

Нейросетевое моделирование годографов сейсмических волн

© М. А. Лазаренко, О. А. Герасименко, 2010

Институт геофизики НАН Украины, Киев, Украина

Поступила 1 февраля 2010 г.

Представлено членом редколлегии А. В. Кендзерой

Запропоновано використання мереж штучних нейронів для створення моделі поля часу пробігу *P*- та *S*-фаз сейсмічних хвиль. Розглянуто принципи функціонування багатозарової, повнозв'язаної, прямострумінної, керованої нейронної мережі, в якій використано для навчання метод зворотної передачі похибки, а також підходи до вибору оптимальної архітектури мережі, оцінок похибок режиму екстраполяції. Показано побудову тривимірних годографів для різних сценаріїв сейсмічного процесу на прикладах сейсмічних записів українських сейсмостанцій, їх використання для моделювання годографів та інверсії останніх. Наведено приклади інверсії годографа у межах моделі Герглотца — Віхерта для одиничних сейсмостанцій, а також узагальнення для випадку довільних координат джерела і точки реєстрації сигналу у Чорноморському регіоні.

The implementation of artificial neural networks for travel-time model of *P*- and *S*-phases of seismic waves arrangement is proposed. The principles of multilayer, fullconnected, feedforward, controlled, and backpropagated neuron network functioning and approach to net topology choice and extrapolation error assessment are considered. The 3D travel-time relationships for various scenario of seismic process using the travel-time inversion based on Ukrainian seismic stations records are considered. The examples of Herglotz — Wiehert inversion for single stations as well as for arbitrary source and station coordinates in circum-Black Sea region are presented.

Приближенное отображение некоего физического феномена либо функции, отражающей его поведение в пространстве и/или времени, обеспечивают модели, осуществляющие это приближение. Полнота наших знаний о природе феномена, развитость математического аппарата, мощность множества примеров его состояния либо поведения диктуют стратегию решения задачи и выбор пути конструирования модели: аналитический либо эмпирический.

Сейсмология предоставляет нам огромную массу наблюдений времен пробега сейсмических волн на фоне ограниченных представлений о физике частного волнового процесса в еще более загадочной среде. Для такого рода задач не построение аналитических решений, а некое отображение входного пространства примеров поведения изучаемой функции в выходное пространство решений может быть осу-

ществлено на сетях искусственных нейронов. Последние формируют модели в виде матриц (таблиц) параметров, обобщающих свойства предложенных им примеров поведения, перенося эти знания на объекты, характер поведения которых неизвестен.

Нейронные сети являются воплощением коннективистской [Caplinskas, 1998] парадигмы искусственного интеллекта [Стюарт, Питер 2007; Мак-Каллок, Питтс, 1956], отражающим представление о способности моделирования процессов (в том числе и мыслительных), происходящих в животном мозге, с помощью систем взаимосвязанных элементарных процессоров. Эти процессоры, имитирующие нервные клетки-нейроны, функционируют в параллельном режиме и способны перестраивать силу (вес) этих связей, управляя таким образом интенсивностью сигналов, циркулирующих в сети.

Нейросеть: схема и функционирование.

Сеть искусственных нейронов — мощный инструмент статистического анализа, позволяющий построить поведенческую модель, используя множество численных либо лингвистических переменных, являющихся примерами этого поведения. Нейронная сеть, “безграмотная” вначале, превращается после обучения в некоторую модель зависимостей между этими “описательными” переменными и поведением, подлежащим объяснению.

Из всего многообразия известных типов сетей искусственных нейронов для наших приложений выберем (как наиболее разработанную) многослойную, полносвязанную, прямопоточную, управляемую сеть [Уоссермен, 1992], использующую для обучения метод обратной передачи ошибки [Chauvin, Rumelhart, 1995]. Схема такой сети представлена на рис. 1, причем для простоты показан один скрытый слой.

Обучение нейронной сети состоит в предоставлении ей последовательности примеров описания изучаемого феномена и целевого значения в виде требуемого решения либо оценки состояния или принадлежности для каждого примера. Эта итерационная процедура позволяет на каждом шаге итерации вычислить поправку к весам (W, wb) синаптических (межузловых) связей сети:

$$\Delta(w) = \mu \delta x, \tag{1}$$

минимизирующую среднеквадратическую ошибку или невязку

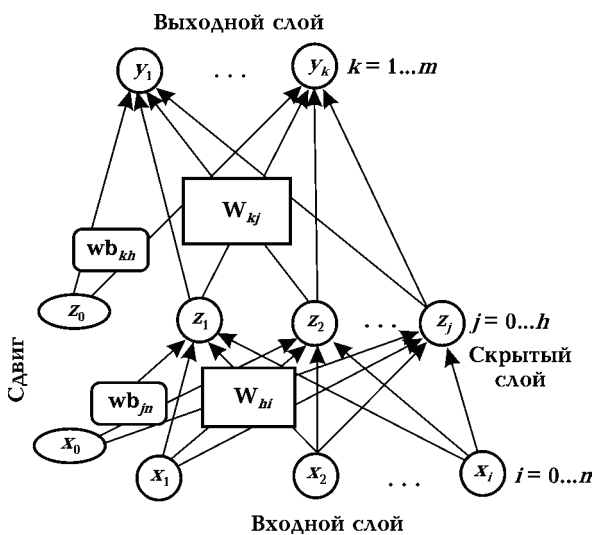


Рис. 1. Схема многослойной сети искусственных нейронов.

$$E = \frac{1}{2} (t - y_{\text{вых}})^2 \tag{2}$$

между выходом сети $y_{\text{вых}}$, возбужденным вектором x , и ожидаемым выходом, т. е. его целевым значением t . Здесь δ — градиент многомерной поверхности ошибок (2) является вектором, определяемым частной производной от функции ошибок по весам, а $1 > \mu \geq 0$ определяет скорость сходимости итераций.

При наличии учебного материала достаточной представительности, нейронная сеть способна отыскать и запомнить в виде таблиц соответствий “вход — выход” закономерности, определяющие отнесение этих примеров к своим целевым значениям (решениям). Эта “память” используется сетью при функционировании в оперативном режиме обобщения для нахождения неизвестных решений у предъявляемых ей примеров, не принимавших участия в обучении.

Представляя в виде векторов: 1) входные сигналы x_i , 2) сдвиговые веса wb_k , 3) позиционируемые по отношению к слоям межузловых связей выходные из предшествующих слоев нейронов сигналы (входные в последующие)

$$y_k = \sum_{k \in A_i} w_{ik} y_k, \tag{3}$$

где $A_i = \{ j : \exists w_{ij} \}$ — множество узлов, предшествующих узлу i , 4) $\delta_k = \partial E / \partial y_k$ — сигналы ошибок на узлах слоя k , а межузловые веса W в виде матриц, и определив последовательность слоев нейронов k как $k = 0, 1, 2, \dots, K$, процесс обучения нейронной сети начинается с возбуждения входного ($k = 0$) слоя нейронов сигналом-вектором, являющимся элементом обучающей выборки. Так как первый слой нейронов сети не несет вычислительной нагрузки, а выполняет только дистрибутивную функцию, то

$$y_0 = x. \tag{4}$$

Продвигаясь в направлении выходного слоя K , выходы каждого k -го слоя запишутся в виде

$$y_k = f_k (W_k y_{k-1} + wb_k), \tag{5}$$

а векторы ошибки на выходе сети как

$$\delta_K = t - y_K, \tag{6}$$

где t — вектор цели (желаемый отклик сети) с размерностью, равной мощности выходного слоя, $f_k(\bullet)$ — передаточная функция k -го слоя. Таким образом, первая (прямая) фаза обучения, начавшись с возбуждения скрыто-

го слоя сигналом из обучающего множества, заканчивается вычислением ошибки для каждого нейрона выходного слоя.

Получить решение (6) для нейрона, находящегося в выходном слое, для которого желаемая реакция сети задана в виде целевого значения, задача тривиальная. Для нейронов скрытого (скрытых) слоя это значение неизвестно, поэтому ошибка должна вычисляться с использованием ошибки нейронов с ним непосредственно связанных, что осуществляется алгоритмом обратной передачи ошибки [Chauvin, Rumelhart, 1995].

