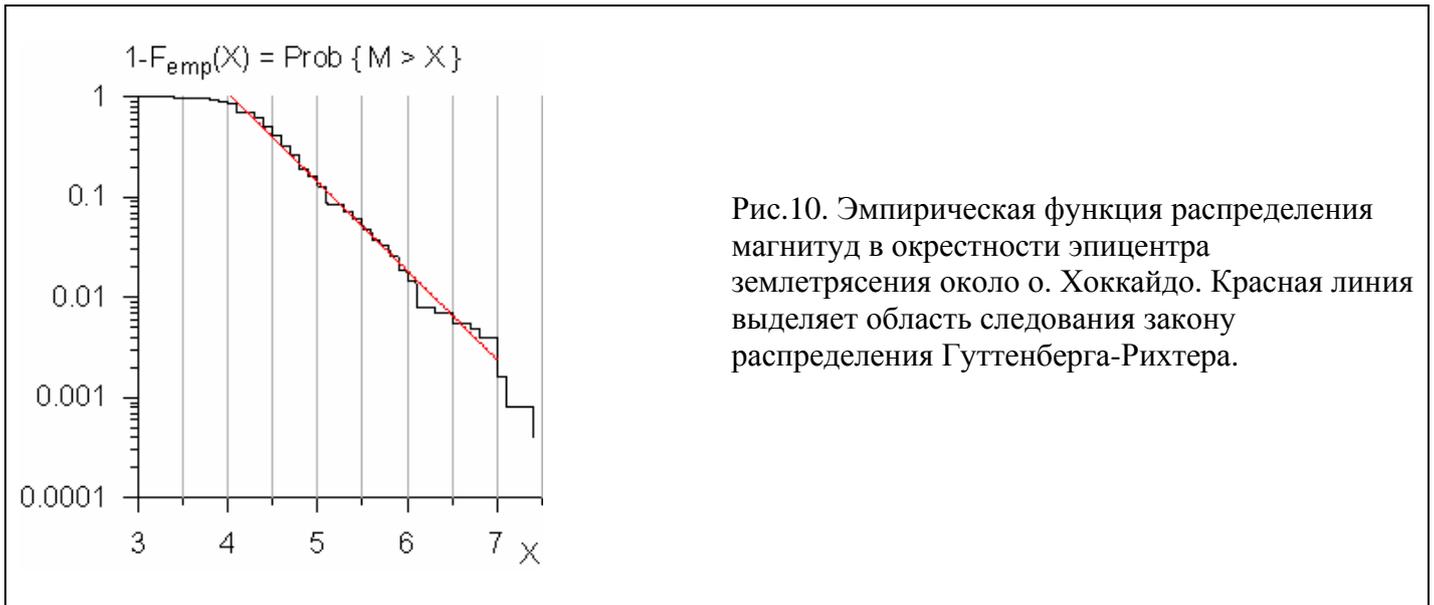


Рис.9.

Как видно из графика на рис.10, в данном случае нижний порог по магнитуде был выбран равным 4.0, поскольку с этого значения начинается соответствие распределения магнитуд закону Гуттенберга-Рихтера.



Частотно-временная диаграмма для эволюции $\Delta \ln(L)$ представлена на следующем рис.11.

