**Прогноз плотности вероятности ветроэнергетики на основе многоядерной нейронной сети**

*Аннотация*

Масштабное применение энергии ветра создает серьезные проблемы для диспетчеризации энергосистем. В настоящее время метод прогнозирования плотности вероятности, сочетающий нейронную сеть квантильной регрессии (НСКР) с функцией ядра Епанечникова, является отличным алгоритмом для прогнозирования ветровой энергии, который может дать комплексное распределение вероятностей будущей ветровой энергии и эффективно оценить неопределенность генерации ветровой энергии. Однако существующие методы прогнозирования плотности вероятности обрабатывают данные последовательно в разных квантилях, и затраты времени на вычисления умножаются с увеличением обучающих данных. Это влияет на практичность метода прогнозирования плотности вероятности. Чтобы решить эту проблему, в этой работе предлагается многоядерная нейронная сеть с параллельной квантильной регрессией (МНСПКР), основанная на параллельной модели «начальник-подчиненный». Алгоритм делит сложные задачи прогнозирования во всех квантилях на несколько параллельных подзадач, которые независимо выполняются на разных ядрах, так что преимущества производительности многоядерного ЦП могут быть полностью использованы для повышения вычислительной эффективности модели совместной работы. Влияние различного количества ядер ЦП на параллельную производительность МНСПКР анализируется с учетом алгоритмической природы ускорения и параллельной эффективности. Чтобы продемонстрировать эффективность предложенной модели, на наборах данных проводятся сравнительные эксперименты других четырех традиционных моделей. Результаты моделирования показывают, что МНСПКР может не только повысить эффективность обучения НСКР, но также получить точные результаты прогнозирования энергии ветра, демонстрируя потенциальную ценность и полезность для сложной энергосистемы.

**1. Введение**

В последние несколько лет из-за увеличения теплового загрязнения и выбросов парниковых газов, вызванных традиционным производством энергии на ископаемом топливе, возобновляемая энергия стала отличным выбором для производства электроэнергии и ключевым компонентом государственных стратегий развития [1]. Среди возобновляемых источников энергии все большее значение приобретают ветровая и солнечная генерация из-за ее универсальности и относительно низкой себестоимости производства [2]. Энергия ветра — это чистый, незагрязненный возобновляемый источник энергии с огромными запасами, обладающий потенциалом для последовательного и устойчивого производства энергии из возобновляемых источников. Она стал одним из самых популярных возобновляемых источников энергии в мире благодаря своей широкой доступности [3]. Благодаря резкому снижению стоимости технологий хранения энергии, к 2050 году почти половина мировой электроэнергии будет производиться за счет этих двух быстро развивающихся возобновляемых источников энергии [4]. Без тени сомнения ожидается, что ветроэнергетика будет интенсивно развиваться в течение длительного времени. Однако одним из основных препятствий для производства ветровой энергии является то, что неопределенность процесса ветрорегенерации приводит к непланируемой загруженности производственных мощностей ветряных электростанций, что представляет собой огромную проблему для диспетчеризации сети и торговли электроэнергией [5]. Кроме того, опыт эксплуатации ветроэнергетики в различных странах показывает, что широкомасштабный доступ к ветроэнергетике может вызвать ряд проблем безопасности и стабильности. Комплексное и точное прогнозирование энергии ветра способствует раннему планированию, оптимизации и планирования нагрузки. Это незаменимый инструмент для повышения стабильности энергосистемы и экономической выгоды от производства ветровой энергии [6]. В последние годы было проведено множество исследований, посвященных моделям прогнозирования энергии ветра, таким как нейронная сеть обратного распространения ошибки [7], нечеткая нейронная сеть [8], вейвлет-анализ [9], метод опорных векторов [10] и регрессионный анализ [11], обеспечивающие надежный прогноз. Среди них метод обратного распределения ошибки обладает весомыми способностями к нелинейному отображению и обобщению, но склонен к медленной сходимости и локальной оптимизации; нейтронная сеть прямого распространения имеет большие преимущества при решении нелинейных и нечетких задач благодаря органичному сочетанию нечеткой модели [12,13], но скорость сходимости модели низкая, а способность глобальной аппроксимации плохая; вейвлет-анализ обладает хорошей аппроксимацией и отказоустойчивостью, но выбор вейвлет-базисной функции и инициализация параметров не основаны на определенных критериях; метод опорных векторов может устранить множество избыточных выборок с хорошей надежностью, но сложно обучать крупномасштабные выборки и решать проблемы мультиклассификации; Метод регрессионного анализа имеет высокую скорость обучения и небольшую частоту ошибок при небольшом объеме данных, но он недоступен для обработки большого количества данных. Эти алгоритмы прогнозирования энергии ветра не могут удовлетворить все более сложные требования прогнозирования энергии ветра из-за ограничения точности и сферы применения.

