

Выбор оптимальной конфигурации нейронной сети при моделировании годографов сейсмических волн

© О. А. Герасименко, 2012

Институт геофизики НАН Украины, Киев, Украина

Поступила 14 сентября 2011 г.

Представлено членом редколлегии А. В. Кендзерой

Виконано вибір оптимальної архітектури мережі штучних нейронів для створення моделі часу вступу фаз сейсмічних хвиль. На прикладі записів на сейсмостанції «Одеса» сигналів, що збуджуються джерелами навколо Чорноморського регіону, оцінено адекватність модельованого часу вступу фаз реально зареєстрованому, сталості роботи моделі в оперативному режимі залежно від щільності покриття простору джерелами, напрямку приходу до точки реєстрації сейсмічної радіації. Наведено графічні приклади оцінок.

The choice of the optimal neural net topology is performed for seismic phase onset simulation. Based on «Odessa» seismic station earthquake signal records excited by the sources of the Black Sea region the reliability of model outputs, operation stability related to source distribution density and seismic radiation flow direction are evaluated. The graphic examples are presented.

Предлагаемое построение годографов сейсмических волн представляет собой задачу аппроксимации функции времени прихода фазы сейсмической волны к точке наблюдения с помощью нейросетевой модели, построенной путем обучения сети искусственных нейронов с использованием множества примеров поведения этой функции, полученных в виде зарегистрированных на сейсмостанциях времен прихода сейсмических сигналов.

Аргументы функции-«примера поведения» времени распространения фазы сейсмической волны, возбужденной p -м землетрясением ($p = \overline{1, N}$) и зафиксированной на отдельной сейсмостанции, запишем в виде вектора

$$\mathbf{x}_i = (x_1, x_2, \dots, x_n, t). \quad (1)$$

Здесь t — целевая функция, определяемая временем пробега фазы сейсмической волны t_p или t_s , отображается в виде нелинейной проекции «вход-выход»:

$$t = f(x), \quad (2)$$

где $f(\cdot)$ — ограниченная, непостоянная, монотонно возрастающая непрерывная функция. Множество (именуемое обучающей выборкой), включающее наблюдаемые времена прихода фаз сейсмических волн и представляющее собой примеры поведения искомой функции, запишем как

$$T = \{(x, t)\}^N, \quad (3)$$

где N — мощность обучающего множества.

В процессе итерационного возбуждения нейронной сети входными сигналами (1) — членами обучающего множества — осуществляется обучение сети с помощью метода обратной передачи ошибки [Горбань, 1990] путем коррекции межнейронных связей нейронов сети за счет передачи невязки между выходным вектором и целевым значением. При этом структуре нейронной сети, которая аппроксимирует неизвестную функцию $y(x)$, предъявляется требование: функция $F(\cdot)$, описывающая отображение входного сигнала в выходной, должна быть достаточно близка к функции $y(x)$ в смысле евклидовой нормы на множестве всех входных векторов \mathbf{x} :

$$\|F(\mathbf{x}) - y(\mathbf{x})\| < \varepsilon, \quad (4)$$

где ε — некоторое малое положительное число [Хайкин, 2006].

От того, насколько представительна обучающая выборка N , правильно сконфигурирована нейронная сеть, достаточно ли свободных параметров — скрытых нейронов и сдвигов, зависит допустимая минимальная величина значения ε — ошибки аппроксимации некоторого непрерывного отображения [Тарков, 2006].

Теорема об универсальной аппроксимации, будучи математическим доказательством возможности аппроксимации любой непрерывной функции, не определяет, является ли один слой скрытых нейронов оптимальной структурой нейронной сети в задаче заданной слож-

ности, как и не регламентирует количество скрытых нейронов в слое. Предполагается, что аппроксимируемая функция непрерывна, достаточно хорошо представлена и для ее приближения можно использовать скрытый слой неограниченного размера [Cybenko, 1989].

