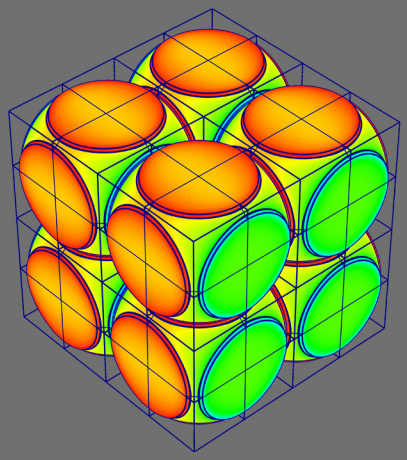
ISSN 2309-3684



Математическое моделирование и численные методы

Воронин Е.А., Носов В.Н., Савин А.С. Нейросетевой подход к решению обратной задачи генерации поверхностных волн. Математическое моделирование и численные методы, 2025, № 4, с. 173–193.

Источник: https://mmcm.bmstu.ru/articles/395/

Параметры загрузки:
IP: 216.73.216.253
06.11.2026 00:28:32

УДК 681.3:534.2 DOI: 10.18698/2309-3684-2025-4-173193

Нейросетевой подход к решению обратной задачи генерации поверхностных волн

© Е.А. Воронин[1], В.Н. Носов[2], А.С. Савин[3]

[1]ФИЦ ИУ РАН, Москва, 119333, Россия
[2]ГЕОХИ РАН, Москва, 119991, Россия
[3]МГТУ им. Н.Э. Баумана, 105005, Россия

Рассмотрен нейросетевой подход к решению обратной задачи генерации поверхностных волн или идентификации источника возмущения по данным, снятым со свободной поверхности жидкости. Выбран тип и построена математическая и компьютерная модель нейронной сети на основе данных, полученных в ходе лабораторного эксперимента по изучению поверхностных волн, возникающих при движении в жидкости крылового профиля и кругового цилиндра. Для оценки адекватности этой модели проведены численные эксперименты по определению глубины и скорости источника возмущений жидкости. Результаты расчёта сопоставлены с данными эксперимента. Разработан программный комплекс, позволяющий эффективно и с минимальными затратами времени решать задачу определения параметров источника возмущения по данным, снятым с водной поверхности.

Ключевые слова: *обратная задача генерации поверхностных волн; нейронные сети; обучение нейронных сетей; метод стохастического градиентного спуска; мониторинг состояния нейронной сети; спектральный анализ; алгоритм обратного распространения ошибки; задача регрессии; сети прямого распространения сигнала; многослойный персептрон*

Введение. В настоящее время освоение Мирового океана приобретает все более возрастающее значение. При осуществлении такой деятельности требуется создавать новые методы получения информации о протекающих в морской среде процессах. Непосредственное наблюдение за явлениями в морской среде, особенно на значительных глубинах, с помощью современных средств носит локальный характер. С другой стороны, оптические и радиолокационные средства дистанционного зондирования морской среды способны предоставлять данные с больших участков морской поверхности. При этом ни оптическое, ни радиолокационное излучения не способны проникать на значительные глубины. Необходимость изучения процессов, протекающих в толще морской среды, по данным о ее свободной поверхности, полученных с помощью радиолокационных средств, привела к созданию нового научного направления - компьютерной радиотомографии морской среды [1]. В основе этого подхода лежит комплексная модель информационного тракта радиотомографического комплекса, включающая в себя гидрофизические и электродинамические блоки, а также блоки классификации и распознавания подводных явлений.

Важнейшей составной частью этого научного направления является разработка методов решения обратных задач возмущения морской поверхности различными источниками. Например, обратная задача генерации поверхностных волн заключается в восстановлении внутренней структуры потока тяжелой жидкости и определении параметров источника возмущений по данным, снятым со свободной поверхности. В работе [2] предложен подход к решению этой задачи, основанный на замене источника возмущений его моделью в виде системы гидродинамических особенностей (источников, стоков, вихрей). Восстановление структуры потока по данным о его свободной поверхности сопряжено с преодолением значительных трудностей. Продолжение полей в сторону источников относится к классу некорректно поставленных задач [3]. Решения таких задач неустойчивы к малым изменениям исходных данных, чем объясняются трудности применения предлагаемых методов в реальных случаях. Например, в работах [4,5] предложено решение этой обратной задачи, основанное на технике аппроксимаций Паде. Однако рассмотренная процедура применима лишь в случае абсолютно точного задания исходных данных.

Настоящая работа представляет собой дальнейшее развитие подходов, к решению обратной задачи генерации поверхностных волн, предложенных в [6, 7]. Эти подходы основаны на применении методов машинного обучения, а именно, теории нейронных сетей, который широко применяется в задачах обработки данных и их анализа [8-17]. Для численной реализации полученных результатов разработан программный комплекс, с помощью которого по экспериментальным данным, содержащим результаты измерений наклона свободной поверхности, можно определить параметры источника возмущения – глубину и скорость. Для этого потребовалось применить методы машинного обучения, в частности искусственную нейронную сеть, для решения обратной задачи генерации поверхностных волн, выбрать тип нейронной сети и построить её математическую модель для предсказания глубины и скорости источника возмущения в толще жидкости, реализовать алгоритм обучения построенной модели в виде соответствующего программного комплекса, провести численные эксперименты и сравнить полученные результаты с данными, найденными в ходе лабораторного эксперимента для обоснования и подтверждения адекватности выбранного решения.

Подход на основе математической модели. Для применения технологии машинного обучения к задаче о восстановлении параметров источника возмущений жидкости по данным о ее свободной поверхности необходимо иметь обучающую выборку. При недостатке экспериментальных данных её можно получить путём численного решения прямой задачи. Приведем постановку и основные результа-

ты решения плоской задачи о генерации поверхностных волн при обтекании точечных гидродинамических особенностей в приближении малых волн. Направим ось Ox прямоугольной декартовой системы координат по невозмущенной свободной границе жидкости в сторону течения, а ось Oy вверх против силы тяжести, тогда при $x = \infty$ каждая жидкая частица обладает скоростью $(V, 0)$, $V = const > 0$ (рис. 1).

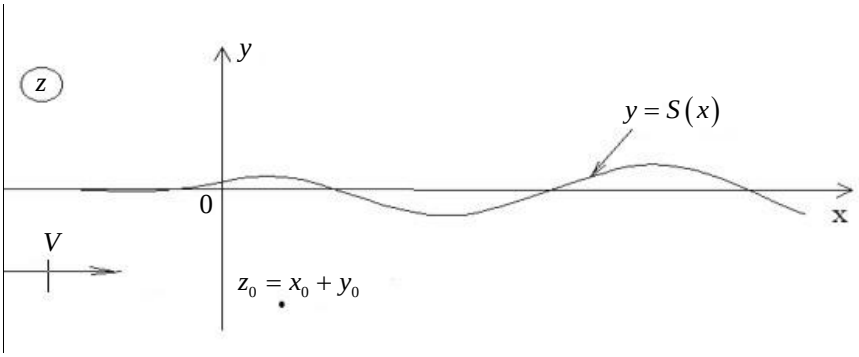


Рис. 1. Общая схема обтекания точечной гидродинамической особенности потоком со свободной границей