Передача ошибки в обратном направлении, т. е. для слоев $k = K - 1, K - 2, \dots, 1$ начинается с вычисления сигнала ошибки

$$\delta_k = \mathbf{W}_{k+1}^T \delta_{k+1} f'_k(\mathbf{y}_k), \quad (7)$$

где T — оператор транспозиции, $f'_k(\bullet)$ — производная передаточной функции нейронов k -го слоя.

Фазу обратной передачи ошибки итерационного процесса обучения нейронной сети завершает операция вычисления коррекции Δ матрицы весов \mathbf{W} и вектора сдвигов \mathbf{wb} :

$$\Delta \mathbf{W}_k = \delta_k \mathbf{y}_{k-1}^T, \quad \Delta \mathbf{wb}_k = \delta_k, \quad (8)$$

причем, об онлайн-овом или стохастическом режиме обучения говорят в случае осуществления коррекции (8) после прохождения через сеть каждого вектора обучающей выборки, а о пакетном, когда в обратной передаче ошибки участвует математическое ожидание ошибки всей обучающей выборки.

Скорость сходимости процесса обучения нейронной сети контролируется численной обусловленностью матриц весов, измеряемой величиной обусловленности $K_{об}$, которая в приложении к нейронным сетям определяется отношением \max/\min собственных значений матрицы Гессе (т. е. матрицы вторых производных функции ошибок по синаптическим весам). $K_{об} = 1$ определяет наилучшую обусловленность, и чем выше это значение, тем обусловленность хуже.

В случае $K_{об}$ близком к 1 в процессе обратной передачи ошибки отрицательное направление градиента практически совпадает с направлением на глобальный минимум функции ошибок, и линии проекции сечения функции потерь (6) будут иметь форму, приближающуюся к кругу. Для величин $K_{об}$ отличных от 1, т. е. в случаях плохой обусловлен-

ности, сечения имеют весьма сложную проекцию, а отрицательный градиент — направление, значительно отклоняющееся от направления к минимуму. Таким образом, в плохо обусловленных задачах поверхность функции потерь в одних координатах квазиплоская, в других — крутопадающая, а большие коэффициенты скорости обучения μ , обеспечивающие сходимость итерационного процесса в первом случае, могут вызвать его расхождение во втором.

В нейронных сетях, использующих нелинейные передаточные функции, функция ошибок содержит много точек, где гессиан не определен (например, седловидные точки), поэтому для сохранения сходимости итерационного процесса необходимо использовать небольшие значения коэффициента скорости обучения, что, правда, увеличивает время обучения сети, в некоторых случаях значительно.

Известно, что нейронные сети, обучаемые с постоянными значениями коэффициента μ , сходятся медленно, часто "застревают" в локальных минимумах [Горбань, 1990]. Использование же переменного значения коэффициента скорости обучения, изменяющегося обратно пропорционально количеству итераций, дает значительный позитивный эффект.

Нами использовались два пути уменьшения μ в процессе обучения. Первый, в алгоритме обучения с экзаменом известным под названием "перекрестная проверка", в котором задается конечное количество итераций цикла обучения и экспоненциальное убывание скорости коэффициента обучения

$$\mu(t) = \mu_i (\mu_t / \mu_i)^{t/t_{\max}}, \quad (9)$$

где μ_i — начальное, μ_t — конечное значение коэффициента скорости обучения, t — текущее и t_{\max} — конечное число итераций.

Использование закона убывания коэффициента скорости обучения (9) позволяет избежать "застревания" процесса обучения в локальных минимумах на начальных этапах обучения, так как этот подход можно рассматривать как внесение в систему случайного шума, уровень которого экспоненциально уменьшается. Известно, что введение аддитивного шума в исходные данные уменьшает чувствительность системы нейронных сетей к ошибкам задания стартового множества синаптических весов и во многих случаях приводит к меньшим (по сравнению с другими подходами) среднеквадратическим ошибкам [Хайкин, 2008].

Второй путь состоит в интерактивном взаимодействии с алгоритмом обучения, позволяющем с задаваемым шагом по количеству итераций манипулировать величинами коэффициента скорости обучения, коэффициентом инерции и модой обучения (пакетной или стохастической), руководствуясь формой кривой функции ошибок, величиной среднеквадратической ошибки, темпом уменьшения последней.

Используемый в методе обратной передачи ошибки метод скорейшего спуска весьма чувствителен к диапазону значений участвующих величин. Поэтому нормировка входных данных может не только улучшить, но часто и обеспечить сходимость процесса обучения. Обученную нейронную сеть можно рассматривать как классификатор, относящий предлагаемые примеры к областям многомерного пространства, ограниченными гиперплоскостями. Каждый узел скрытого слоя определяет свою гиперплоскость. Каждая гиперплоскость есть геометрическое место точек, в которых вектор-вход в этот узел (см. рис. 1) равен нулю. Вектор весов, соответствующий этому входному вектору, определяет положение гиперплоскости. Величина сдвига определяет смещение гиперплоскости от начала координат. Если в начале обучения сети величины смещений заданы в виде небольших равномерно распределенных случайных чисел, то все гиперплоскости располагаются вблизи начала координат. Следовательно, если входные данные занимают области пространства вдали от нуля, то облако их значений не может быть пересеченным гиперплоскостями и процесс обучения "застрянет" в локальном минимуме или будет неоправданно долгим.

Очевидно, что для успешного обучения входные данные должны быть нормированными. Известно [Stornetta, Huberman, 1987], что симметричный диапазон изменения стартовых весов и сигналов, возбуждающих сеть, дает приращение скорости обучения на 30—40 %, поэтому в качестве передаточной функции нами использовался гиперболический тангенс, областью существования которого есть интервал $[-1, 1]$, а переменные x_i входных данных нормируются как

$$\hat{x}_i = 2(x_i - m) / r \quad (10)$$

к величине полу суммы крайних значений $m = 0$ и размаху $r = 2$, что обеспечивается значениями $m = (\max x_i + \min x_i) / 2$ и $r = \max x_i - \min x_i$.

Широко используемая в практике обработки сейсмических сигналов зависимость времени прихода к точке наблюдения фазы сигнала от расстояния до эпицентра, или годограф, обычно определяется в виде таблиц значений функции одной переменной для различных глубин гипоцентра. Аппроксимация этой зависимости, выполненная с помощью нейросетевого моделирования, позволяет решить задачу не только в виде линейной постановки множественной регрессии, но и нелинейной [Лазаренко, Герасименко, 2007]. Предлагается обобщение стандартного одномерного распределения путем увеличения числа аргументов функции времени распространения фаз сейсмических волн, представляя вектор, описывающий ее поведение для одной сейсмической станции, в виде

$$\mathbf{x}_i = (x_1, x_2, \dots, x_n, t_p), \quad n = \overline{1, N}. \quad (11)$$

Здесь N — мощность вектора параметров, компоненты вектора, x_1 — глубина гипоцентра, x_2 — магнитуда, x_3 — обратный азимут, x_4 — эпицентральное расстояние, t_p — время пробега фазы волны, возбужденной p -м событием.

Экзамен. Несмотря на априорную эффективность аппроксимации нейронными сетями достаточно сложных функций [Cybenko, 1989], в каждом случае решения конкретной задачи желательно получить оценки точности аппроксимации как для подбора оптимальной архитектуры сети, так и для оценок адекватности реакции создаваемой модели реальным ситуациям при ее использовании в оперативном режиме, т. е. в режиме прямого потока без коррекции весов. Теоремы о полноте [Hecht-Nielson, 1987], доказывая возможность многослойной нейронной сети с заданной точностью аппроксимировать любую непрерывную функцию нескольких переменных, не определяют количества нейронов скрытого слоя, позволяющего решить конкретную задачу.