Вышеупомянутые модели делают упор на детерминистические точечные прогнозы энергии ветра. Однако они не способны выявить неопределенность характеристик ветроэнергетики. Напротив, вероятностное прогнозирование энергии ветра в последние годы приобрело большую популярность, поскольку оно способно обеспечить точное детерминистическое предсказание, количественно оценить неопределенность выходной мощности ветра и построить непрерывную кривую плотности вероятности прогнозируемой энергии [14]. Это более важно для энергетических компаний и системных операторов, чем точечные прогнозы. В исследованиях вероятностного прогнозирования энергии ветра квантильная регрессия является одним из известных методов регрессионного анализа, часто используемых в статистике и эконометрике. Не задумываясь о характере распределения случайных величин, с помощью квантильной регрессии можно получить общие характеристики распределения переменной отклика с помощью указанной соответствующей функции регрессии, которая более надежна при обработке выбросов в интерпретируемых измерениях [15]. Поэтому квантильная регрессия широко применяется в прогнозировании ветроэнергетики [16]. Чтобы исчерпывающе выявить нелинейную связь между зависимой и независимой переменной, Тейлор [18] сначала предложил модель НСКР, которая представляет собой метод непараметрической композиции искусственной нейронной сети (ИНС) [19] и квантильной регрессии. Нелинейный метод сочетает в себе преимущества обоих. Прадипкумар и др. [20] спрогнозировали волатильность финансовых временных рядов, используя алгоритм оптимизации роя частиц для оптимизации обученной НСКР. Кэннон [15] продемонстрировал потенциальное превосходство модели НСКР для задачи уменьшения масштаба осадков. Авторы [16] использовали гибридный метод, основанный на нейронной сети квантильной регрессии и функции ядра Епанечникова с применением несмещенной перекрестной проверки, для выполнения прогноза плотности вероятности, который достоверно характеризовал неопределенность выработки ветровой энергии. Эти эмпирические исследования еще раз демонстрируют эффективность подхода НСКР. НСКР не только демонстрирует все условное распределение объясняемой переменной, но и решает сложные нелинейные задачи. Применяя метод НСКР для прогнозирования энергии ветра, можно получить конкретные условные квантили энергии ветра. Однако у него есть ключевой недостаток. Механизм агрегации НСКР сложен, а вычислительная сложность его обучения увеличивается экспоненциально из-за традиционной структуры нейронной сети. Кроме того, интеграция информатизации и индустриализации в энергетике способствовала быстрому росту данных о ветроэнергетике, что неизбежно усложняло прогнозирование в ветроэнергетики и создавало серьезные проблемы для эффективности расчетов модели [21]. Упомянутый выше алгоритм НСКР может обеспечить точные прогнозы мощности ветра, но затраты времени вычислений резко возрастают с увеличением размера данных и сложности модели. В частности, когда масштаб ветроэнергетической системы достигает определенного уровня, затраты времени на расчеты становятся абсолютно невыносимым. Хотя многие исследователи за прошедшие годы оптимизировали модель НСКР и получили хорошие результаты в некоторых аспектах, все еще существуют некоторые недостатки, такие как низкая скорость обучения. Эти усовершенствованные алгоритмы до сих пор не могут удовлетворить требованиям практического применения. В целом существует два основных способа повышения эффективности вычислений. Один из них — улучшить классический алгоритм или найти новые. Другой – использовать новые компьютерные технологии, включая аппаратное и программное обеспечение. Поэтому большое количество исследований было посвящено эффективному использованию платформ высокопроизводительных вычислений для повышения эффективности работы моделей [22]. Эта работа будет нацелена на последнее, и наше внимание сосредоточено на распараллеливании метода НСКР, особенно на том, как максимально эффективно использовать многоядерные ресурсы. Многоядерные процессоры также известны как многоядерные процессоры или однокристальные мультипроцессоры [23]. Многоядерный процессор означает объединение двух или более неповрежденных вычислительных механизмов (ядер) в один процессор, каждый из которых поддерживает отдельный поток. В 1996 году Стэнфордский университет впервые предложил идею многоядерного процессора. Затем, в 2001 году, IBM представила первый коммерческий многоядерный процессор POWER4, после чего в 2005 году последовало крупномасштабное применение многоядерных процессоров Intel и AMD [24]. До сих пор многоядерные процессоры развивались более десяти лет и стали основным направлением рынка. В этом процессе диапазон применения многоядерного процессора охватывает множество областей, таких как мультимедийные вычисления, встроенные устройства, персональные компьютеры, коммерческие серверы и высокопроизводительные компьютеры.