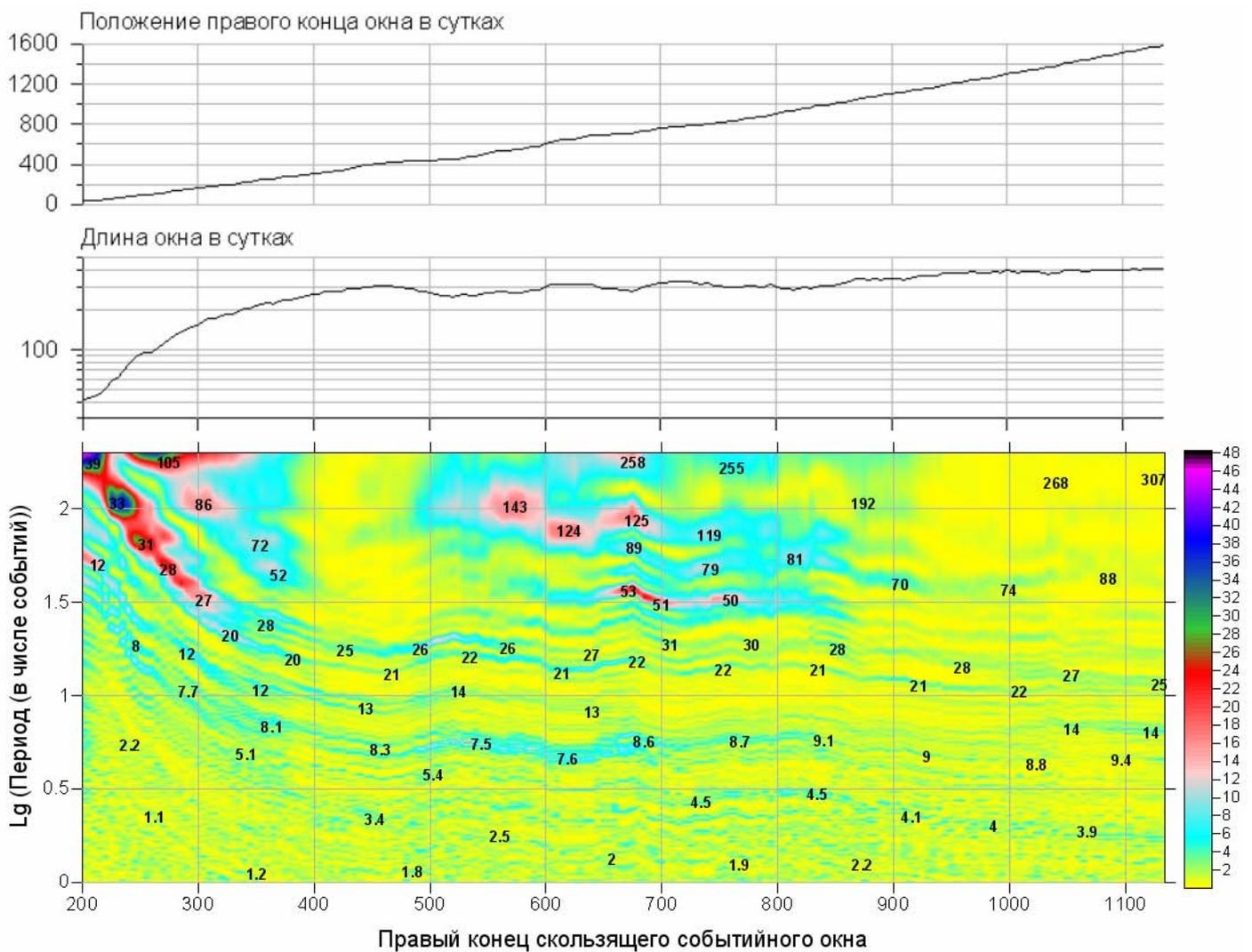


Рис.11. Черными цифровыми метками на частотно временной диаграмме проставлены значения периодов в сутках.

После окна 400 сейсмический режим выходит на стационарное поведение с примерно одинаковой длиной окна. Основная особенность графиков на рис.15 – это наличие сильной периодичности интенсивности афтершоков с периодом 25-32 суток. Отметим, что эти периоды близки к 28 суткам – длительности лунного месяца, который, как известно, является одним из наиболее часто встречающихся периодов, модулирующих геофизические процессы различной природы.

Курильское (Симуширское) землетрясение 2006 г.

Рассмотрим последовательность из 1561 времен событий (сутки) с магнитудой ≥ 4.0 в окружности радиуса 5 градусов с центром в эпицентре Симуширского землетрясения 15.11.2006, $M=8.3$, (46.59° С.Ш., 153.27° В.Д.) и глубиной эпицентра не более 100 км – рис.12. Время в сутках отсчитывается сразу после момента главного толчка.