Выражение для профиля свободной границы потока, обтекающего точечную гидродинамическую особенность (источник, вихрь, диполь, мультиполь), локализованную в точке $z_0 = x_0 + y_0$, можно представить в виде

$$S(x) = -\frac{2}{V}\int_{-\infty}^{x} \mathrm{Re}\left\{ \mathrm{i}\bar{C}\mathrm{e}^{\mathrm{i}\nu(\xi-x)}\,\frac{\mathrm{d}}{\mathrm{d}\xi}\mathrm{q}(\xi-\overline{z_0})\right\}\mathrm{d}\xi,$$
$$\nu = g/V^2, \quad \mathrm{q}(z-z_0) = \begin{cases} \ln(z-z_0), & n=0, \\ (z-z_0)^{-n}, & n=1,2,\ldots. \end{cases} \tag{1}$$

С вычислительной точки зрения использование этой формулы для получения профиля свободной границы жидкости оправдано, если это требуется сделать однократно или небольшое количество раз. Дело в том, что при вычислении отклонения поверхности жидкости от равновесного положения в некоторой точке по формуле (1) никак не используется информация, полученная при вычислении этой величины в предыдущих точках. Это приводит к замедлению вычислительного процесса, скорость протекания которого, очевидно, существенна при создании обучающей выборки, включающей в себя множество реализаций. В этой связи задачу для определения профиля свободной границы потока, обтекающего точечную гидродинамическую особенность, целесообразно переформулировать и свести к задаче Коши для обыкновенного дифференциального уравнения,

численное решение которого основано на определении текущего значения искомой функции по ее значениям, вычисленным на предыдущих шагах. Можно показать, что уравнение профиля свободной границы потока имеет вид

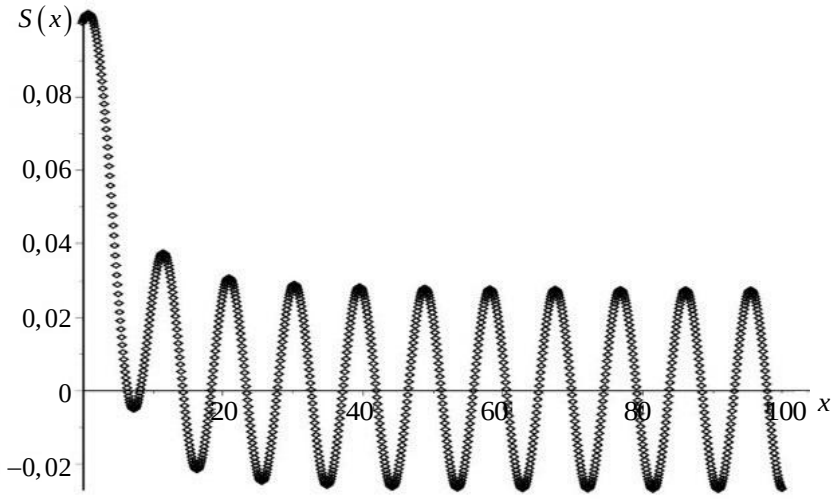
$$\frac{\mathrm{d}^2\mathrm{S}}{\mathrm{dx}^2}+\nu^2\mathrm{S}=\frac{2(\mathrm{n}\text{-}\delta_{0\mathrm{n}})}{\mathrm{V}}\,\mathrm{Re}\left\{\mathrm{C}\left(\frac{\mathrm{i}(\mathrm{n}+1)}{(x-z_0)^{\mathrm{n}+2}}+\frac{\nu}{(x-z_0)^{\mathrm{n}+1}}\right)\right\}. \qquad (2)$$

Поскольку гравитационные поверхностные волны развиваются в основном за обтекаемой особенностью, а вверх по потоку быстро затухают, профиль свободной поверхности $S=S(x)$ находится как решение соответствующего уравнения (2), удовлетворяющее условиям излучения $S(-\infty)=0$, $S'(-\infty)=0$.

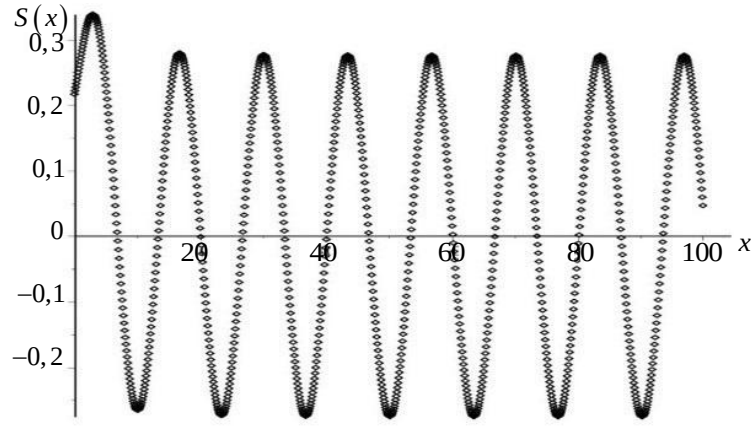
Используя формулу (1), можно найти отклонение поверхности жидкости и его производную в любой выбранной точке. Эти значения представляют собой данные Коши для уравнения (2). Например, обтеканию точечного вихря интенсивности Γ, локализованного в точке $z_0=-ih$, соответствует задача

$$\begin{gathered}\frac{\mathrm{d}^2\mathrm{S}}{\mathrm{dx}^2}+\nu^2\mathrm{S}=\frac{\Gamma}{\pi V}\cdot\frac{1}{h^2+\mathrm{x}^2}\cdot\left(\frac{h^2-\mathrm{x}^2}{h^2+\mathrm{x}^2}+\nu h\right),\\ S(0)=\frac{\Gamma}{\pi V}e^{-\nu h}Ei(\nu h),\ \frac{dS}{dx}(0)=\frac{\Gamma\nu}{V}e^{-\nu h},\end{gathered} \qquad (3)$$

Пусть в качестве гидродинамической особенности выбран точечный вихрь. Тогда, численно решая задачу (3) методом Рунге-Кутты 4-го порядка, можно получить обучающую выборку.



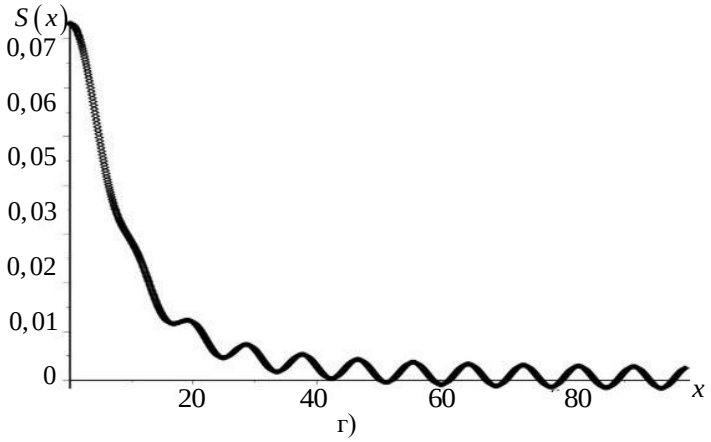
а)



б)



в)



г)

Рис. 2. Решение задачи, соответствующей обтеканию точечного вихря интенсивности Γ, локализованного в точке $z_0 = -ih$, при заданных параметрах:

1) $\Gamma = 4.389$, $h = 3.762$, $\nu = 0.469$, $V = 4.569$ 2): $\Gamma = 3.123$, $h = 5.624$, $\nu = 0.674$, $V = 3.813$ 3) $\Gamma = 0.436$, $h = 1.565$, $\nu = 3.052$, $V = 1.792$

4) $\Gamma = 2.441$, $h = 9.064$, $\nu = 0.857$, $V = 3.381$

Это модельное решение, для получения обучающей выборки его можно смешать с некоторым случайным сигналом. Однако для того, чтобы приблизиться к реальным условиям окружающей среды, выгоднее использовать данные для обучения, полученные из реального эксперимента. В Институте проблем механики РАН было создано специальное оборудование и разработан метод отражённой сетки, позволяющий наблюдать возмущение свободной поверхности жидкости с высокой точностью [18-21]. В частности, проводились измерения возмущений поверхности воды, вызываемых движущимися цилиндром и крыловым профилем [18-20]. В этих экспериментах были получены значения тангенса угла наклона поверхности воды в точках с горизонтальными координатами x при различных скоростях V движения цилиндра и крылового профиля. Эти данные использованы в настоящей работе для решения поставленной задачи.