Стандартный прием — экзамен сети с помощью подмножества обучающей выборки, состоящего из элементов, не принимавших участия в данном сеансе обучения и отбирающихся из последней случайным образом. По количеству неверно классифицированных объектов этого экзаменационного подмножества или идентификаций, превышающих некий заданный порог, можно делать выводы об уровне обученности и ожидаемой эффективности работы нейросетевой модели. Для этой цели

из множества обучающей выборки \mathbf{B} случайным образом выбирается подмножество \mathbf{T} , в дальнейшем именуемое экзаменационным, $\mathbf{T} \in \mathbf{B}$, причем, $|\mathbf{T}| = \sim 10^{-1} |\mathbf{B}|$. Далее нейронные сети различной архитектуры обучаются на множестве $\mathbf{B} - \mathbf{T}$ и "экзаменуются" на подмножестве \mathbf{T} , а ошибка экзамена каждой архитектуры позволяет оценить пригодность последней для решения задачи, использующей выборку \mathbf{B} .

На рис. 2 приведен результат использования этого алгоритма экзамена для множества параметров землетрясений, зарегистрированных тремя сейсмическими станциями. Тестировались нейронные сети с одним и двумя скрытыми слоями и количеством скрытых узлов от 2 до 30. Приведенные графики позволяют оценить относительную пригодность структуры скрытых слоев для решения нашей задачи.

Проверке способности нейросетевой модели в рассматриваемой постановке адекватно воссоздавать поведение во временном пространстве функции вступления фаз сейсмических волн мешает отсутствие эталонных ситуаций, которые можно было бы использовать для прямой проверки результатов, воспроизводимых обученной нейронной сетью в оперативном режиме. Один из типов косвенных экзаменов, использовавшихся нами, состоит в так называемом перекрестном или обучении "без одного", при котором для конкретной ситуации методом проб и ошибок определяется архитектура сети, оптимальное количество итераций в одном цикле обучения, стартовые значения и темп убывания коэф-

фициентов скорости обучения и инерции [Ногник et al., 1989].

Для количества циклов обучения, равного $|\mathbf{B}|$ — мощности множества обучающей выборки, выполняется цикл обучение сети с обучающей выборкой мощностью $|\mathbf{B}| - 1$, причем, член выборки, не принимавший участия в данном цикле обучения, предлагается обученной сети в качестве экзаменатора, т. е. для оценки отличия выхода модели, возбужденной вектором экзаменатора, от его целевого значения. Ошибкой считалось отличие $> 5\%$ от целевого значения. Количество ошибок ϵ таких оценок, отнесенное к мощности обучающей выборки, характеризует разрешающую способность модели. Так, для трех сейсмостанций: "Сквира", "Одесса" и "Полтава" $|\mathbf{B}_i| = 330, 344$ и 324 , а $\epsilon_i = 9, 8$ и 10 соответственно при $i = 1, 3$, что известным способом дает возможность оценить вероятность ошибки для каждой сейсмостанции.

Представление годографа как геометрического места концов многомерных векторов, проведенных из одной точки, вызывает естественный вопрос о точности аппроксимации реального годографа — нелинейной функции очень большого числа аргументов функцией, или суперпозицией функций, заведомо меньшего их числа. Корректность такой аппроксимации доказана как аналитически [Колмогоров, 1956], так и эмпирически [Jeffreys, Bullen, 1940]. Адекватность отображения наблюдаемых данных реальным годографом может быть оценена путем сравнения с неким эталоном, обобщающим в решении относительно времени распространения упругой волны

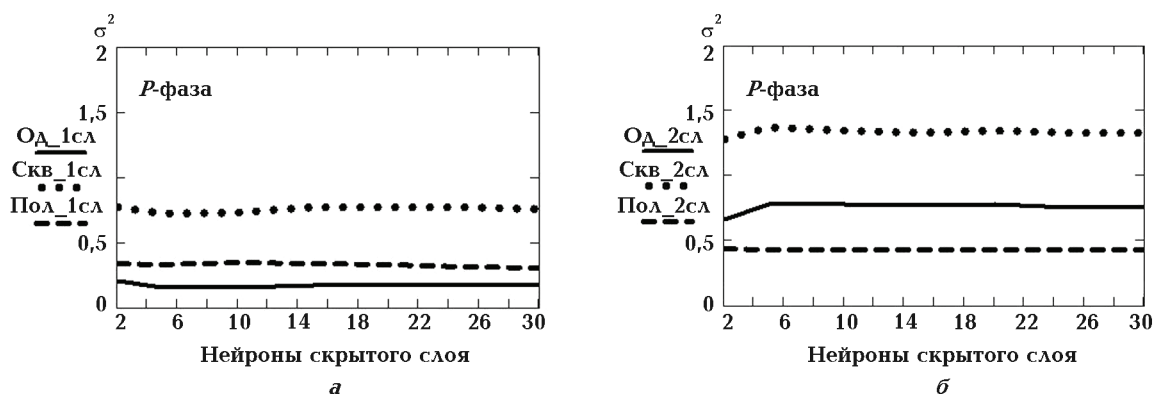


Рис. 2. Зависимость дисперсии ошибок экзамена нейронной сети при моделировании годографов P -фазы сейсмических волн, зарегистрированных с/ст "Одесса", "Сквира", "Полтава", от числа узлов q скрытого слоя для вариантов: a — нейронная сеть $NN : 4 : q : 1$ с одним скрытым слоем, $q = 2, 4, 6, \dots, 30$; b — нейронная сеть $NN : 4 : 2 : q : 1$ с двумя скрытыми слоями, $q = 2, 5, 10, 15, \dots, 30$.

набор параметров, достаточный для описания среды распространения сейсмического сигнала и его источника.

В настоящее время в сейсмологии широко известны фундаментальные осреднения времен пробега сейсмических волн для сферической модели Земли с симметричными относительно центра свойствами в виде таблиц глобальных годографов [Jeffreys, Bullen, 1940; Kennet, 1991]. Эти таблицы использовались для опосредованного экзамена качества обучения нейронной сети, трансформируя их в обучающую выборку нейросетевой модели, которая организовывалась после обучения сети (с архитектурой NN 2 : 25 : 7 : 1) на множестве двумерных векторов $\mathbf{t}(h, r)$, скомпилированных из стандартных таблиц. Далее из обучающего множества в каждом сеансе обучения исключались векторы $\mathbf{t}(h, r)$ для определенной глубины источника h . С помощью обученных на таких усеченных множествах нейросетевых моделей генерировались годографы для этих изъятых из обучающего множества глубин источника.

Сравнивая смоделированные эти отсутствующие в обучающей выборке годографы с табличными оригиналами, можно оценить способность модели к обобщению в решении задач этого типа в виде невязки

$$\Delta t(L)|_h = t_{\text{ДЖ}}(L|h) - t_{\text{NN}}(L), \quad (12)$$

где L — эпицентральное расстояние, h — глубина гипоцентра, $t_{\text{ДЖ}}$ — годограф Джеффриса — Буллена, t_{NN} — годограф, сгенерированный моделью, обученной на выборке с изъятыми временами прихода фазы при глубине гипоцентра h .

Последовательное изъятие времен прихода P - и S -фаз для ряда глубин гипоцентра из использовавшихся в качестве обучающих табличных данных Джеффриса — Буллена дает возможность оценить относительные ошибки в моделировании времен прихода сейсмических волн, возбуждаемых на различных эпицентральных расстояниях и глубинах.

На рис. 3 для различных глубин гипоцентра приведены невязки (12) смоделированных годографов с табличными глобальными. Понятно, эти невязки могут служить мерой адекватности NN-модели модели Джеффриса — Буллена. Особенно отчетливо видны погрешности при работе нейросети в режиме экстраполяции в модели рис. 3, а для $h = 150$ км.

Можно отметить два очевидных источника ошибок, искажающих эти разностные кривые.

Высокочастотные помехи обуславливаются ошибками приведения неравномерного по расстоянию шага табличных годографов Джеффриса — Буллена к равномерному (10 км), выполненного с помощью интерполяции кубическими сплайнами. Уход от нуля разностных кривых связан с режимом экстраполяции, в который попадает модель при воспроизводстве (опущенных при обучении) крайних годографов Джеффриса — Буллена из ряда по глубинам источника.

Увеличение мощности обучающей выборки за счет табличных значений времен пробега от источников на глубинах ≤ 250 км позволяет перевести функционирование нейросети при моделировании годографа $h = 150$ км из режима экстраполяции в режим интерполяции, что и показано на рис. 3, б. Следует отметить, что нейросетевое моделирование глобальных годографов, кроме оценочной функции в вышерассмотренном случае, может использоваться в прикладных задачах для генерации годографов от источника произвольной глубины, не прибегая к неудобной практике линейной интерполяции дискретных табличных значений глубин.