В стохастическом режиме обучение сети осуществляется для каждого сигнала (входного вектора) путем оценки среднеквадратической ошибки, определяющей уровень обученности сети:

$$E = 1/2 \sum_{j=1}^m (t_j - y_j)^2, \quad (5)$$

где m — количество нейронов выходного слоя, t_j — целевое значение для j -го узла, y_j — текущий выход. Для минимизации ошибки сети (5) методом скорейшего спуска синаптические веса нейронов w_{ij} корректируются следующим образом:

$$w_{ij}(t+1) = w_{ij}(t) - \mu \frac{dE}{d w_{ij}}, \quad (6)$$

где w_{ij} — вес связи между i -м узлом слоя l и j -м узлом слоя $l-1$, t — номер итерации, $0 \leq \mu \leq 1$ — коэффициент скорости обучения.

Процесс обучения считается сошедшим в случае достижения среднеквадратической ошибкой (5) некоторого заданного уровня, не изменяющегося в течение достаточного количества итераций.

Оценка точности аппроксимации функции времени распространения фаз сейсмических волн была проведена в режиме «проб и ошибок» для одиночной сейсмологической станции. Нахождение оптимальной конфигурации нейронной сети проводилось с помощью сетей с варьирующим набором скрытых слоев и количеством нейронов в них, основываясь на регистрации 488 сейсмических сигналов землетрясений Черноморского региона, зарегистрированных с/с «Одесса». Из совокупности векторов-примеров параметризации исследуемых сейсмических событий формируется два множества — обучающее и экзаменационное. Первое используется для обучения нейронных сетей разных архитектур (с одним и двумя скрытыми слоями), в которых набор нейронов в разных слоях изменялся от 1 до 30.

Для нейронной сети с одним скрытым слоем поведение функции ошибки обучения изучалось для конфигурации нейронных сетей (НС) от НС=4:2:1 до НС=4:30:1 с интервалом изменения количества нейронов скрытого слоя, равным 2. Для варианта двух скрытых слоев

использовались сети с фиксированным количеством скрытых нейронов первого слоя, равным 2, 5, 10, 30, и переменными наборами скрытых нейронов второго слоя с интервалом изменения 5.

На рис. 1 представлен сводный график поведения функции погрешностей в процессе обучения сетей заданных топологий, где в качестве целевой функции использовались времена прихода P -волн.

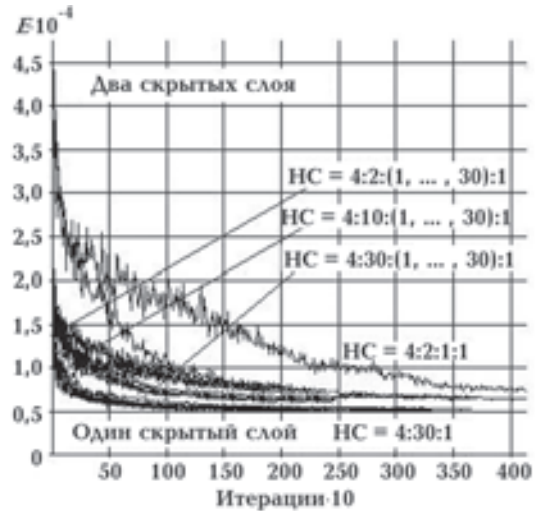


Рис. 1. Поведение функции потерь в зависимости от количества итераций при обучении нейронных моделей распространения P -волн.

Члены экзаменационного множества, выбранные случайным образом из обучающего множества, не принимали участия в обучении и использовались для проверки работы НС. Уровень обученности НС с определенным числом нейронов скрытого слоя оценивался по величине дисперсии ошибок идентификации сетью членов экзаменационного множества. Результаты экзамена для P -волн показаны на рис. 2.

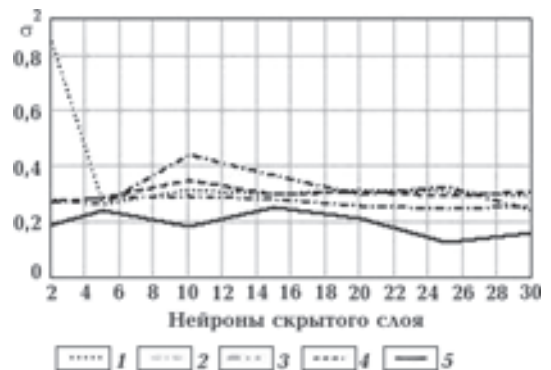


Рис. 2. Зависимость дисперсии величин ошибок экзамена НС от числа узлов скрытого слоя: 1 — НС=4:2:(1,...,30):1; 2 — НС=4:5:(1,...,30):1; 3 — НС=4:10:(1,...,30):1; 4 — НС=4:30:(1,...,30):1; 5 — НС=4:30:1 (1–4 — двухслойные НС, 5 — однослойная НС).