Подход на основе машинного обучения и нейронных сетей. Формально данная задача обучения нейросети формулируется как задача аппроксимации, постановка которой имеет следующий вид. Пусть задана область $\Omega : \vec{r} = (x_1, x_2, x_3) \in \Omega \subset \mathbb{R}^3$ и заданы два множества точек:

$Q_N = \{r_n\}_{n=1}^N \in \Omega$ — обучающее множество;

$Q_M = \{r_m\}_{m=1}^M \in \Omega$ — целевое множество.

$f(Q_N)$ — значения функции в точках обучающего множеств,

$F(r)$ — аппроксимирующая функция, которая строится с помощью нейронной сети как непрерывное отображение $F : Q_n \to Q_m$.

Необходимо построить нейронную сеть (аппроксимирующую функцию), которая будет принимать совпадающие значения (с заданной точностью) не только на данных, участвовавших в обучении, но и на данных контрольного множества (целевого) не участвовавших в обучении. На вход поступают координаты точки $r \in \Omega$, далее происходит преобразование сигнала в нейронной сети. На выходе получаем выходной сигнал $F(r)$. Для этого подходит сеть прямого распространения сигнала такие как многослойный персептрон, обучаемый алгоритмами, основанными на методе обратного распространения ошибки [10].

Важной проблемой наряду с получением данных для обучения и их обработкой служит нахождение оптимальной архитектуры нейронной сети, а именно поиском количества слоем, числа нейронов на скрытом слое. После проведения достаточного количества численных экспериментов была выбрана архитектура полносвязной нейронной сети, которая имеет следующий вид ориентированного графа (рис. 3). Во входном слое имеется 3 нейрона (вершины), на ко-

торые в процессе обучения и применения приходят значения амплитуд. В первом (скрытом) слое имеется некоторое количество (от 20 до 70 в зависимости от эксперимента) нейронов с активационной функцией $\mathrm{Re}\,LU$, либо фукцией активации гиперболический тангенс, каждый из которых соединен синапсом (ребром) с каждым из входящего слоя. Такую сеть здесь и далее будем обозначать следующим образом: 3-20-3. Это означает, что сеть имеет 3 входных, 20 скрытых и 3 выходных нейрона.

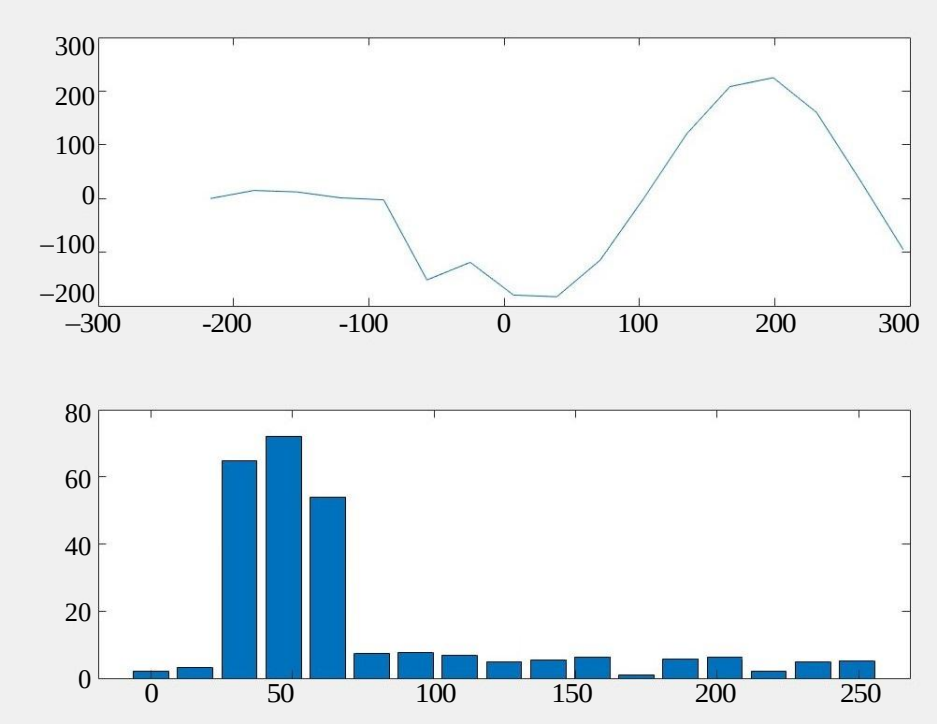


Рис. 3. Тангенс угла наклона волны поверхностного возмущения и её спектр, полученный методом быстрого преобразования Фурье

Важнейшим этапом нейросетевого моделирования является выбор характерных признаков изучаемого процесса или объекта. В последнее время для этого применяют сети глубокого обучения, но для этого требуется большой объём исходной информации и значительные вычислительные ресурсы или длительное время процесса обучения, что недостижимо в реальных условиях и не отвечает требованиям оперативной обработки информации и идентификации параметров возмущающего объекта. Поэтому для выявления характерных признаков процесса и сокращения обучающей выборки было принято решение к данным, полученным в ходе лабораторного эксперимента, применить спектральный анализ по алгоритму быстрого преобразования Фурье. Он позволяет удалять шумовые компоненты изучаемо-

го процесса, а спектр инвариантен относительно смещения результатов наблюдения, что очень важно в реальных условиях исследований. Визуально полученный спектр Фурье представлен на рис. 3.

Установлено, в результате спектрального анализа экспериментальных данных, что частоты имеют примерно одинаковые значения и имеет три характерные частоты, а корреляционный анализ между его характеристиками и параметрами возмущающего воздействия показал, что, для данной выборки экспериментальных данных корреляция с частотой основных гармоник слабая. Поэтому значения частот можно отбросить в виду их низкой вариативности и слабой связи между целевыми метками — H и V (глубиной и скоростью соответственно). Из всего этого следует решение о том, что для обучения нейронной сети достаточной будет выборка с тремя входами – значениями получившихся амплитуд.

Несмотря на то, что для применения нейронных сетей обычно во всем используют экспериментальный подход, для определения размера обучающей выборки, как правило, рекомендуется придерживаться следующего неравенства [10]:

$$N > \frac{W}{\varepsilon}, \tag{4}$$

где N — число учебных образцов, W — число весовых коэффициентов в сети, ε — доля ошибок, которая допустима в ходе тестирования. Для нашей архитектуры нейронной сети имеем: число весовых коэффициентов — $W = 3 \cdot 50 + 50 \cdot 20 + 20 = 1170$, доля ошибок, которая допустима в ходе тестирования — $\varepsilon = 0{,}1$. Получаем, что для заданной ошибки количество примеров обучения должно в 10 раз превосходить количество весовых коэффициентов сети. Но реальных данных значительно меньше требуемых, поэтому было решено на основе имеющихся данных сгенерировать выборку большего размера путем зашумления данных. Это позволит придать сети свойство обобщения и избежать излишне точного соответствия нейронной сети конкретному набору обучающих примеров или переобучения.