Нейросетевая модель наиболее эффективно функционирует в режиме интерполяции, т. е. обладает наилучшей обобщающей способностью при условии, когда у объектов, предложенных модели к распознаванию или классификации, значения параметров не выходят за пределы диапазона значений, которыми они ограничивались в обучающей выборке. Для нашей задачи естественной является ситуация неравномерного покрытия моделируемой территории источниками и пунктами наблюдения и большими пропусками в ряду магнитуд. Поэтому не исключена работа модели в режиме хотя бы "частичной" экстраполяции, когда значения некоторых составляющих выходят за границы упоминавшегося диапазона. В таком случае наилучшие результаты получаются, когда передаточная функция выходного слоя линейна: она не изменяет уровня активации выходного слоя, последний не насыщается и такая сеть способна экстраполировать.

Обычно в границах изучаемого региона длины годографов превосходят расстояния до эпицентров событий, входящих в обучающее множество, и годограф на некотором участке строится моделью в режиме экстраполяции. Для оценки ожидаемых ошибок проведено сравнение "базовых" годографов по моделям,

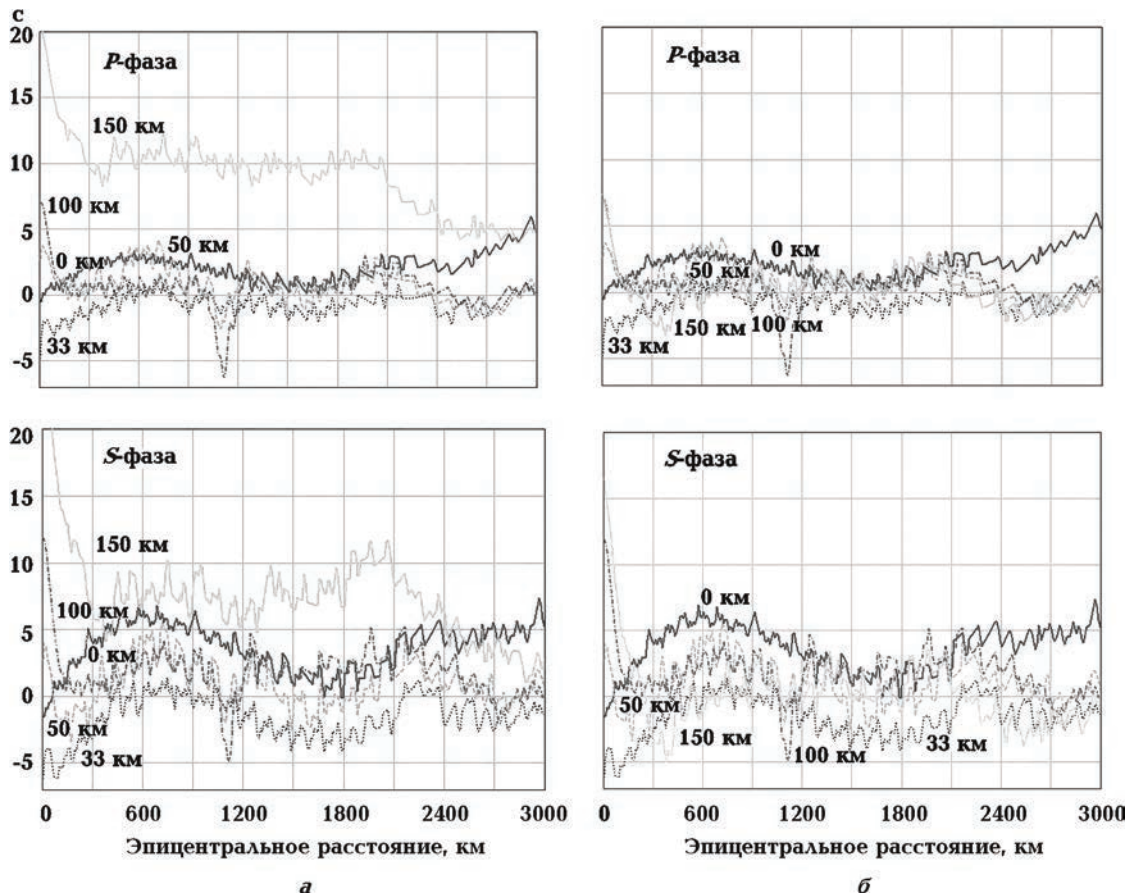


Рис. 3. Ошибка воспроизводства нейросетевой моделью глобальных годографов для значений глубин гипоцентров, отсутствующих в обучающей выборке: *а* — интервал глубин источника 0—150 км, *б* — 0—250 км.

обученным на множестве событий, покрывающих весь диапазон существования функции, и той же длины "текущих" годографов, но смоделированных на сетях, обученных на землетрясениях, покрывающих лишь часть эпицентрального расстояния. Рис. 4 позволяет оценить величину таких ошибок продолжения моделью функции годографа в область отсутствия информации о поведении модели на примере годографов *P*- и *S*-волн для двух сейсмостанций.

Построение годографов. Плоскостью приложения наших теоретических экскурсов был выбран черноморский регион с координатами $\varphi = 47 \div 35^\circ$, $\lambda = 25 \div 42^\circ$, ограниченный для каждой из трех упоминавшихся сейсмостанций прямоугольной областью. В избранном регионе "поля", контролируемые с/ст "Сквира" и "Полтава", имеют размеры 1250×1500 км, с/ст "Одесса" — 1250×1200 км. На множестве записей землетрясений, выбранных из

$\sim 15^\circ$ окрестности этих сейсмостанций, были созданы нейросетевые модели времен прихода *P*- и *S*-фаз сейсмических волн. Модели позволяют прогнозировать для каждой станции эти времена в виде трехмерных годографов, а представление (11) делает возможным розыгрыш различных сценариев влияния на времена пробега сейсмических волн параметров очага, таких как магнитуда и глубина гипоцентра, входящих в число аргументов функции годографа.

На рис. 5 представлены наблюдаемые на с/ст "Одесса", "Сквира", "Полтава" времена прихода *P*- и *S*-фаз сигналов, генерируемых землетрясениями Румынии, Крыма, Кавказа и Средиземноморья, а также обобщенные годографы, смоделированные для обратного азимута 180° и магнитуды 5.

На схеме (рис. 6) видна неравномерность покрытия исследуемого пространства сейсмическими источниками, которые практичес-