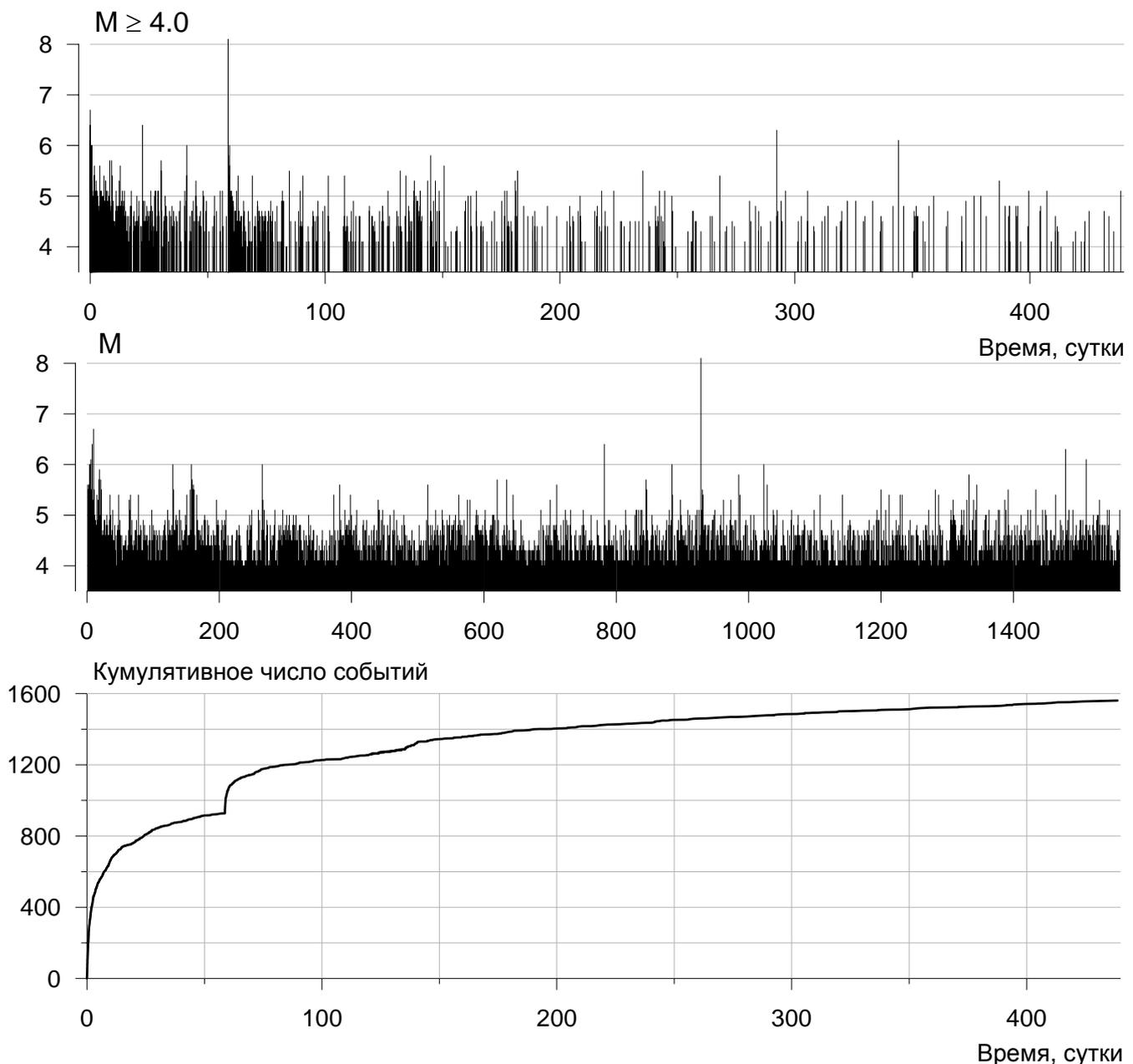