Анализ приведенных зависимостей показывает избыточность двухслойной конфигурации скрытых слоев НС для заданного уровня сложности поставленной задачи и мощности входного вектора. Минимум абсолютного значения среднеквадратической ошибки обеспечивает прямопоточная, полностью связанная, управляемая НС с одним скрытым слоем, состоящим из 25 нейронов.

Благодаря способности обученной НС к обобщению была построена универсальная нейросетевая модель поведения функции времени прихода P - и S -фаз сейсмической волны на с/с «Одесса», способная давать правильные оценки времен прихода фаз при возбуждении векторами со значениями составляющих, не принимавших участия в обучении.

Составляющие входного вектора (1) x_i : x_1 — глубина гипоцентра очага землетрясения; x_2 — магнитуда, x_3 — азимут прихода сейсмической волны, x_4 — эпицентральное расстояние и целевое значение, t_p , t_S — время вступления фазы сейсмической волны на сейсмической станции при проигрывании сценария изменения компонента x_4 и прочих постоянных дают возможность моделировать годограф для единичной сейсмологической станции с избранных направлений прихода волны, глубины гипоцентра и магнитуды события. Очевидно, что время пробега сейсмических волн в геологической среде, при всех прочих равных условиях, зависит от сложности строения последней, и степень этой сложности определяет ошибки результатов моделирования. Поэтому