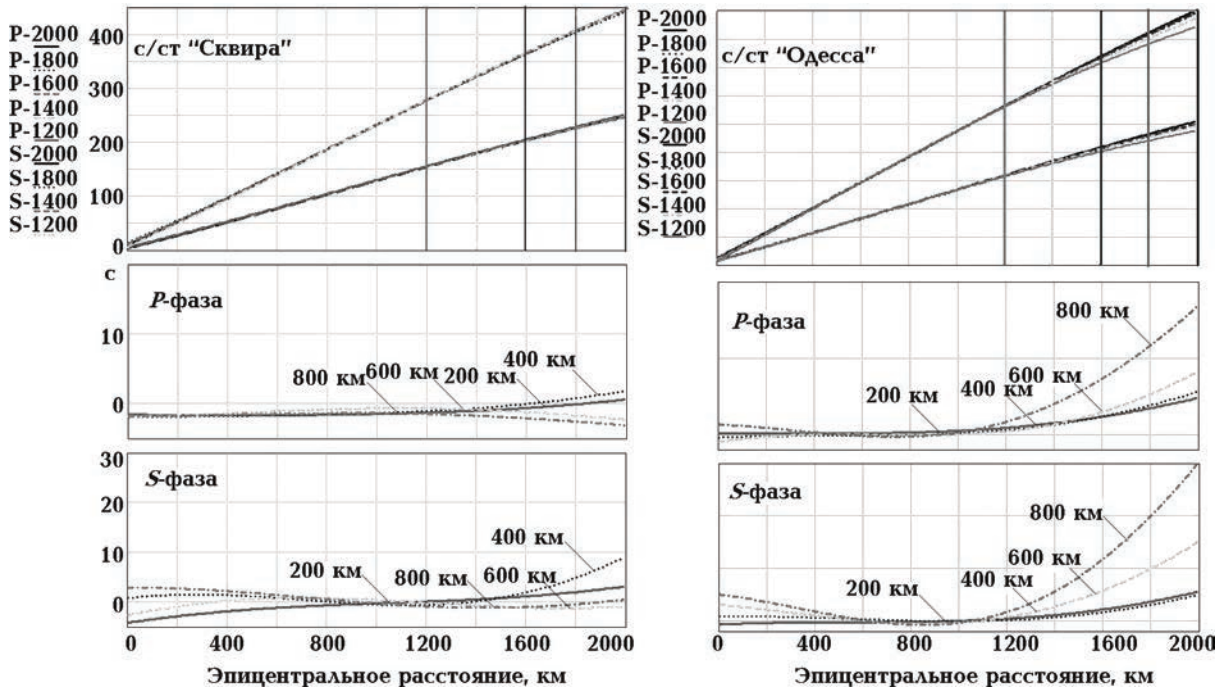
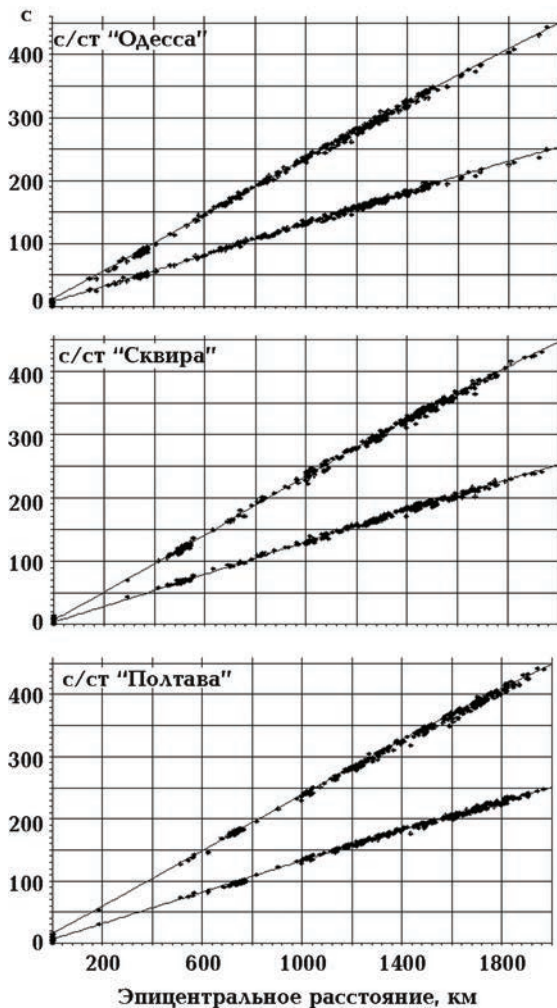


Рис. 4. Ошибки, связанные с различными значениями длины экстраполяции годографа P - и S -фаз сейсмических волн, зарегистрированных на с/ст "Одесса" и "Сквира".



ки отсутствуют вблизи сейсмостанции. Отсутствие здесь информации заставляет модель продолжить функцию годографа в область около нулевых времен по закону, определяемому максимальной информационной нагрузкой — временами прихода рефрагированных волн. Поэтому для данной представительности обучающей выборки модель склонна давать максимальные ошибки вблизи сейсмостанции.

Однако амплитуды изменений времен вдоль годографа, вызываемых сейсмической анизотропией либо связанные с вариаций параметров (11), не экспликативны при рассмотрении их модулей. Значительно выразительней такие изменения могут оказаться, если будут отнесены к некому уровню, физическая природа которого известна и является своего рода эталоном. В случае моделируемых годографов таким референц-уровнем может служить го-

→ Рис. 5. Годографы P - и S -фаз сейсмических волн на фоне зарегистрированных на с/ст "Одесса", "Сквира", "Полтава" времен их прихода, использованных в качестве целевых значений при обучении нейронной сети.

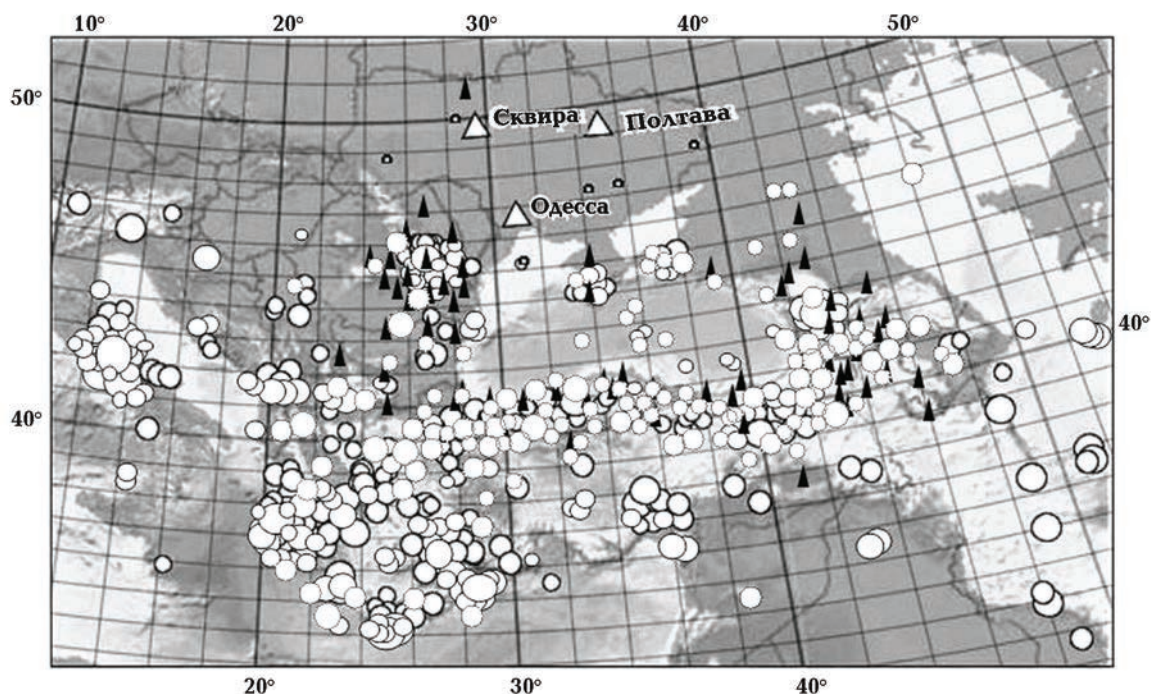


Рис. 6. Схема региона сбора данных. Светлые треугольники — пункты регистрации сети сейсмических станций Украины, черные — международные сейсмологические центры.

дограф Джеффриса — Буллена, относящийся к стандартной сферически-симметричной модели Земли.

Многочисленное приближение годографа (11) нейросетевой моделью предполагает возможность розыгрыша различных сценариев поведения функции времени прихода фаз сейсмических волн в зависимости от величин параметров. На рис. 7 приведена трехмерная зависимость этой функции от глубины гипоцентра в терминах невязки моделируемого времени пробега фазы с таковым глобальным. В такой "разностной" постановке отклонения от "опорного" уровня, или невязки, могут оказаться полезными в выявлении пространственных закономерностей и аномалий в поведении времен пробега (скоростей?) фаз сейсмических волн в изучаемом пространстве.

Отклонение времен пробега от табличных Джеффриса — Буллена в некоторых случаях могут лежать в интервалах, определенных выше как возможный диапазон ошибок. Область вблизи сейсмостанций требует более детального освещения близлежащими источниками, а общерегиональные закономерности деформаций поля времен прихода фаз сейсмических волн, очевидные из рисунков, могут ока-

заться ценными в качестве аргументов при тектоно-геологических построениях.