Была сформировано выборка большего объема путем 5 % зашумления (выбор интервала зашумления был выявлен путем эксперимента). Это помогло достичь хорошего обобщения. Свойство обобщения показывает то, насколько хорошо сеть выполняет свои функции на новых данных, которые не участвовали в процессе обучения. Для того, чтобы провести процесс обучения по вариативности факторов, вначале было проведено нормирование данных, т.е. их привели к относительным единицам в интервале (0,1) путем деления каждого параметра на максимальное значение.

Выбор метода обучения нейронной сети – основной этап её проектирования. Обучение нейронной сети сводится к минимизации функции ошибки путем корректировки весовых коэффициентов синаптических связей между нейронами.

$$E = \frac{1}{2p}\sum_{j=1}^{p}\left(d_j - y_j\right)^2 \tag{5}$$

здесь y_j — фактический результат, d_j — желаемый отклик сети.

Классический метод градиентного спуска в алгоритме обучения нейронной сети заключается в выборе направления и величин весов связей между нейронами для достижения минимальной ошибки аппроксимации целевой функции за минимальное число шагов обучения.

Математически это представляется следующим образом.

Пусть у нас есть некий функционал:

$$Q(W) = \sum_{j=1}^{l} L((W, X_j)Y_j) \to \min . \tag{6}$$

Вычислим градиент от него (вектор частных производных):

$$gradQ(W) = \frac{\partial Q(w_j)}{\partial w_j} . \tag{7}$$

После подстановки и преобразований имеем:

$$w^{(t+1)} = w^t - \mu \cdot gradQ(w^t) . \tag{8}$$

Обновление весов осуществляется по следующей формуле, которое представляет собой некое аддитивное выражение (сумму по всем объектам выборки):

$$w^{t+1} = w^t - \mu\sum_{j=1}^{t} L^j((w^t, x_j y_j)) . \tag{9}$$

Недостаток данного метода заключается в том, что в некоторых случаях он выдаёт локальный минимум функционала, т.е. процесс обучения не сходится. Методом стохастического градиентного спуска (SGD), можно обойти эту проблему [10]. Преимущества метода SGD заключаются в том, что он помогает избежать «застревания» модели в локальном минимуме, проще в вычислениях, поскольку позволяет не хранить все вычисления в оперативной памяти. Однако, заметим, что он не позволяет доказать существование и единственность решения. При выборке объекта случайным образом осуществляется перемешивание данных (shuffling), то есть случайно выбираем объекты, но попеременно из разных классов. Такой прием оказывает

существенное влияние на улучшение сходимости полученной модель нейронной сети. После выбора архитектуры нейросети проводятся процедуры валидации и оценки качество построенной модели нейронной сети. Схема оценивания качества обучения состояла из нескольких этапов:

1) настройка всех алгоритмов обучения,

2) сравнение качества на валидационной выборке,

3) выбор лучшего алгоритма, того, который допускает наименьшую ошибку,

4) после того, как алгоритм выбран, измерялось качество на контрольной выборке и проверим модель на адекватность.

Сначала использовалась простая валидационная процедура (simple validation) с простым разбиением генеральной выборки в соотношении 70 к 30 (70-тренировочное множество, размер обучающей выборки, 30-валидационное, измеряем качество на ней). Но после некоторого количества проведенных экспериментов обнаружилось, что текущая валидационная процедура дает очень высокий разброс оценки качества предсказания алгоритма при фиксированных настройках. Для того, чтобы справиться с этой проблемой было принято решение перейти к процедуре кроссвалидации по k блокам (k-fold cross-validation) как более объективной. Основная идея состоит в том, чтобы случайным образом разбивает данные на k непересекающихся блоков примерно одинакового размера. Далее каждый блок по очереди будет выступать в качестве тестового. Возьмем первый блок в качестве тестового, а все остальные в качестве обучения. Обучим алгоритм на обучающей выборке, измерим качество на тестовом блоке, запомним. Возьмем второй блок в качестве тестового. Все остальные сольем в обучающую выборку, обучим все, измерим качество и т.д. Каждый блок побывает один раз тестовым. Получим k оценок качества, усредним их и получим итоговую оценку модели. После проведения экспериментов с разными значениями k работе сделан вывод, что оптимальным является k= 20.

В процессе обучения проводилось варьирование или выбор гиперпараметров нейросети, что является очень важным и напрямую влияет на сходимость. Гиперпараметры — значения, которые нужно подбирать вручную и зачастую методом проб и ошибок при обучении. В качестве примеров гиперапараметров можно привести следующие:

1) Момент и скорость обучения.

2) Количество скрытых слоев.

3) Количество нейронов в каждом слое.

4) Количество эпох.

После проведения численных экспериментов были найдены оптимальные настройки гиперпараметров для построенной модели:

– оптимальная скорость обучения (learning_rate) — 0,2;

– оптимальная конфигурация нейронной сети: 3-50/tanh-20/tanh-3 tanh,

– количество эпох — $>= 200$.

– $k = 20$ (кроссвалидация по 20 блокам).

Численные эксперименты показали, что, по крайней мере, до 2000 эпохи метрика качества улучшается и достаточно быстро. С большой долей уверенности можно предположить, что улучшение качества предсказания будет продолжаться при дальнейшем увеличении количества итераций обучения. Однако численные эксперименты были остановлены ввиду высоких и увеличивающихся пропорционально количеству эпох затрат по времени. Для оценки качества работы построенной сети были проведены численные эксперименты.

При отсутствии специфических данных о наборе гиперпараметров (включая структуру нейросети) для данной задачи было принято решение начать с настроек, оптимальных для близкой (пусть и отличающейся) задачи о предсказании скорости. В экспериментах было проведено варьирование гиперпараметров и оценка разброса валидационной процедуры при фиксированном наборе гиперпараметров. В качестве разумного приближения был выбран набор параметров:

$$learning_rate = 0,2\ ,\ layers = [50 - \tanh/\ 30 - \tanh].$$

В экспериментах 7-10 по аналогии с предыдущей задачей было проведено увеличение количества эпох, давшее радикальное улучшение качества итогового алгоритма. Далее один запуск алгоритма занимает уже несколько десятков минут (2000 эпох × 20 валидаций). По некоторым экспериментам была построена аналогичная изображенной на (рис. 5) гистограмма, показывающая зависимость среднего отклонения значения глубины H от настроек нейронной сети, а также приблизительные значения дисперсии этого отклонения (рис. 4). На первых трех измерениях проиллюстрированы эксперименты с архитектурой нейронной сети при фиксированном количестве эпох ($epoch$) и скорости обучения (l_r). Дальнейшие измерения показывают вычислительно сложные эксперименты с выбранной архитектурой (рис. 5) и увеличением количества эпох ($epoch$). Из полученной гистограммы видим постепенное снижение ошибки при экспериментах с архитектурой сети и более стремительное (но при этом вычислительно затратное) при увеличении количества эпох.