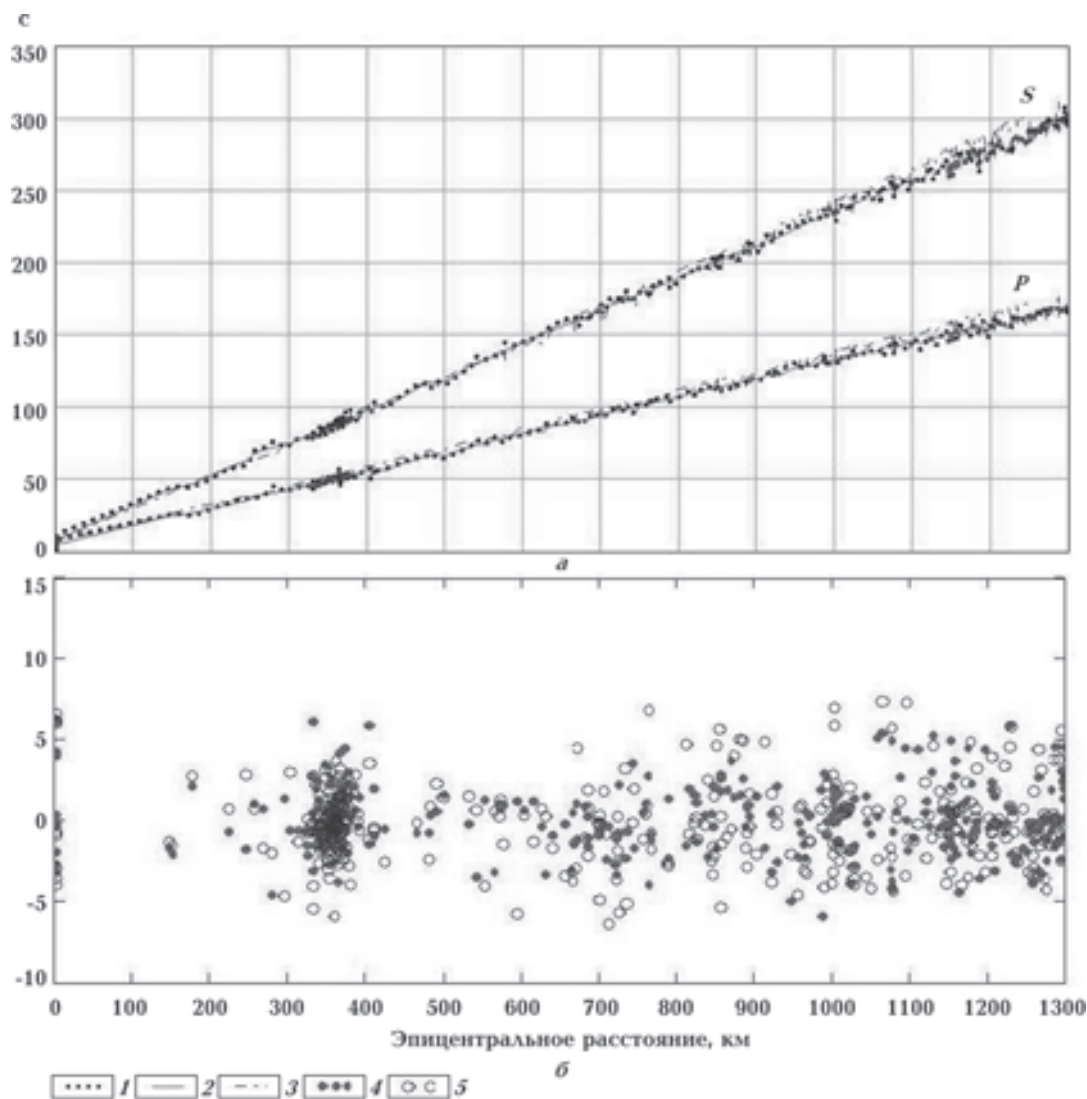


Рис. 3. Времена пробега P -, S -волн (а) по зарегистрированным на с/с «Одесса»; оценка ошибок нейросетевого моделирования (б): 1 — наблюдаемые, 2 — моделируемые, 3 — определенные по глобальным оценкам [Jeffreys, Bullen, 1940], 4, 5 — времена вступлений (4 — t_p , 5 — t_S).