Обращение годографов. Нейросетевая модель 4-мерного поля времен прихода фаз сейсмических волн к точке наблюдения позволяет генерировать годографы для произвольного азимута в плоскости исследуемого региона. Известно [Lay, Wallace, 1995], что в случае среды, обладающей в горизонтальном направлении нулевым градиентом скорости распространения упругих волн и неубывающим градиентом с глубиной, обратная задача для годографа имеет единственное решение Герглотца — Вихерта и в форме Абеля записывается

$$z(x) = \frac{1}{\pi} \int_0^x \ln \left(k + \sqrt{k^2 - 1} \right) dx, \quad (13)$$

где $k = \frac{\left(\frac{\partial x}{\partial t} \right)_x}{x/t}$, x — расстояние до точки

выхода луча на земную поверхность, t — время в точке выхода,

$$v [z(x)] = \left(\frac{dx}{dt} \right)_x \quad (14)$$

— фазовая скорость в точке максимального погружения луча в среду.

Представим некоторый участок исследуемого региона (см. рис. 6) в виде сеточной области (рис. 8). Для любой длины годографа, смоделированного вдоль трассы, проведенной от сейсмостанции через произвольный узел области, могут быть получены оценки максимальной глубины $z(x)$ (13) проникновения луча и фазовой скорости $v[z(x)]$ (14) на этой глубине. Фактически эти оценки получаются для некой бесконечно тонкой пластины, вдоль верхней грани которой лежит обсуждаемая трасса.

Так как по определению горизонтальный градиент скорости в среде равен нулю, то все оценки скоростей и глубин, полученные для годографов различной длины вдоль трассы, могут быть отнесены к произвольной точке на трассе, в частности к узлу сетки, подсоединяемому трассой, и записаны в виде двух последовательностей: глубин и соответствующих этим глубинам скоростей. Очевидно, что такие два вектора могут быть получены для каждого узла сетки, и с помощью интерполяции возможно построить в пределах сеточной области распределение фазовых скоростей (14) для любой глубины из определенного диапазона (13).

Одно из основных несоответствий модели Герглотца — Вихерта реальной среде содержится в требовании положительности градиента скорости с глубиной, т. е. в запрете "слоев" с пониженной скоростью, хотя высокая вероятность наличия таких зон широко известна как в масштабах разведочной сейсмики [Pavlenkova, 2004], так и сейсмологии [Aleksiev et al., 1973]. Современные представления о глубинном строении Земли [Соллоуб, 1982; Литосфера ..., 1989; Чекунов и др., 1990] содержат сведения о высокой вероятности встречи в верхней мантии на глубине ~ 60—70 км с зонами пониженной скорости, из-за чего адекватность реальному строению среды результатов использования одномерной модели обращения годографа (13, 14) ниже этих глубин может вызывать сомнения. Действительно, среда Герглотца — Вихерта предполагает увеличение максимальной глубины проникновения луча с увеличением длины годографа, поэтому отклонения от этой закономерности могут служить признаком нарушения условия $\text{grad } V(z) \geq 0$.

На рис. 9 приведена зависимость максимальной глубины проникновения луча от направления прихода сейсмического сигнала, полученная при обращении годографов, генерируемых нейросетевой моделью для с/ст "Сквира", по направлениям основных сейсмоактивных "центров" региона (см. рис. 6): Кавказ, Турция, Греция, Вранча, Италия. Видно, что искажение в кривой роста глубин максимального проникновения луча наблюдается вблизи глубины ~ 60 км. Подобное поведение решения (13) требует определенной осторожности при использовании модели обращения Герглотца — Вихерта в настоящем приложении для рассматриваемого региона.

Как отмечалось ранее, в каждом узле сеточной области, согласно схеме рис. 8, с помощью решений (13, 14) возможно получить значения максимальных глубины проникновения луча и соответствующих им скоростей в виде двух последовательностей, что позволяет путем интерполяции оценить скорость на любой глубине в пределах полученного интервала.

На рис. 10 для с/ст "Сквира" в качестве примера приведен результат такой интерполяции, выполненной методом кубических сплайнов, для глубины 50 км.

Оценки скоростей в узловых точках сеточной области, выполненные по "верной" схеме (см. рис. 8), не учитывают вариации изменения скорости вдоль трассы, проходящей через данный узел, а моделируют лишь зависимость от направления прихода волны к сейсмостанции, определяемую градиентом скорости между трассами. Очевидно, что шаг трасс, контролируемый густотой сетки, определяет дискретность этого квазитрехмерного изображения.

Характер поведения кривой максимальных глубин проникновения луча в зависимости от длины годографа может оказаться полезным для глубинных оценок состояния геологической среды и локальных вариаций взаимоотношения элементов структуры изучаемого региона.

Понятно, что построение годографов для отдельных сейсмостанций, использующее обобщающий (4-мерный) подход (11), хотя и расширяет палитру изучаемых характеристик среды распространения сейсмических волн по сравнению с традиционным подходом, использующим 1-мерный годограф, но сужает его приложение, позволяя получать лишь градиентные оценки (см. рис. 10) скоростей фаз сейс-

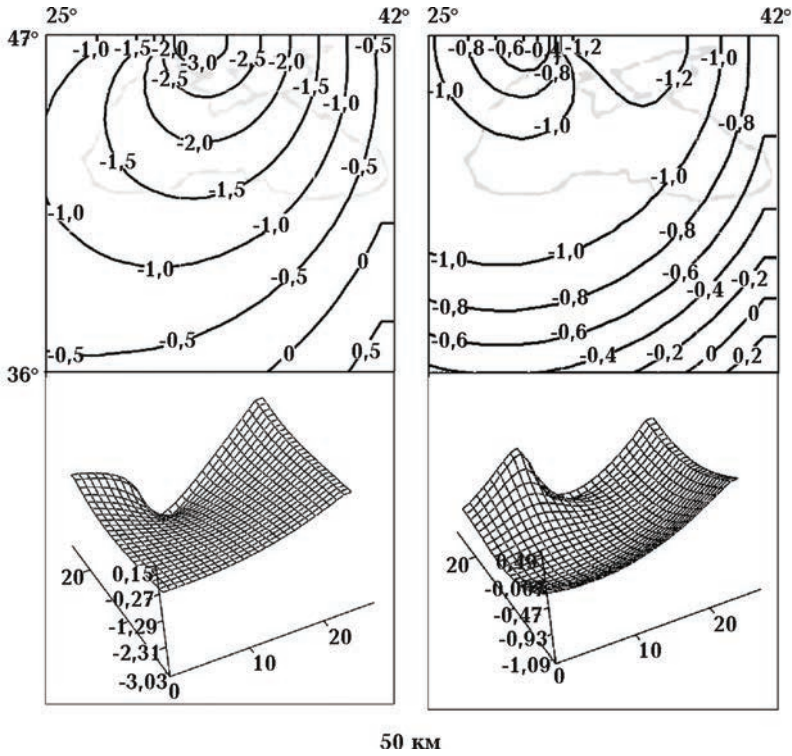


Рис. 7 (окончание).

мических волн на глубинах, ограничиваемых "гладкостью" поведения зависимости максимальной глубины проникновения луча от эпицентрального расстояния.

При построении локальных годографов для единичной сейсмостанции координаты последней не являются информативными параметрами, так как для всех векторов-примеров поведения функции времени прихода фазы они единственны, а положение эпицентра однозначно определяется составляющими вектора: эпицентральной расстоянием и азимутом. Создание же обобщенной модели для некоего региона предполагает компиляцию обучающей выборки в широком диапазоне эпицентральных расстояний, содержащей примеры событий, возбужденных не только во многих источниках, но и зарегистрированных на многих сейсмостанциях. Такой подход направлен на обеспечение покрытия трассами распространения сигналов интересующего региона с возможно большей равномерностью и, предполагая наличие сейсмической анизотропии, со всех направлений. В этом случае положение эпицентра уже однозначно не определяется эпицентральной расстоянием и азимутом, а должно задаваться в явном виде.

Поэтому мощность вектора (11) увеличивается до $N=9$, где x_5 — долгота станции, x_6 — широта станции, x_7 — превышение сейсмостанции над уровнем моря, x_8 — долгота источника, x_9 — широта источника.