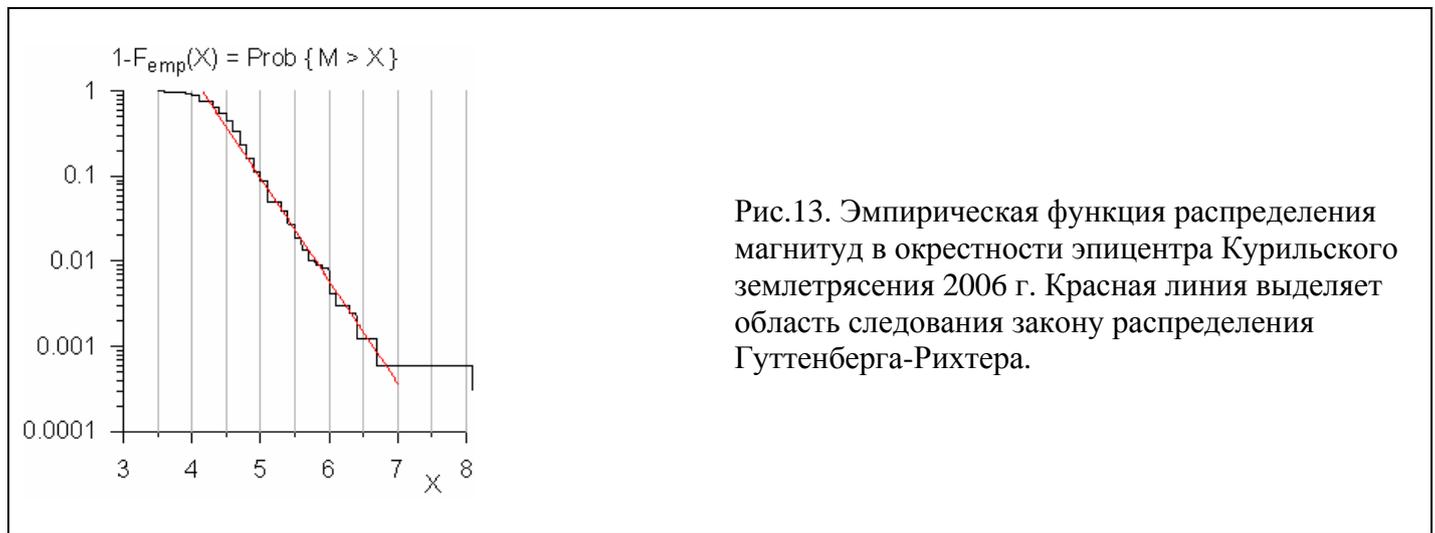