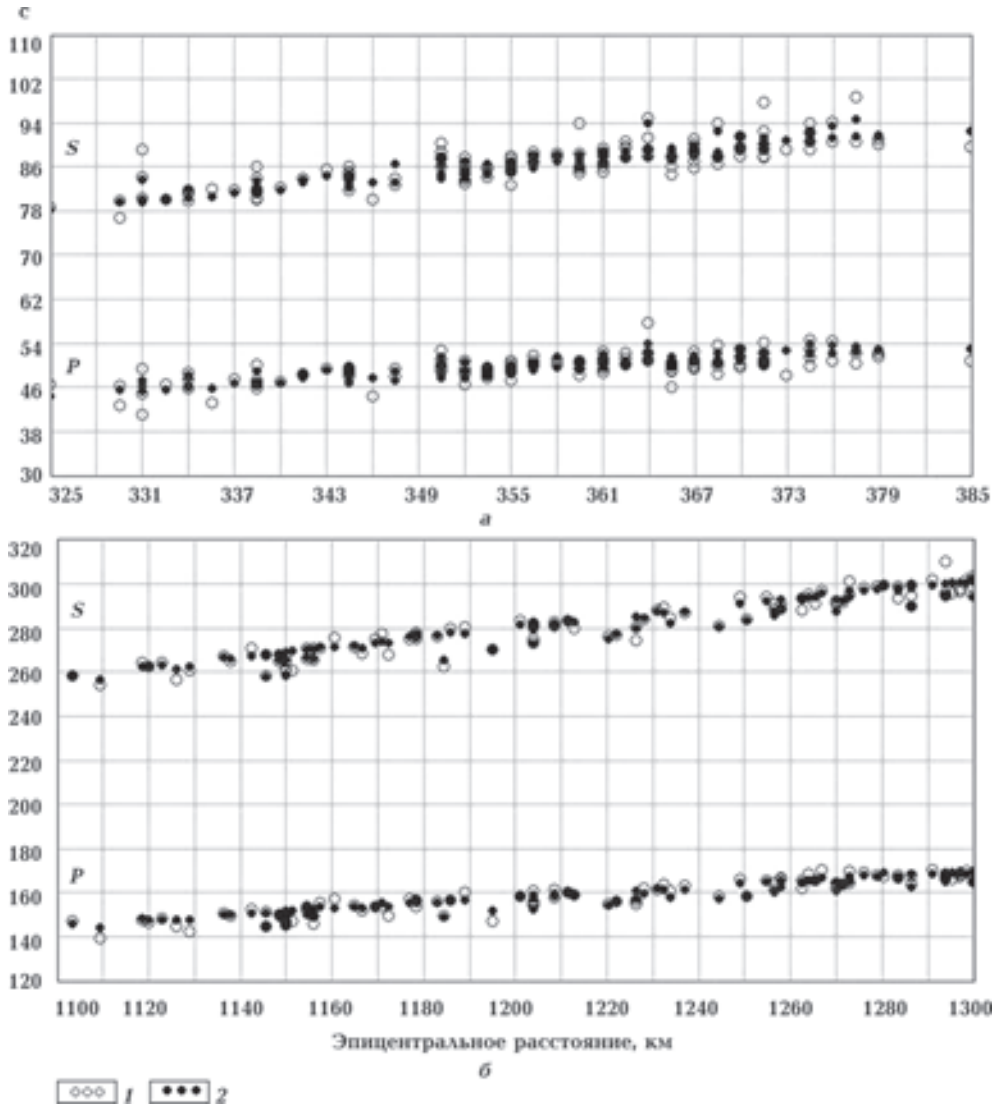


Рис. 4. Времена пробега P - и S -волн на с/с «Одесса» (1) и их оценки (2) по нейросетевой модели для эпицентральных расстояний 325—385 км (а) и 1000—1300 км (б).

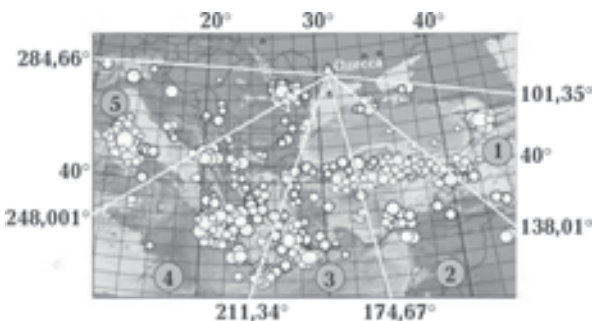


Рис. 5. Карта-схема моделируемого региона и секторы основных направлений сейсмической радиации.

точность результатов моделирования, в нашем случае — времен пробега волн, зарегистрированных на с/с «Одесса» от землетрясений в пределах $\Delta \approx 13^\circ$, может служить косвенным

показателем уровня сложности геологического строения среды.

Показанные на рис. 3 результаты нейросетевого моделирования приходов P - и S -волн хорошо согласуются с наблюдаемыми данными, но масштаб рисунка не дает представления о деталях, поэтому на рис. 4 приведены интервалы эпицентральных расстояний, наиболее «населенных» событиями.

Ранее рассмотренная в работах [Лазаренко, Герасименко, 2008; Лазаренко и др., 2011] азимутальная зависимость времен прихода фаз волн может быть вызвана как условиями среды распространения сейсмических волн, так и ошибками нейросетевого моделирования и/или глобального годографа. Для выяснения природы этих невязок была оценена погреш-

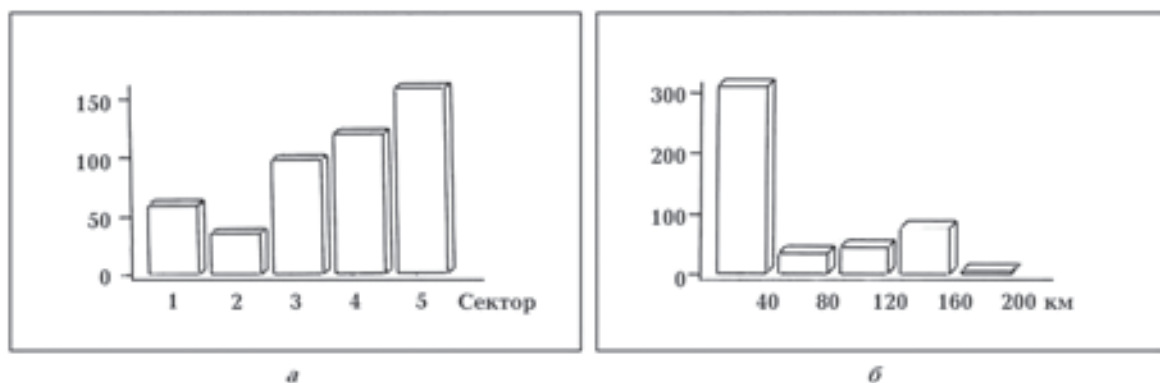


Рис. 6. Распределение очагов землетрясений, зарегистрированных с/с «Одесса»: *a* — количество землетрясений в секторах; *б* — то же по глубине очагов.