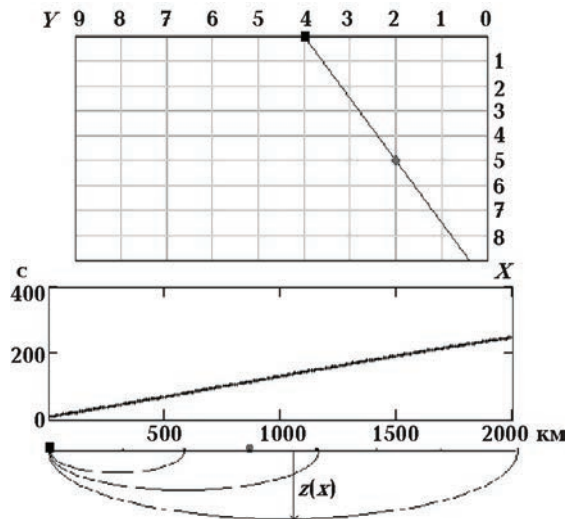


Рис. 8. Схема получения оценок максимальных глубин проникновения луча и фазовых скоростей на этих глубинах.

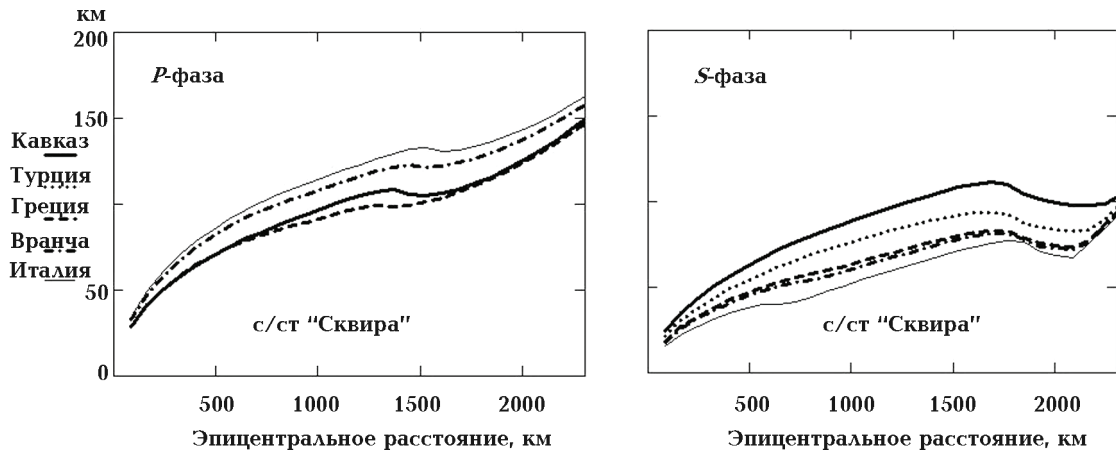


Рис. 9. Зависимость максимальной глубины проникновения луча P - и S -фаз сейсмической волны от эпицентрального расстояния в зависимости от направления на источник сейсмического возбуждения для с/ст "Сквира".

Для обсуждаемых результатов использовались данные регистрации 4142 событий в исследуемом регионе с привлечением качественной информации международных сейсмологических центров ISC, NEIC, ORF (1982—2009 гг.). Такая выборка, используемая в качестве обучающей для нейронной сети, позволяет построить нейросетевую модель, отображающую времена прихода сейсмических фаз от произвольной точки возбуждения до произвольной точки регистрации региона исследования.

Рассматривая узлы сеточной области в качестве срединной точки семейства годографов, расположенных вдоль проходящей че-

рез узел трассы, 9-мерная нейромодель позволяет смоделировать годографы вдоль трасс, проходящих с произвольными азимутами через данный узел. Такая модель обобщает результаты 4-мерной, устраняя "одноточечную" привязку к конкретной сейсмостанции и дает возможность осуществить покрытие пространства вокруг узла, рассматриваемого в качестве срединной точки для семейства годографов, с желаемой плотностью расположения трасс.

На рис. 11 приведен пример распределения градиента скоростей P - и S -фаз сейсмических волн в черноморском регионе на глубине 60 км как математического ожидания

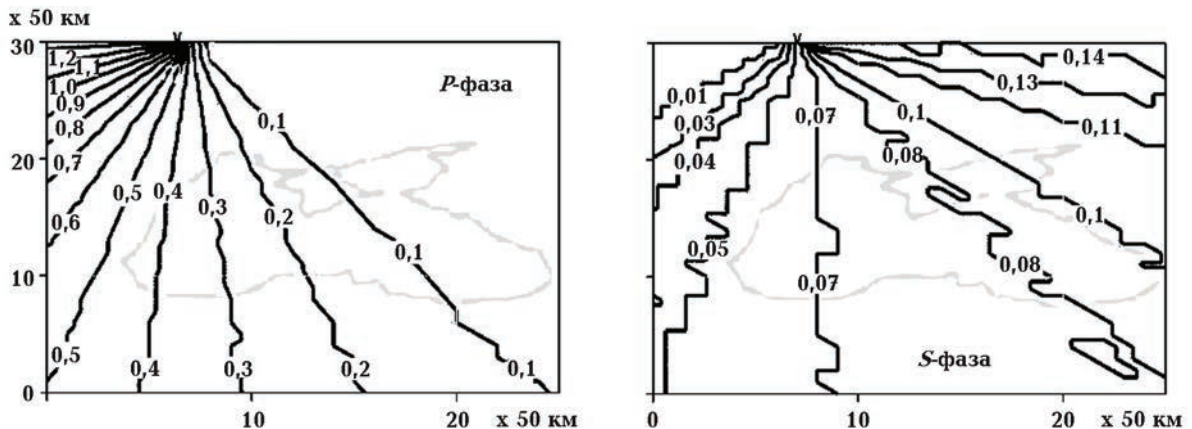


Рис. 10. Градиент скоростей P - и S -фаз сейсмических волн на глубине 50 км, полученный в результате обращения годографов, сгенерированных нейросетевой моделью времен вступления фаз, зарегистрированных на с/ст "Сквира".

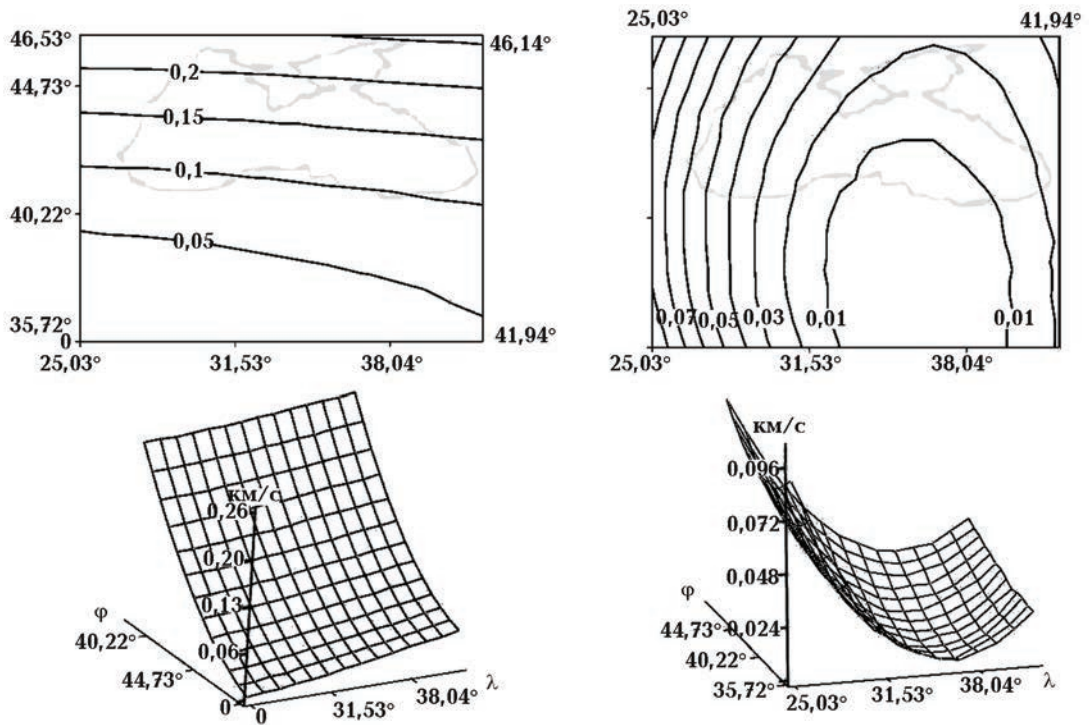


Рис. 11. Градиент скоростей P - и S -фаз сейсмических волн в черноморском регионе на глубине 60 км, полученный в результате Герглотца — Вихерта обращения годографов, сгенерированных 9-мерной нейромоделью.