Рис.12.

Как видно из графика на рис.13, в данном случае нижний порог по магнитуде был также выбран равным 4.0, поскольку с этого значения начинается соответствие распределения магнитуд закону Гуттенберга-Рихтера.



Частотно-временная диаграмма для эволюции $\Delta \ln(L)$ представлена на следующем рис.14.

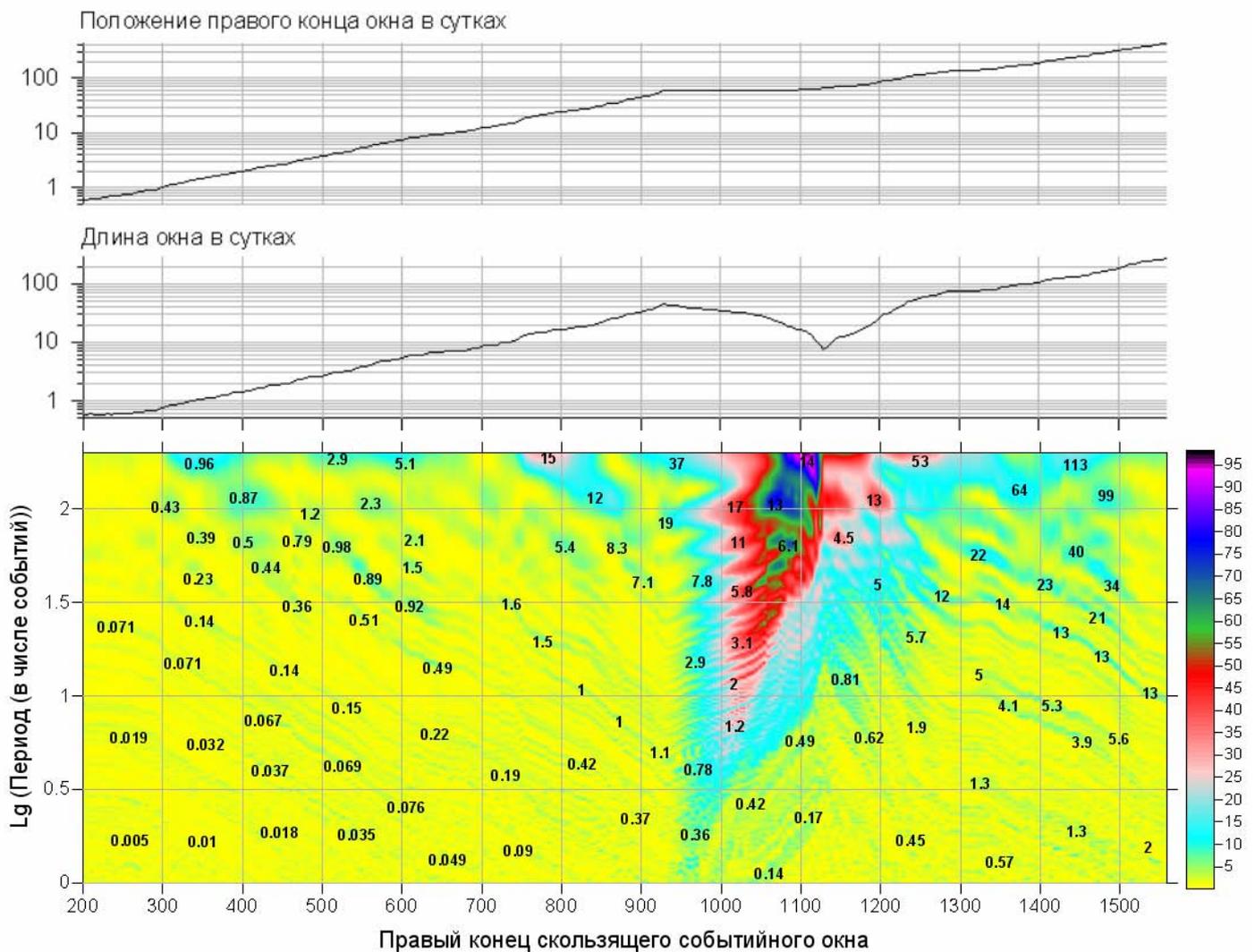


Рис.14. Черными цифровыми метками на частотно временной диаграмме проставлены значения периодов в сутках.

Как видно на рис.12, исследуемая последовательность землетрясений содержит сильное событие с магнитудой 8.1 за номером 928 – событие 13.01.2007 (46.24° С.Ш., 154.52° В.Д.). Наличие такого сильного события и последующей серии афтершоков видно и на частотно-временной диаграмме на рис.14. Исследуем подробнее фрагменты этой диаграммы. Начнем с интервала окон 200-900:

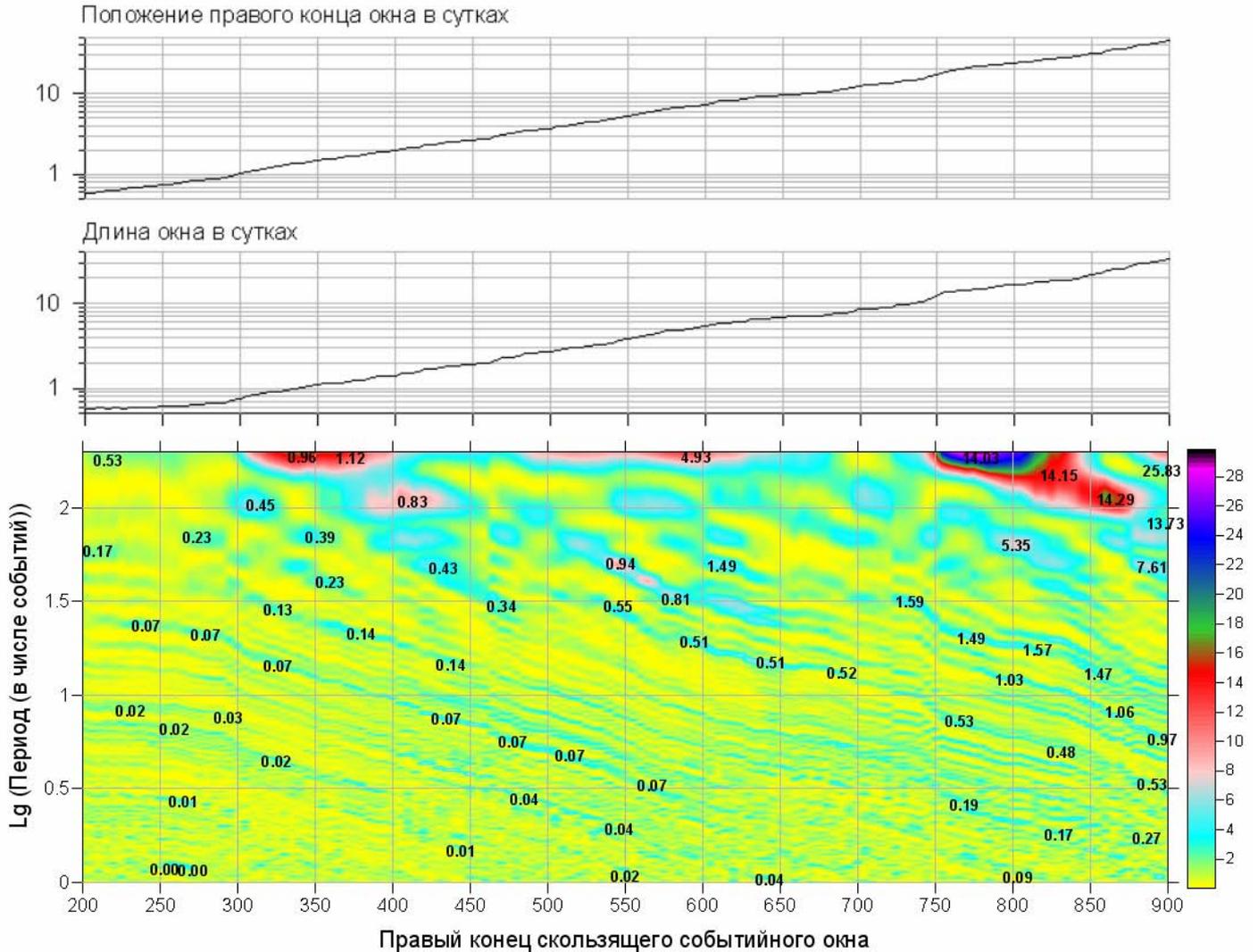


Рис.15. Черными цифровыми метками на частотно временной диаграмме проставлены значения периодов в сутках.

Эта диаграмма, подобно диаграмме на рис.6, содержит всплеск периодических компонент интенсивности на низких частотах (здесь – с периодами около 14 суток) на фоне роста длины окна также перед сильным землетрясением. Следовательно, эту аномалию можно рассматривать как его предвестник. Из интересных особенностей этой диаграммы следует отметить устойчивые периоды 0.07 и 0.48-0.55 суток, хотя их статистическая значимость находится на пороге 98%-ой.