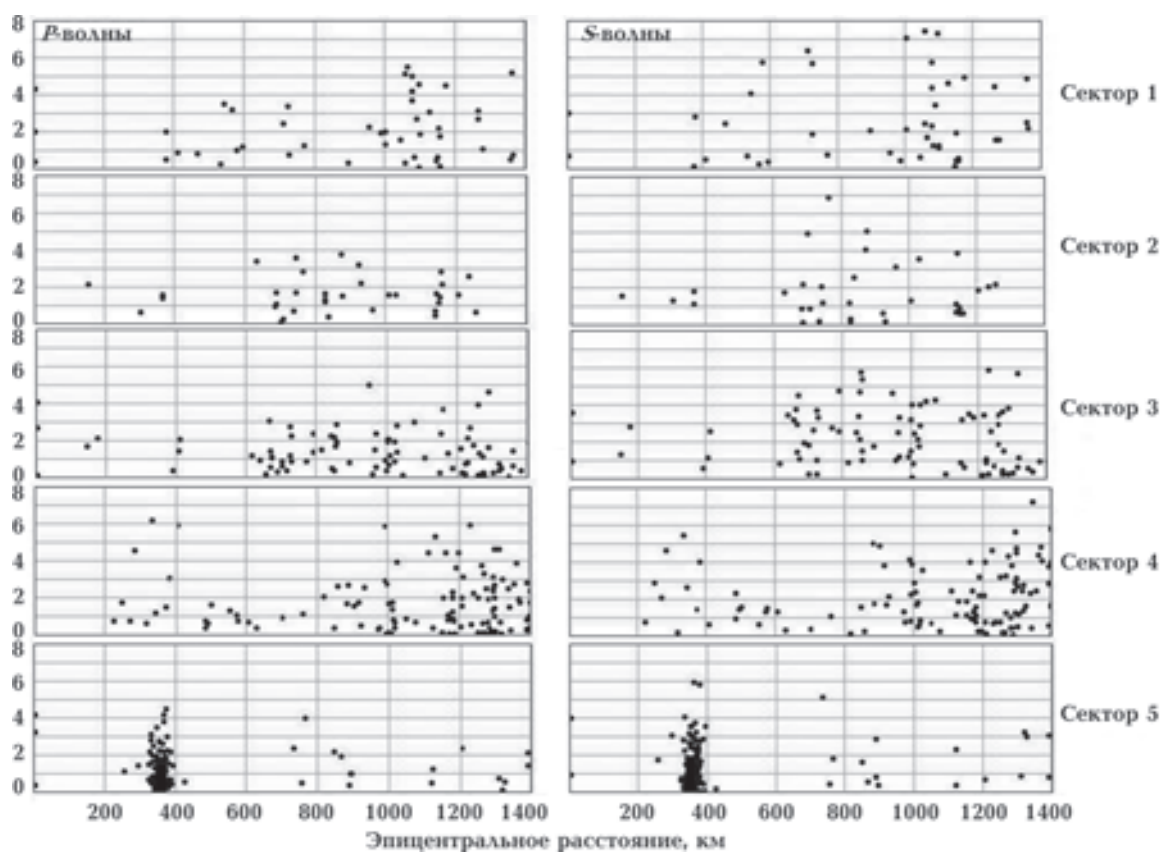


Рис. 7. Среднеквадратическая погрешность прогноза прихода сейсмических *P*- и *S*-волн на с/с «Одесса», разнесенная по секторам 1—5.