результатов обращения (13, 14) семейства прямых и обратных годографов вдоль четырех трасс, проходящих через каждый узел сеточной области с шагом 45° .

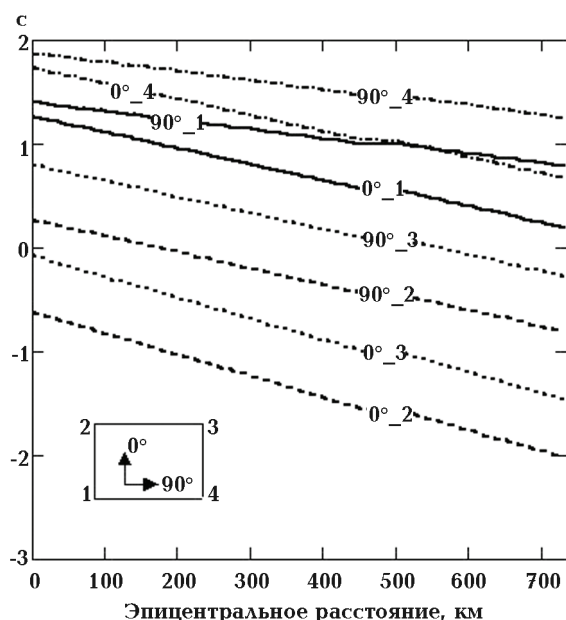
Наблюдаемая зависимость времен пробега сейсмических волн от направления их распространения может обуславливаться многими факторами, включая анизотропию среды распространения волн, зашумленность каталожных данных, вызванную погрешностями в определении координат эпицентров землетрясений и времен вступления фаз, погрешность аппроксимации нейронной моделью реальной функции времени прихода фазы сейсмической волны. Такую зависимость времен вступления можно обнаружить, оценивая разность времен пробега фаз в прямом и обратном годографах. На рис. 12 показана разность между прямым (источник на север от узла) и обратным (источник на юг) годографами вдоль трасс с азимутом 0° . Для азимута 90° прямым считается годограф с источником на запад от узла и обратным с источником на восток от узла.

Ошибки определения координат эпицентров и времен прихода фаз для более чем че-

тырех тысяч событий, зарегистрированных на 56 сейсмостанциях, вероятнее всего будут случайным образом искажать разностные кривые рис. 12. Однако наблюдаются определенные закономерности, довольно четко сегрегирующие север и юг рассматриваемого участка черноморского региона (азимут 0°) с позиций анизотропии.

Настоящее сообщение не ставит своей целью связать результаты различных сценариев поведения моделируемой функции с геологическим строением конкретного региона исследований. Широко доступные сейсмологические данные в черноморском регионе использовались как для демонстрации возможностей построения эмпирической модели пространственного поведения функции времен пробега сейсмических волн, так и проигрывания различных сценариев этого поведения, контролируемых параметрами, ранее не используемыми.

Основное внимание уделялось вопросам выбора оптимальной архитектуры нейронной сети с точки зрения как процесса обучения, так и оценок различных режимов ее функционирования в операционном режиме, а так-



→ Рис. 12. Разность прямых и обратных годографов P-фазы для четырех срединных точек (см. врезку) в углах региона (согласно схеме рис. 11) для двух взаимно перпендикулярных направлений (азимуты 0° и 90°).

же адекватности результатов моделирования реальным ситуациям.

Авторы надеются, что продемонстрированная относительная простота получения поведенческих характеристик и зависимостей окажутся привлекательными для широкого круга геофизической общественности и помогут преодолеть основное препятствие на пути широкого использования этого эффективного метода статистического анализа — неумение формулировать свои задачи в терминах, позволяющих их интеграцию в нейросетевых решениях.

Список литературы

- Горбань А. Н. Обучение нейронных сетей. — Москва: СП ПараГраф, 1990. — 160 с.
- Колмогоров А. Н. О представлении непрерывных функций нескольких переменных суперпозициями непрерывных функций меньшего числа переменных // Докл. АН СССР. — 1956. — **108**, № 2. — С. 179—182.
- Лазаренко М.А., Герасименко О.А. Імплементация нейронно-мережевого підходу до розв'язання задачі прогнозу інтенсивності стрясань // Вісн. КНУ. Геологія. — 2007. — Вип. 41. — С. 21—25.
- Литосфера Центральной и Восточной Европы. Восточно-Европейская платформа / Отв. ред. В. Б. Соллогуб. — Киев: Наук. думка, 1989. — 180 с.
- Мак-Каллок У. С., Питтс В. Логическое исчисление идей, относящихся к нервной активности // Автоматы. — Москва: Изд-во иностр. лит., 1956. — С. 363—384.
- Соллогуб В. Б. Земная кора Украины // Геофиз. журн. — 1982. — **4**, № 4. — С. 3—25.
- Стюарт Р., Питер Н. Искусственный интеллект: современный подход (АИМА). — 2-е изд. — Москва: Вильямс, 2007. — 1424 с.
- Уоссермен Ф. Нейрокомпьютерная техника. — Москва: Мир, 1992. — 184 с.
- Хайкин С. Нейронные сети: Полный курс. 2-е изд., испр. — Москва: Вильямс, 2008. — 1103 с.
- Чекунов А. В., Соллогуб В. Б., Старостенко В. И., Красовский С. С., Кутас Р. И., Пашкевич И. К., Ильченко Т. В., Калюжная Л. Т., Соллогуб Н. В., Оровецкий Ю. П., Щербаков И. Б., Гавриленко Н. М., Зарицкий А. И., Галецкий Л. С., Деметреску К., Корня И., Рэдулеску Ф., Онческу М., Станика Д. Геотрансект Днепровско-Донецкий палеорифт — Украинский щит — Южные Карпаты // Геофиз. журн. — 1990. — **12**, № 6. — С. 3—32.
- Alekseev A. S., Belonosova A. V., Burmakov I. A., Krasnovevtseva G. V., Matveeva N. N., Nercesov G. L., Pavlenkova N. I., Romanov V. G., Ryaboy V. Z. Seismic studies of low-velocity layers and horizontal inhomogeneities within the crust and upper mantle on the territory of the USSR // Tectonophysics. — 1973. — **20**. — P. 47—56.
- Caplinskas A. All paradigms // J. Intelligent Manufacturing. — 1998. — **9**, № 6. — P. 493—502(10).
- Chauvin Y., Rumelhart D. E. Back Propagation: Theory, Architectures, and Applications, Lawrence Erlbaum Associates. — New Jersey: Publishers Hillsdale, 1995. — 564 p.
- Cybenko G. Approximation by superpositions of sigmoidal functions // Math. Control Signal Systems. — 1989. — **22**. — P. 303—314.

- Hecht-Nielsen R.* Kolmogorov's mapping neural network existence theorem // IEEE First Annual Int. Conf. on Neural Networks. — San Diego, 1987. — **3**. — P. 11—13.
- Hornik K., Stinchcomb M., White H.* Multilayer Feedforward Networks are Universal Approximators // Neural Networks. — 1989. — **2**. — P. 359—366.
- Jeffreys H., Bullen K.E.* Seismological Tables. — London: British Association Seismological Committee, 1940. — 145 p.
- Kennet B.L.N.* Seismological Tables. — Canberra: Res. School Earth Sci., Australian National University, 1991. — 167 p.
- Lay T., Wallace T.C.* Modern Global Seismology. — New York: Acad. press Inc., 1995. — 517 p.
- Pavlenkova N.I.* Low velocity and low electrical resistivity layers in the middle crust // Ann. Geophys., Editrice Compositori srl. — 2004. — **47**, № 1. — P. 157—169.
- Stornetta W. S., Huberman B. A.* An improved three-layer, back-propagation algorithm // Proc. of the IEEE 1st Intern. Conf. on Neural Networks. — San-Diego: SOS Printing, 1987. — P. 37—44.