Кроме того, растущая популярность многоядерных процессоров обеспечивает незаменимую аппаратную основу для выполнения вычислительных задач с применением параллельных механизмов. Технология высокопроизводительных параллельных вычислений привлекает все большее внимание во всем мире и постепенно становится решающим путем повышения эффективности вычислений [25]. Он широко использовался в научных исследованиях и инженерных технологиях, таких как прогнозирование климата, сейсмический анализ и другие сложные задачи [26]. Параллельные вычисления в многоядерной среде означают разложение всей традиционной задачи последовательной обработки на несколько подзадач с последующим распределением подзадач по нескольким различным ядрам для обработки [27]. В ядрах подзадачи могут выполняться независимо, чтобы ускорить процесс вычислений. Основная цель параллельных вычислений: одна — обеспечить более высокую скорость вычислений, чем у традиционного компьютера; другая — решать крупномасштабные задачи, которые трудно решить традиционным компьютерам. По сравнению с традиционными последовательными вычислениями, многоядерные вычисления позволяют осуществлять параллельные вычисления на обычных машинах, экономя большое количество аппаратных и сетевых ресурсов и повышая эффективность вычислений. Таким образом, полное применение преимуществ многоядерных параллельных вычислений является жизненно важным средством устранения узкого места в скорости последовательной обработки и реализации сложных, многомерных и ресурсоемких задач. Сегодня повышение эффективности вычислений за счет распараллеливания успешно доказано во многих областях. В [28] предложен новый параллельный алгоритм искусственной пчелиной семьи для эффективного решения задачи выбора маршрута транспортного средства. Ченг и др. [24] представили алгоритм мелкозернистого параллельного дискретного дифференциального динамического программирования, основанный на параллельной структуре в многоядерной среде, для повышения эффективности вычислений для долгосрочной эксплуатации гидроэнергетических систем с несколькими водохранилищами. В [29] Брайан использовал методы высокопроизводительных вычислений для комплексной пространственно-временной оценки и моделирования сложных социально-экологических систем. В [30] моделирование с высоким разрешением производства пшеницы, динамики почвенного углерода и азота в выращиваемых районах Австралии было выполнено путем сочетания параллельной обработки и сетевых вычислений. Хуо и др. [31] предложили алгоритм многоядерной параллельной искусственной роевой оптимизации (MPABC), основанный на параллельной иерархической модели и структуре, для оптимизации гидрологической модели Синаньцзяна. Для удовлетворения требований эффективности алгоритма прогрессивной оптимизации (POA) при решении задач планирования гидроэнергетики в [32] представлен новый эффективный параллельный алгоритм прогрессивной оптимизации. Ша и др. [33] предложили новую глубокую параллельную остаточную сеть, которая является эффективным алгоритмом для решения проблемы сверхразрешения одиночного изображения. Мэдсен и др. [34] предложили новый метод распараллеливания обучения структуры байесовской сети, который может значительно улучшить временные характеристики. Масудд и др. [35] предложили параллельный алгоритм на основе графического процессора для расчета альфа-комплексов, и вычислительные эксперименты на нескольких биомолекулах доказали его превосходную производительность. Поэтому исследование параллельных алгоритмов также имеет кардинальное значение для решения трудоемких задач крупных ветроэнергетических систем.