Рассмотрим теперь детальнее фрагмент 900-1300, содержащий вышеупомянутое сильное землетрясение:

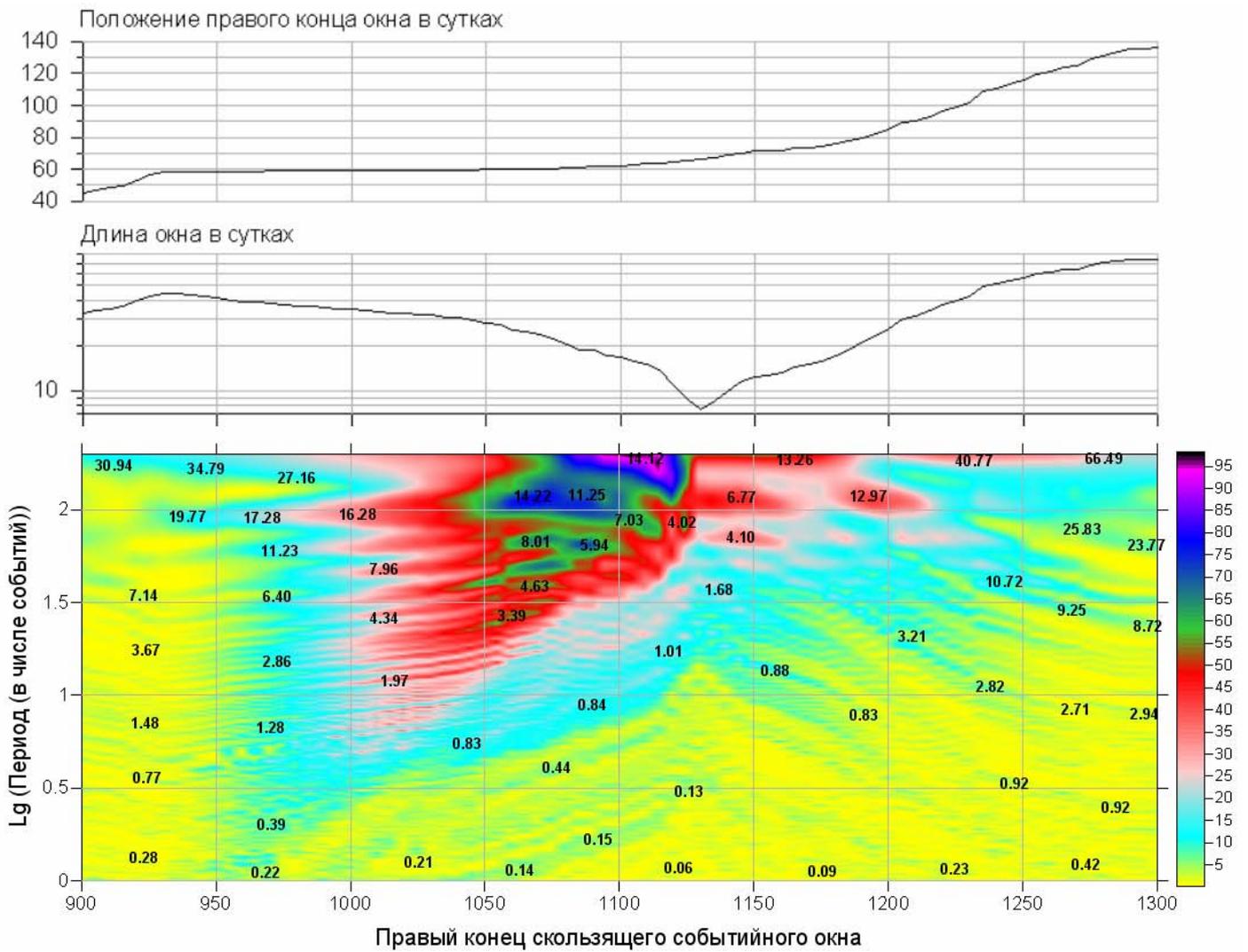


Рис.16. Черными цифровыми метками на частотно-временной диаграмме проставлены значения периодов в сутках.

Основная периодичность возникает в афтершоках на периодах от 10 до 18 суток. В отличие от диаграммы на рис.7 здесь не возникает устойчивого полусуточного колебания после главного толчка.

Наконец, исследуем финальную часть выборки для окон с 1300.