ность реальных измерений и моделируемого годографа на примере записей землетрясений, очаги которых находятся в разноплановых геолого-тектонических системах. Для этого зарегистрированные с/с «Одесса» события были разделены по значениям обратных азимутов прихода сейсмических волн на пять секторов в диапазоне 101—284° (рис. 5).

Характеристика сейсмической ситуации в регионе генерации волн показана на рис. 6.

Очевидна неравномерность информативной базы в разных секторах для выбранного эпицентрального расстояния. Как видно, основное количество очагов лежит в интервале глубин 0—30 км при весьма неоднородном «информативном» наполнении.

Единственная возможность проверить адекватность нейросетевого прогноза времен пробега фаз сейсмических волн реальной ситуации — сравнить результаты прогноза с на-

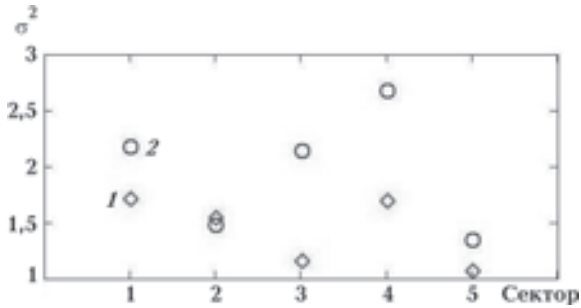


Рис. 8. Стандартное отклонение ошибки прогноза времен прихода *P*-волн (1) и *S*-волн (2) на с/с «Одесса» по секторам.

блюденными данными. На рис. 7 приведены по секторам (см. рис. 5) среднеквадратические отклонения наблюдаемых от смоделированных значений времен *P*- и *S*-волн, зарегистрированных на с/с «Одесса».

Несмотря на разницу в информативной обеспеченности секторов, визуальная оценка погрешности в них практически одинакова, что, учитывая современные представления о сложности архитектуры региона, свидетельствует о хорошей обобщающей способности нейросетевой модели и подтверждается распределением по секторам величин дисперсии ошибок прогноза, показанном на рис. 8.

Сравнение результатов оценок по вступлениям *P*- и *S*-волн указывает на значительно более высокий уровень зашумленности последних, что говорит о больших ошибках в оценках времен вступления *S*-волн, чем *P*. Кроме того, подобное представление ошибок прогноза при наличии достаточного объема базы наблюдаемых данных для обучения нейросетевой модели позволяет по анализу «ураганных» вы-