В этой работе представлена параллельная модель НСКР «ведущий-подчиненный» (далее МНСПКР). Параллельный алгоритм НСКР, основанный на разделении задач прогнозирования, в основном использует параллельный режим «главный-подчиненный» и одновременно сочетает в себе характеристики структуры и механизм работы НСКР. Хотя инфраструктура параллельного не является наиболее эффективным методом обеспечения хорошей параллельной производительности, она имеет явные преимущества для привлекательного выбора. Во-первых, модель может рекурсивно разлагать задачу на множество подзадач, используя стратегию «разделяй и властвуй», которая применима для задач с интенсивными вычислениями. Во-вторых, параллельная платформа позволяет максимально эффективно использовать многоядерные ресурсы, что делает ее широко распространенным приложением на компьютерах с несколькими ядрами. Наконец, модель является одним из наиболее эффективных и простых в реализации методов параллельных вычислений, поэтому гораздо проще разрабатывать параллельные алгоритмы для конкретной задачи. Джин [36] использовал модель в сочетании с алгоритмом оптимизации роя частиц для оптимизации конструкции антенны. В [37] исследователи подробно описали распараллеливание алгоритма с помощью модели «начальник-подчинённый» для решения задачи идентификации биомеханической системы среднего масштаба. В [38] была представлена параллельная модель кооперативного метода, основанная на разложении исходного пространства поиска на подпространства меньшей размерности. Тан и др. [39] применили программы и многопопуляционного распараллеливания Эпсилон-недоминируемого сортированного генетического алгоритма-II к проблемам водных ресурсов. Ли и др. [40] предложили модель параллельного динамического программирования в системе HPC и применили метод, основанный на знаниях, для определения выработки электроэнергии в рамках проекта «Три ущелья» и каскадной гидроэлектростанции проекта «Гечжоуба» в Китае. Чжао и др. [41] предложили метод параллельных вычислений с применением глубоких сетей доверия и использовали его для прогнозирования транспортного потока. Хотя параллельные вычисления широко используются в других моделях предметной области, они не были полностью исследованы на территории операций по прогнозированию ветроэнергетики, и лишь немногие статьи в технической литературе рассматривают применение параллельного алгоритма на территории прогнозирования ветроэнергетики.

Таким образом, распараллеливание алгоритма НСКР все еще находится в зачаточном состоянии, а применение параллельной структуры для оптимизации модели прогнозирования энергии ветра встречается относительно редко. Насколько нам известно, это первая параллельная реализация на основе НСКР. Осуществимость и полезность представленного алгоритма проверены на архивных данных о ветровой энергии. Тематические исследования показывают, что, используя метод МНСПКР, пользователи могут легко добиться более высокой параллельной эффективности и полностью обеспечить точность исходного предсказания НСКР. По сравнению с недавними работами авторов [16,42,43], в этой работе основное внимание уделяется эффективности прогнозирования плотности вероятности ветровой энергии. Модель МНСПКР создается путем объединения параллельной модели и НСКР. Параллельная обработка является жизнеспособным подходом к повышению эффективности вычислений для прогнозирования энергии ветра и очень полезным методом в будущем, особенно с учетом быстрого роста и масштабирования ветроэнергетических систем и постоянно растущего уровня компьютерного оборудования и программного обеспечения.