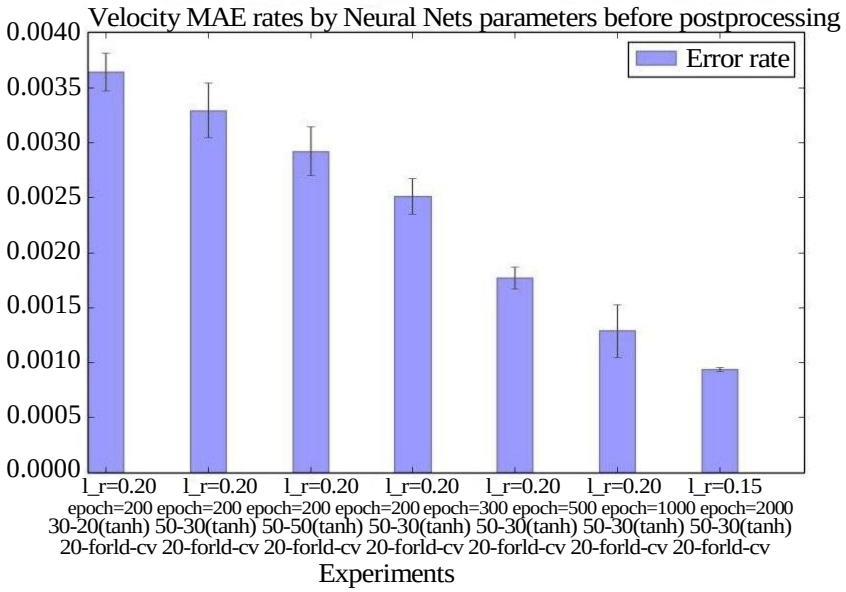


Рис. 4. Значения среднемодульных отклонений и их дисперсий, l_r — скорость обучения, $k-fold-cv$ — кроссвалидация по k блокам, $epoch$ — количество эпох, $т-n\,(\tanh)$ — m нейронов на первом скрытом слое, n — нейронов на втором скрытом слое с функцией активации гиперболический тангенс

На рис. 5 представлены графики зависимости MSE и MAE от количества эпох. Как видим, значения монотонно уменьшаются с ростом количества эпох.

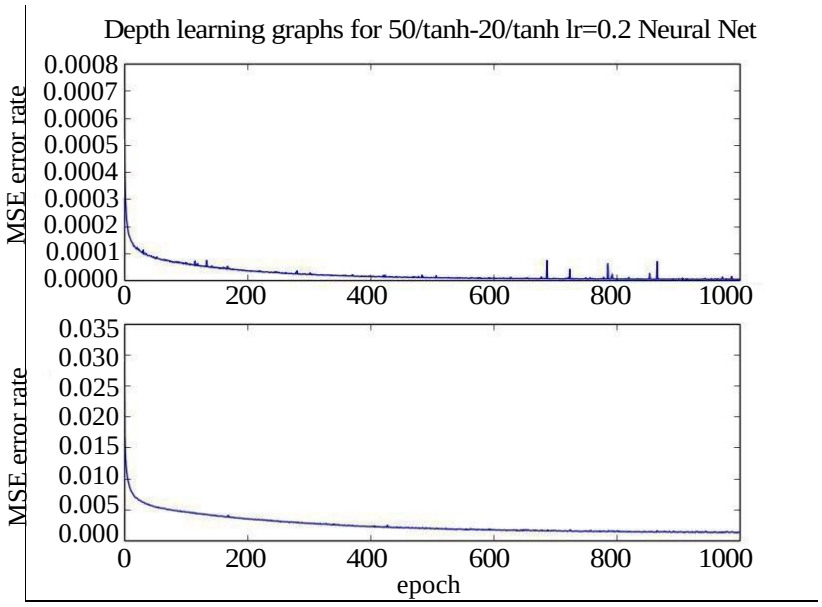


Рис. 5. Графики зависимостей MSE, MAE от количества эпох соответственно

Скорость обучения $learning_rate = 0,02$, функция активации – гиперболический тангенс.

1)однослойная нейронная сеть, 30 нейронов на первом слое с функцией активации гиперболический тангенс.

2) двухслойная нейронная сеть, 50--20 нейронов на первом и втором слое.

На рис.6 изображен график зависимости ошибки до постпроцессинга от количества нейронов в скрытом слое однослойной сети. Интервал зашумления: (0, 0.05), коэффициент зашумления: 100. Валидационная процедура — $20-fold-cv$. На графике наблюдаем не очень стабильное, но очевидное падение значения величины потерь на валидации при увеличении количества нейронов в скрытом слое. Предположительно, что при дальнейшем увеличении количества нейронов ошибка будет продолжать падать.

На рис. 7 изображен график зависимости ошибки до постпроцессинга от количества эпох. Предсказываемая величина – глубина прохождения цилиндра, интервал зашумления: $(0,\ 0.05)$, коэффициент зашумления: 100. Архитектура сети: $3-50/tanh-20/tanh-1$, коэффициент скорости обучения $learning_rate = 0.02$, валидационная процедура — $20-fold-cv$.

На рис. 8 изображен график зависимости ошибки скорости прохождения цилиндра до постпроцессинга от количества эпох. Интервал зашумления: $(0,\ 0.05)$, коэффициент зашумления: 100. Архитектура сети: $3-50/tanh-20/tanh-1$, коэффициент скорости обучения $learning_rate = 0.02$, валидационная процедура: $20-fold-cv$.

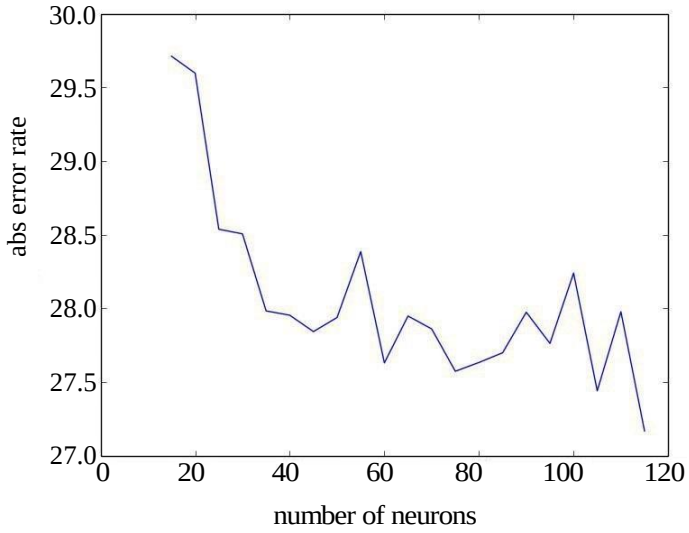


Рис. 6. График зависимости среднего модуля отклонения до постпроцессинга от количества нейронов в скрытом слое однослойной нейронной сети: *numbers of neurons* – количество нейронов,

abs error rate - средний модуль отклонения

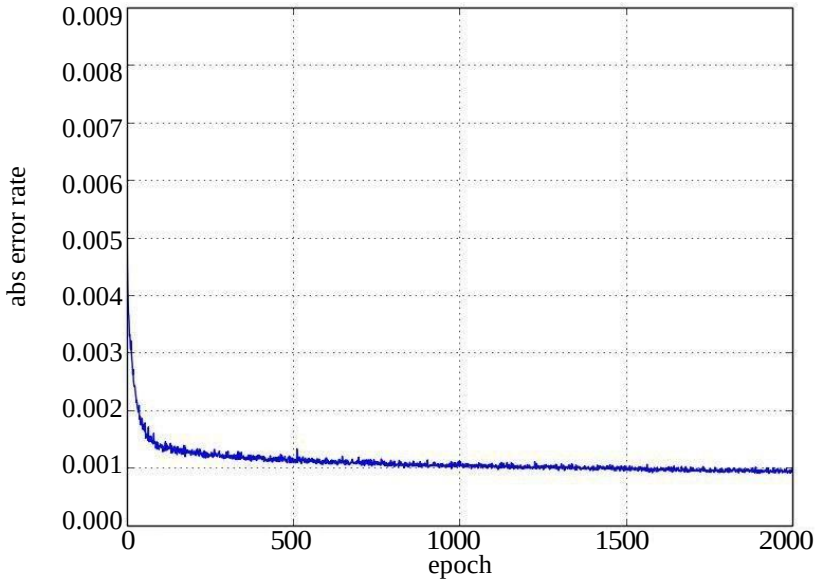


Рис. 7. График зависимости среднего модуля отклонения до постпроцессинга от количества нейронов в скрытом слое однослойной нейронной сети:
numbers of neurons — количество нейронов,
abs error rate — средний модуль отклонения.