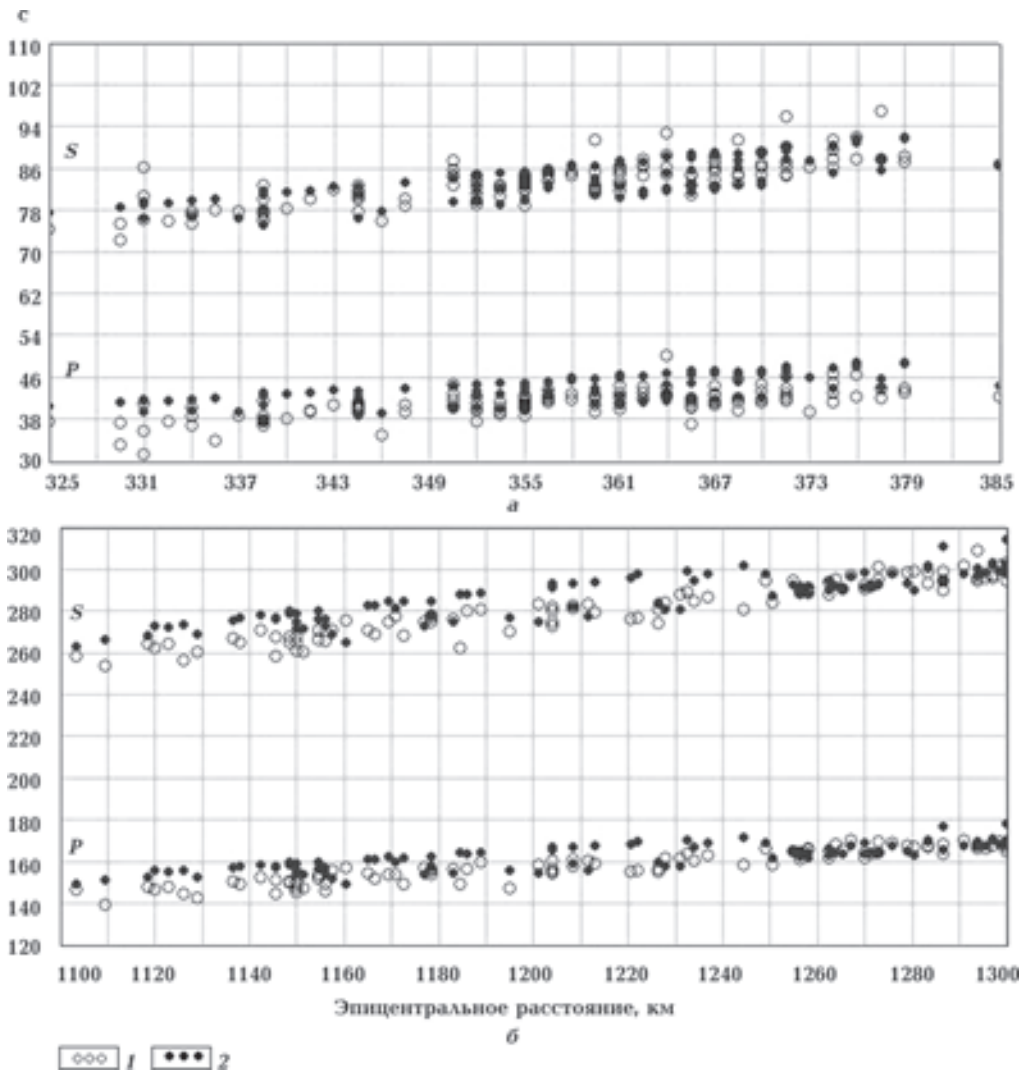


Рис. 9. Наблюдаемые времена вступления *P*- и *S*-волн на с/с «Одесса» (1) и их оценки (2) по глобальным моделям для эпицентральных расстояний 325—385 км (а) и 1000—1300 км (б).

бросов редактировать результаты первичной обработки сейсмологических наблюдений.

Невязки вступления фаз по секторам между наблюдаемыми и определенными по глобальным оценкам [Лазаренко, Герасименко, 2010], приведены на рис. 9.

Как видим, эти результаты проигрывают в сравнении с полученными для нейросетевых модельных оценок. Не обсуждая причин появления такого результата, для современного уровня точности сейсмических наблюдений,

по крайней мере, на порядок большего, чем у выполнявшихся в прошлом столетии, использование нейросетевого моделирования позволит создавать адаптированные к конкретной сейсмологической станции модели вступления фаз сейсмических волн, способные повысить точность аппроксимации реальных годографов по мере накопления информации, а также легко инкорпорироваться различными алгоритмами обработки сейсмических данных.

Список литературы

- Горбань А. Н. Обучение нейронных сетей. — Москва: СП ПараГраф, 1990. — 160 с.
- Лазаренко М. А., Герасименко О. А. Моделювання сейсмічної анізотропії кори і верхньої мантії акваторії Чорного моря за допомогою мереж штучних нейронів // Вісник Київ. нац. ун-ту. — 2008. — Вип. 45. — С. 18—23.
- Лазаренко М. А., Герасименко О. А. Нейросетевое моделирование локальных годографов сейсмических волн // Геофиз. журн. — 2010. — **32**, № 5. — С. 126—141.
- Лазаренко М. А., Герасименко О. А., Остапчук Н. Н. Особенности нейросетевого моделирования локальных годографов сейсмических волн // Геофиз. журн. — 2011. — **39**, № 6. — С. 112—114.
- Тарков М. С. Нейрокомпьютерные системы. — Москва: Бинум, 2006. — 140 с.
- Хайкин С. Нейронные сети: Полный курс. — 2-е изд., испр. — Москва: Издательский дом Вильямс, 2006. — 1103 с.
- Cybenko G. Approximation by superpositions of sigmoidal functions // Math. Control Signal Systems. — 1989. — **2**. — P. 303—314.
- Jeffreys H., Bullen K. E. Seismological Tables. — London: British Association Seismological Committee, 1940. — 145 p.