**2. Математическое обоснование**

*2.1. Квантильная регрессия*

Квантильная регрессия, предложенная в [18] в 1978 году является дополнением и расширением обычного метода наименьших квадратов. Надежность квантиля может дополнительно гарантировать надежность квантильной регрессией, что дополнительно компенсирует недостатки обычного метода наименьших квадратов при обработке выбросов. Учитывая выборочный набор  для обучения КР-модель формулируется следующим образом:

|  |  |
| --- | --- |
| , ; | (1) |

где — независимая переменная с  компонентами,  — -е значение ,  — соответствующая зависимая переменная,  представляет -й условный квантиль зависимых переменных  при независимых переменных , и для любых квантиль  и  – вектор коэффициентов регрессии в зависимости от -го условного квантиля. Оценочное значение  в КР-модели можно получить путем оптимизации следующего уравнения.

|  |  |
| --- | --- |
|  | (2) |

где  — размер выборки, а  — асимметричная функция потерь в -м условном квантиле. Потери изображаются кусочно-линейной функцией, показанной следующим образом:

|  |  |
| --- | --- |
|  | (3) |

в которой  — индикаторная функция, иллюстрируемая следующим образом:

|  |  |
| --- | --- |
|  | (4) |

С помощью индикаторной функции остаточным абсолютным значениям присваиваются разные веса. Затем оценочные значения  можно получить путем корректировки квантиля

*2.2. Нейронная сеть квантильной регрессии*

Традиционная КР — это всего лишь простая линейная формулировка, но в действительности взаимосвязь между объясняющими переменными и переменными отклика, скорее всего, будет проявляться как нелинейные характеристики. Тейлор впервые предложил метод нейронной сети квантильной регрессии (НСКР) в 2000 году, который представляет собой непараметрический метод КР-моделирования, основанный на сети прямого распространения с одним скрытым слоем. Этот метод сочетает в себе преимущества Кр с нейронными сетями и демонстрирует высокую производительность. Здесь мы можем получить модель НСКР на основе условного квантиля , оцениваемого следующим образом:

|  |  |
| --- | --- |
|  | (5) |

в котором  обозначает оцененные веса между входным и скрытым слоями нейронов доступен весовой вектор  для соединения скрытого и выходного слоев нейронов, а  — количество скрытых узлов.  — нелинейная функция, состоящая из весов  и . Функция активации скрытого слоя  обычно представляется сигмовидной функцией , а функцию активации выходного слоя  традиционно представляют линейной функцией [44]. Решив следующую задачу программирования, значения оценки параметров  и  в построенной модели НСКР можно получить по формуле

|  |  |
| --- | --- |
|  | (6) |

где  — функция потерь, определенная, как показано в уравнении. (3),  и  — штрафные параметры модели, которые могут эффективно предотвратить переобучение модели. Наилучшие оценки  и  получаются с помощью метода градиентной оптимизации, подставленного в уравнение (5)

|  |  |
| --- | --- |
|  | (7) |

где  и  — оценочные значения  и .