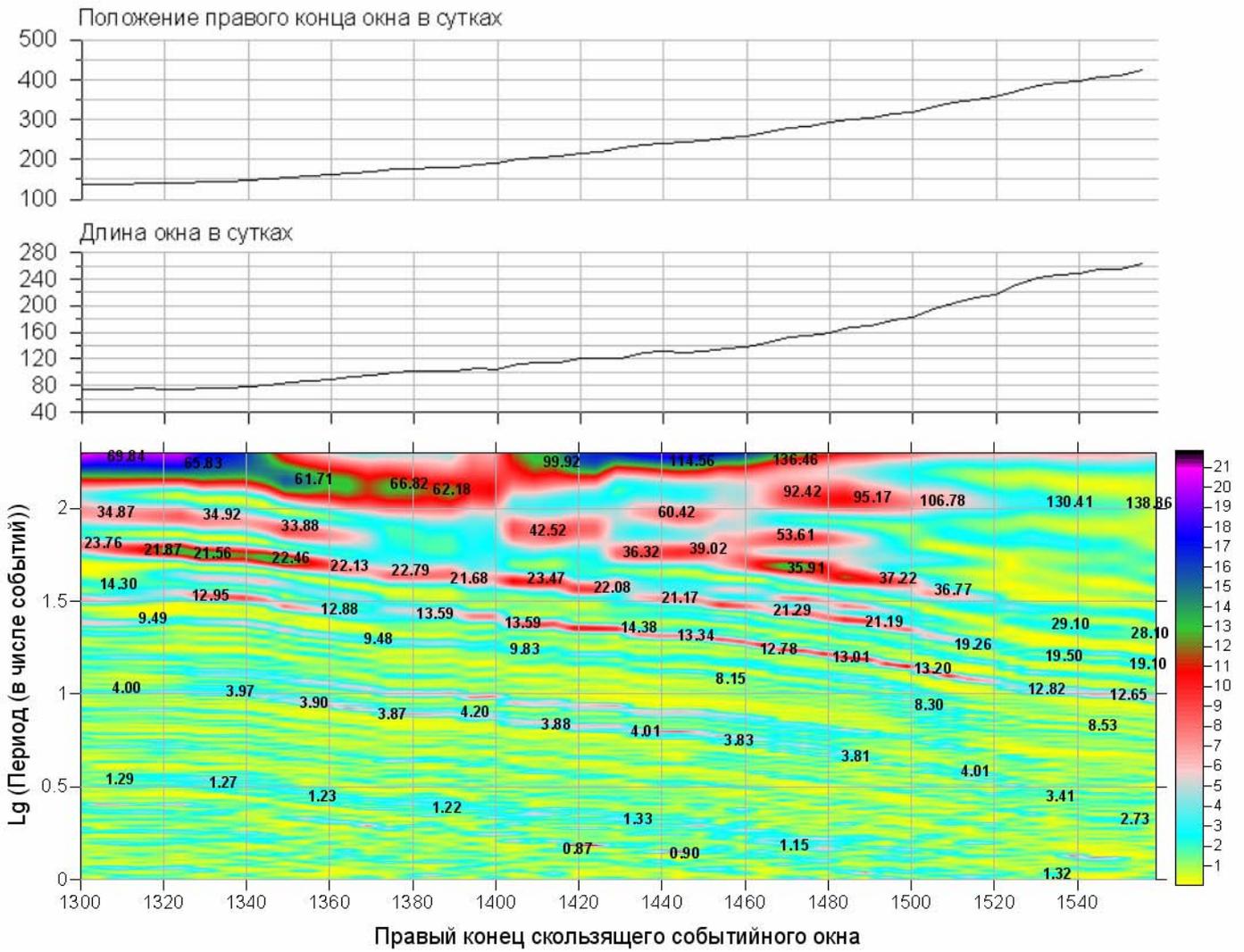


Рис.17. Черными цифровыми метками на частотно-временной диаграмме проставлены значения периодов в сутках.

Помимо низкочастотных периодических компонент с периодами 60-130 и 30-60 суток, на этой диаграмме выделяются устойчивые периодичности с периодами 19-24 суток и 12-13 суток.