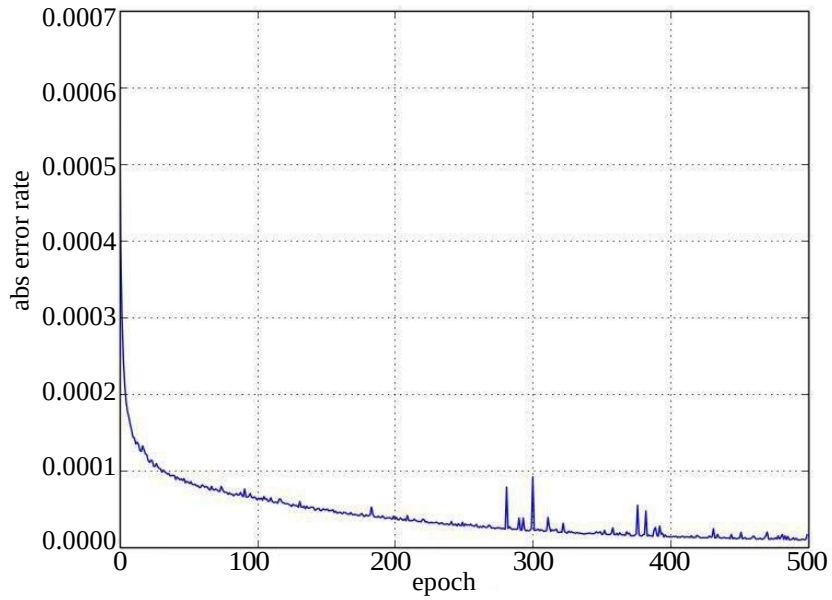


Рис. 8. График зависимости среднего модуля отклонения скорости прохождения цилиндра до постпроцессинга от количества эпох:
epoch — количество эпох обучения перед валидационной оценкой,

abs error rate — средний модуль отклонения

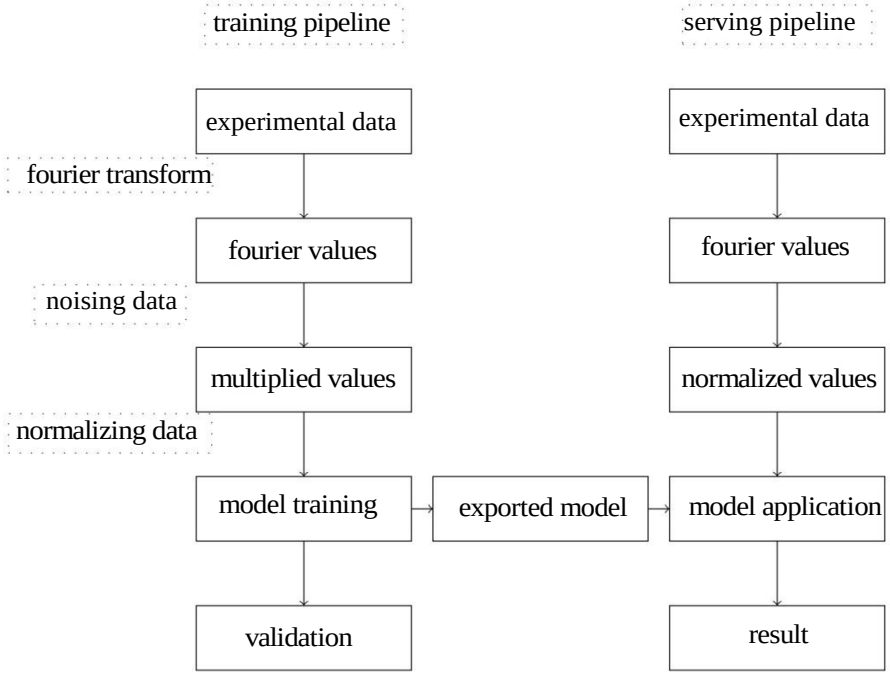


Рис. 9. Структурная схема разработанного алгоритма обучения нейронной сети для предсказывания глубины и скорости

По результатам сравнения оценок, полученных с помощью нейронной сети с данными, полученными в ходе лабораторного эксперимента установлено, установлено, что наблюдается точное совпадение качества предсказания модели обученной нейросети с результатами, полученными в ходе лабораторного эксперимента на данных, участвовавших в процессе обучения. Для корректного сравнения некоторых значения были исключены из обучающего множества (1-3). Для них также получаем результаты с достаточной точностью.

В результате исследований и разработок получен алгоритм и программный комплекс, позволяющий идентифицировать характеристики возмущающего объекта по возмущениям водной поверхности. Блок-схема этого комплекса представлена на рис. 9.

В блок-схеме приняты следующие обозначения:

- *experimental data* (экспериментальные данные) — исходные данные, состоящие из значений тангенса угла наклона поверхности воды в точках с горизонтальными координатами,
- *fourier values* (значения, получившиеся после применения спектрального анализа) — несколько младших значений преобразования Фурье, примененного к исходным данным,

- *multiplied values* (размноженные значения) — увеличенный с помощью зашумления датасет,
- *model training* (обучение модели) — процесс обучения модели на значениях, получившихся после нормирования зашумленного датасета,
- *validation* (валидация) – процесс оценки качества алгоритма с помощью одной из валидационных процедур,
- exported model (экспортированная модель) — обученная модель сохраненная на диск для дальнейшего использования,
- model application (применение модели) — загрузка с диска и применение экспортированной модели,

result (результат) — результат применения модели к новым данным (предсказание).

Структура блок-схемы состоит из двух параллельных последовательностей применений цепочки преобразований входящих данных. Одна последовательность (trainig pipeline, обучающая цепь) применяется на этапе обучения модели, ее валидации и экспорта (сохранения на диск). Вторая (serving pipeline, обслуживающая цепь) применяется на этапе использования экспортированной модели. Через первую, обучающую цепь, проходят данные всей обучающей выборки. Результатом работы этой цепи является либо оценка качества работы алгоритма машинного обучения той или иной заранее заданной валидационной процедурой, либо сохраненная на диск экспортированная модель. Вторая, обслуживающая цепь, используется для получения предсказаний на новых, неизвестных на этапе обучения, данных – как правило, в условиях применения обученной модели в продукте. Значительная часть описанных выше цепей совпадает – как по смыслу, так и на уровне программного кода. Часть, связанная с предварительной обработкой данных (от исходных до подающихся на вход алгоритму машинного обучения) совпадает буквально. Этапу тренировки модели (с последующей валидацией либо экспортом на диск) в обучающей цепи соответствует этап применения модели в обслуживающей цепи. Результатом в этом случае является предсказание модели для входящего в цепь объекта.

Завершая описание нашего подхода к решению обратной задачи генерации поверхностных волн, основанного на применении методов машинного обучения, а именно, теории нейронных сетей, заметим, что аналогичные подходы все шире проникают в различные области математической физики и технических наук, где ранее традиционно применялись классические численные методы. Например, в работе [22] методы машинного обучения применены для моделирования поля давления в потоке газа, обтекающего сферу. В работе [23] реализована модель нейронной сети для выполнения классификации эле-

ментов поверхностей летательного аппарата. Сгенерирована выборка, содержащая параметры поверхностей объектов классификации. По синтетическим данным проведено обучение нейронной сети. В работе [24] задача моделирования продольного движения самолета транспортной категории и параметрическая идентификация аэродинамических характеристик продольного движения: составляющих безразмерных коэффициентов аэродинамической подъемной силы и момента тангажа решена в классе модульных полуэмпирических динамических моделей, созданных объединением теоретического и нейросетевого моделирования.

Заключение. Предложенный нейросетевой подход позволил решить обратную задачу генерации поверхностных волн движущимися крыловым профилем и цилиндром. Построенная оптимальная модель нейронной сети и метод стохастического градиентного спуска (SGD) показали свою эффективность. Сопоставление полученных параметров крылового профиля и кругового цилиндра с данными лабораторного эксперимента подтвердило высокую точность предложенного подхода. Разработанный программный комплекс эффективно и с небольшими затратами времени решает задачу определения параметров источника возмущения по данным, снятым с водной поверхности. Эффективность методов машинного обучения, показанная в рассмотренной задаче, дает основания полагать, что они применимы и к решению других обратных задач гидрофизики и описанные подходы могут быть распространены на более широкий класс некорректно поставленных задач.

ЛИТЕРАТУРА

[1] Нестерова С.В., Шамаева А.С., Шамаева С.И. *Методы, процедуры и средства аэрокосмической компьютерной радиотомографии приповерхностных областей Земли*. Москва, Научный мир, 1996, 272 с.

[2] Савин А.С. Определение параметров гидродинамических особенностей в плоском потоке по данным о его свободной поверхности. *Известия РАН. Механика жидкости и газа*, 2001, № 2, с. 139–146.

[3] Тихонов А.Н., Арсенин В.Я. *Методы решения некорректных задач*. Москва, Наука, 1979, 286 с.

[4] Коновалов А.В., Левченко Е.С., Савин А.С. Восстановление плоского течения тяжелой идеальной жидкости по форме ее свободной поверхностию. *Докл. АН СССР*, 1989, т. 305, № 2, с. 294 – 296.

[5] Исиченко И.В., Коновалов А.В., Левченко Е.С., Савин А.С. Обратная задача обтекания особенностей плоским потоком идеальной жидкости со свободной границей. *Прикладная механика и техническая физика*, 1989, № 6, с. 86 – 91.

[6] Voronin E. A., Nosov V. N., Savin A. S. Neural network approach to solving the inverse problem of surface-waves generation. *Journal of Physics: Conference Series*. 2019, 1392, vol. 1, art. 012022. DOI: 10.1088/1742-6596/1392/1/012022.

[7] Воронин Е.А., Носов В.Н., Савин А.С. Определение параметров

погруженного источника по возмущениям поверхности жидкости на основе методов машинного обучения. *Доклады РАН. Науки о Земле*, 2020, т. 493, № 1, с. 103-106.

[8] Агарвал Ч. *Нейронные сети и глубокое обучение: учебный курс*. СПб, ООО «Диалектика», 2020, 752 с.

[9] Просиз Дж. *Прикладное машинное обучение и искусственный интеллект для инженеров*. Астана, АЛИСТ, 2024, 432 с.

[10] Хайкин С. *Нейронные сети: полный курс. 2-е изд*. Москва, Издательский дом «Вильямс», 2006, 1104 с.

[11] Manning C., Raghavan P., Schutze H. *Introduction to Information Retrieval*. Cambridge University Press, 2009, 544 p.

[12] Kuhn M., Johnson K. *Applied predictive modeling*. New York, Springer, 2013, 600 p.

[13] Nathan M., James W. *Big Data: Principles and best practices of scalable realtime data systems*. Manning Publications, 2015, 308 p.

[14] Рашка С. *Python и машинное обучение*. Москва, ДМК-Пресс, 2017, с. 418.

[15] Серрано Л.. *Грокаем машинное обучение*. Санкт-Петербург, Питер, 2024, 512 с.

[16] Мэрфи К.П. *Вероятностное машинное обучение*. Москва, ДМК Пресс, 2022, 770 с.

[17] Постолит А.В. *Основы искусственного интеллекта в примерах на Python. Самоучитель. 2-е изд*. Санкт-Петербург, БХВ-Петербург, 2024, 448 с.

[18] Бояринцев В.И., Леднев А. К., Фрост В.А. *Движение погруженного цилиндра под поверхностью жидкости*. *Препринт*. Москва, Институт проблем механики АН СССР, 1988, № 332, 39 с.

[19] Бояринцев В.И., Леднев А.К., Прудников А.С., Савин А.С., Савина Е.О. *Моделирование и экспериментальное исследование возмущений свободной границы плоского потока погруженными источниками*. *Препринт*. Москва, Институт проблем механики РАН, 2002, № 720, 37 с.

[20] Бояринцев В.И., Леднев А.К., Прудников А.С., Савин А.С., Савина Е.О. Возмущение свободной поверхности жидкости крыловым профилем. *Известия РАН. Механика жидкости и газа*, 2004, № 6, с. 145- 152.

[21] Бармин А.А., Бояринцев В.И., Леднев А.К., Савин А.С., Савина Е.О. *Моделирование и экспериментальное исследование возмущений свободной поверхности жидкости шаром и эллипсоидом. Препринт*. Москва, Институт проблем механики РАН, 2004, № 763, 43 с.

[22] Котенев В.П., Пучков А.С., Сапожников Д.А., Тонких Е.Г. Применение методов машинного обучения для моделирования распределения давления в возмущенной области около сферы, обтекаемой невязким потоком. *Математическое моделирование и численные методы,* 2017, № 4, с. 60–72.

[23] Булгаков В.Н., Рацлав Р.А., Сапожников Д.А., Чернышев И.В. Моделирование нейронной сети для решения задачи классификации элементов корпуса летательного аппарата. *Математическое моделирование и численные методы*, 2018, № 4, с. 57–71.

[24] Крееренко С.С., Крееренко О.Д. Моделирование и параметрическая идентификация аэродинамических характеристик самолета транспортной категории с использованием нейросетей в среде Tensorflow. *Математическое моделирование и численные методы*, 2024, № 3, с. 81–99.

Статья поступила в редакцию 15.05.2025

Ссылку на эту статью просим оформлять следующим образом:

Воронин Е.А., Носов В.Н., Савин А.С. Нейросетевой подход к решению обратной задачи генерации поверхностных волн. *Математическое моделирование и численные методы*, 2025, № 4, с. 173–193.

Воронин Евгений Алексеевич — д-р техн. наук, ведущий научный сотрудник ФИЦ ИУ РАН. e-mail: e.voronin1@gmail.com

Носов Виктор Николаевич — д-р техн. наук, заведующий лабораторией ГЕОХИ РАН. e-mail: viktor_nosov@mail.ru

Савин Александр Сергеевич — д-р физ.-мат. наук, профессор кафедры «Высшая математика» МГТУ им. Н.Э. Баумана. e-mail: assavin@list.ru

Neural network approach to solving the inverse problem of surface wave generation

© A.A. Valishin[1], A.V. Zaprivoda[2], S.S. Tsukhlo[1]

[1] FRC CSC RAS, Moscow, 119333, Russia
[2]Vernadsky Institute of Geochemistry and Analytical Chemistry RAS, Moscow, 119991, Russia
[3]Bauman Moscow State Technical University, Moscow, 105005, Russia

The article considers the neural network approach to solving the inverse problem of generating surface waves or identifying a disturbance source based on data taken from the free surface of a liquid. The type is selected and a mathematical and computer model of the neural network is constructed based on data obtained during a laboratory experiment to study surface waves that occur when an airfoil and a circular cylinder move in a liquid. To assess the adequacy of this model, numerical experiments are conducted to determine the depth and velocity of the disturbance source in the liquid. The calculation results are compared with the experimental data. A software package has been developed that allows one to effectively and with minimal time expenditure solve the problem of determining the parameters of the disturbance source based on data taken from the water surface.

Keywords*: inverse problem of surface wave generation; neural networks; neural network training; stochastic gradient descent method; neural network state monitoring; spectral analysis; backpropagation algorithm; regression problem; feedforward networks; multilayer perceptron*

REFERENCES

[1] Nesterova S.V., SHamaeva A.S., SHamaeva S.I. *Metody, procedury i sredstva aerokosmicheskoj komp'yuternoj radiotomografii pripoverhnostnyh oblastej Zemli* [Methods, procedures and means of aerospace computed radio tomography of the near-surface regions of the Earth.]. Moscow, «Nauchny Mir» Publishers, 1996, 272 p.

[2] Savin A.S. Opredelenie parametrov gidrodinamicheskih osobennostej v ploskom potoke po dannym o ego svobodnoj poverhnosti [Determination of the parameters of hydrodynamic features in a flat flow from data on its free surface]. *Fluid Dynamics*, 2001, no. 2, pp. 139–146.

[3] Tihonov A.N., Arsenin V.YA. *Metody resheniya nekorrektnyh zadach* [Methods for solving ill-posed problems]. Moscow, Nauka Publ., 1979, 286 p.

[4] Konovalov A.V., Levchenko E.S., Savin A.S. Vosstanovlenie ploskogo techeniya tyazheloj ideal'noj zhidkosti po forme ee svobodnoj poverhnostiyu [Reconstruction of the plane flow of a heavy ideal fluid from the shape of its free surface]. *Doklady Akademii Nauk SSSR* [Proceedings of the USSR Academy of Sciences], 1989, vol. 305, no. 2, pp. 294 – 296.

[5] Isichenko I.V., Konovalov A.V., Levchenko E.S., Savin A.S. Obratnaya zadacha obtekaniya osobennostej ploskim potokom ideal'noj zhidkosti so svobodnoj granicej [The inverse problem of a plane flow of an ideal fluid with a free boundary past singularities]. *Journal of Applied Mechanics and Technical Physics*, 1989, no. 6, pp. 86 – 91.

[6] Voronin E. A., Nosov V. N., Savin A. S. Neural network approach to solving the inverse problem of surface-waves generation. *Journal of Physics: Conference Series*. 2019, 1392, vol. 1, art. 012022. DOI: 10.1088/1742-6596/1392/1/012022.

[7] Voronin E.A., Nosov V.N., Savin A.S. Determination of submerged source parameters from liquid surface disturbances based on machine learning methods. *Doklady Earth Sciences*, 2020, vol. 493, no. 1, pp. 103-106.

[8] Agarval CH. *Nejronnye seti i glubokoe obuchenie: uchebnyj kurs* [Neural networks and deep learning: training course]. St. Petersburg, Dialektika [Dialectics LLC], 2020, 752 p.

[9] Prosiz Dzh. *Prikladnoe mashinnoe obuchenie i iskusstvennyj intellekt dlya inzhenerov* [Applied machine learning and artificial intelligence for engineers]. Astana, Alist, 2024, 432 p.

[10] Khaikin S. *Neural networks: a complete course, 2nd ed*. Moscow, Williams Publishing House, 2006, 1104 p.

[11] Manning C., Raghavan P., Schutze H. *Introduction to Information Retrieval*. Cambridge University Press, 2009, 544 p.

[12] Kuhn M., Johnson K. *Applied predictive modeling*. New York, Springer, 2013, 600 p.

[13] Nathan M., James W. *Big Data: Principles and best practices of scalable realtime data systems*. Manning Publications, 2015, 308 p.

[14] Raschka S. *Python and Machine Learning*. Birmingham, Packt Publishing, 2017, p. 418.

[15] Serrano L. *Grokking Machine Learning* [Grokaem machine learning]. St. Petersburg, Piter [Piter Publishing House], 2024, 512 p.

[16] Murphy K.P. *Probabilistic Machine Learning: Advanced Topic*. Cambridge, MITP Press, 2022, 770 p.

[17] Postolit A.V. *Osnovy iskusstvennogo intellekta v primerah na Python. Samouchitel'. 2-e izd* [Basics of artificial intelligence with examples in Python. Self-instruction manual. 2nd ed.]. St. Petersburg, BHV-Petersburg, 2024, 448 p.

[18] Boyarincev V.I., Lednev A. K., Frost V.A. *Dvizhenie pogruzhennogo cilindra pod poverhnost'yu zhidkosti. Preprint* [Movement of a submerged cylinder under the surface of a liquid. Preprint]. Moscow, Institut problem mekhaniki AN SSSR [Institute of Mechanical Problems of the USSR Academy of Sciences], 1988, no. 332, 39 p.

[19] Boyarincev V.I., Lednev A.K., Prudnikov A.S., Savin A.S., Savina E.O. *Modelirovanie i eksperimental'noe issledovanie vozmushchenij svobodnoj granicy ploskogo potoka pogruzhennymi istochnikami. Preprint* [Modeling and experimental study of disturbances of the free boundary of a plane flow by immersed sources. Preprint]. Moscow, Institut problem mekhaniki AN SSSR [Institute of Mechanical Problems of the USSR Academy of Sciences], 2002, no. 720, 37 p.

[20] Boyarincev V.I., Lednev A.K., Prudnikov A.S., Savin A.S., Savina E.O. Vozmushchenie svobodnoj poverhnosti zhidkosti krylovym profilem [Disturbance of the free surface of a liquid by an airfoil]. *Fluid Dynamics*, 2004, no. 6, pp. 145- 152.

[21] Barmin A.A., Boyarincev V.I., Lednev A.K., Savin A.S., Savina E.O. *Modelirovanie i eksperimental'noe issledovanie vozmushchenij svobodnoj poverhnosti zhidkosti sharom i ellipsoidom. Preprint* [Modeling and experimental study of perturbations of the free surface of a liquid by a ball and an ellipsoid. Preprint]. Moscow, Institut problem mekhaniki AN SSSR [Institute of Mechanical Problems of the USSR Academy of Sciences], 2004, no. 763, 43 p.

[22] Kotenev V.P., Puchkov A.S., Sapozhnikov D.A., Tonkikh E.G. imulation of the pressure distribution in the disturbed region near the sphere streamlined by the inviscid flotation by means of the machine learning methods. *Mathematical Modeling and Computational Methods,* 2017, no. 4, pp. 60–72.

[23] Bulgakov V.N., Ratslav R.A., Sapozhnikov D.A., Chernyshev I.V. Modeling a neural network to solve the problem of classifying air frame elements. *Mathematical Modeling and Computational Methods*, 2024, no. 3, pp. 81–99.

[24] Kreerenko S.S., Kreerenko S.S. Parametric identification of aerodynamic characteristics of a transport category aircraft using recurrent semi-empirical neural networks in the Tensorflow environment. *Mathematical Modeling and Computational Methods*, 2024, no. 3, pp. 81–99.

Voronin E.A, Dr. Sc. (Eng.), Leading Researcher, Federal Research Center «Computer Science and Control» of the Russian Academy of Sciences. e-mail: e.voronin1@gmail.com

Nosov V.N., Dr. Sc. (Eng.), Head of the Laboratory, Vernadsky Institute of Geochemistry and Analytical Chemistry of the Russian Academy of Sciences. e-mail: viktor_nosov@mail.ru

Savin A.S., Dr. Sc. (Phys.-Math.), Professor, Department of Higher Mathematics, Bauman Moscow State Technical University. e-mail: assavin@list.ru