*2.3. Параллельная модель «начальник-подчиненный»*

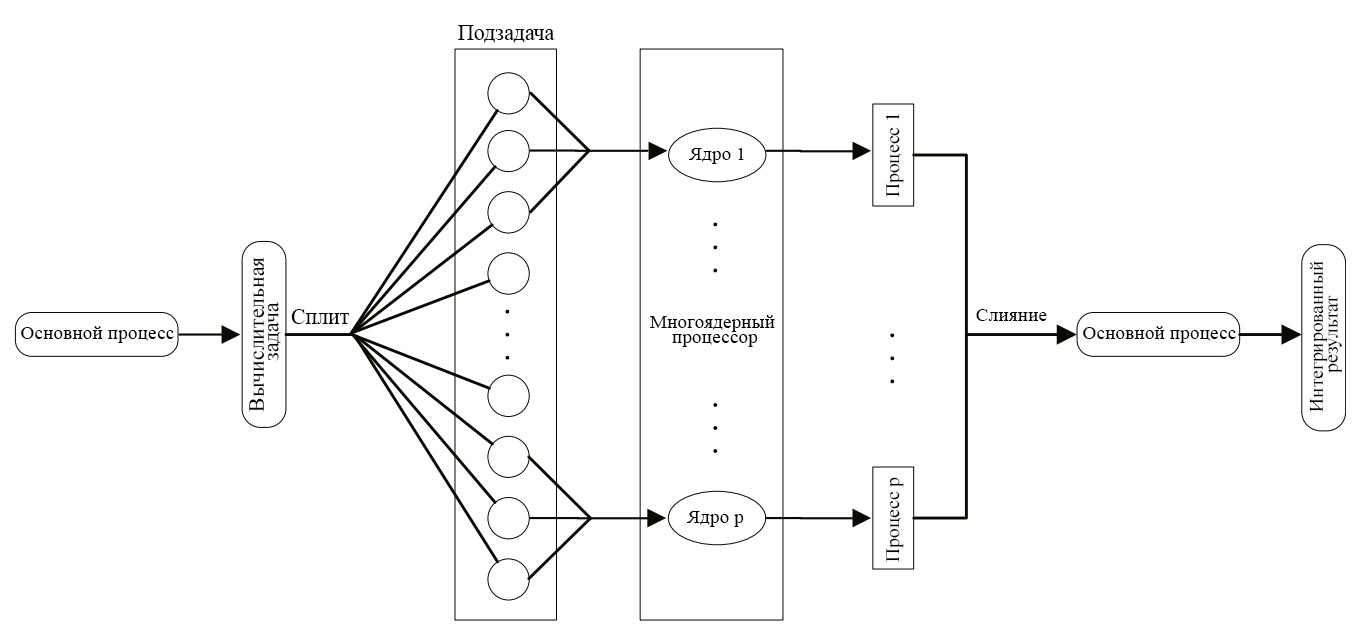
В настоящее время модель «начальник-подчиненный» является одним из наиболее распространенных методов параллельных вычислений, а также одним из самых простых и понятных способов реализации глобального параллелизма. Его популярность в основном объясняется прямым применения возможностей распараллеливания современных компьютерных систем и простотой реализации. Для разработки устоявшихся вариантов параллельного алгоритма «начальник-подчиненный» обычно требуется лишь небольшой объем работы по программированию и фундаментальные знания гомологичной компьютерной системы [38]. Кроме того, требуются лишь незначительные изменения в структуре алгоритма. По этим причинам модель «начальник-подчиненный» часто считается первым прорывом в преобразовании алгоритма от последовательного проектирования к параллельному. Модель «начальник-подчиненный» реализует параллельную обработку задачи в три этапа: разделение задач, параллельные вычисления и интеграцию результатов. Главный процесс генерирует и разделяет сложную работу на меньший набор независимых подзадач, а затем передает множество независимых подзадач и данные для расчета на различные ядра. Каждое ядро создает отдельный подчиненный процесс для одновременного выполнения подзадач, и между подчиненным процессом и ядром существует взаимно однозначное соответствие. После того как все рабочие процессы завершат выполнение, параллельная обработка прекращается, и результаты возвращаются главному процессу. Главный процесс отвечает за объединение результатов всех подзадач для получения окончательного результата сложной задачи. Среди них все подзадачи выполняются одновременно в разных ядрах ЦП; следовательно, он может максимально эффективно использовать компьютерные ресурсы для решения вычислительных задач и повышения скорости вычислений компьютерных систем. Рабочий процесс модели «начальник-подчиненный» показан на рис. 1.

Ожидаемые преимущества параллельной модели «начальник-подчиненный» заключаются в значительном сокращении общего времени выполнения алгоритма. В среде, состоящей из  «починных» ,  «начальников» могут одновременно работать на разных ядрах, тем самым сокращая общее время вычислений. Также обратите внимание, что в случае серьезной проблемы общая производительность распараллеливания относительно желаемого времени выполнения будет более заметной. Естественно, это не ухудшается и даже не затрагивается тривиальными проблемами обработки мелкомасштабных и временных задач. Но в последней ситуации обычно лучше придерживаться серийной операции. Кроме того, размер рабочей нагрузки параллельных задач (называемый степенью детализации) в модели «начальник-подчиненный» может оказывать огромное влияние на ее параллельную производительность. Если параллельная задача слишком мала или к главному процессу подключено слишком много подчиненных процессов, это увеличит рабочую нагрузку и приведет к увеличению времени выполнения. Следовательно, параллельное программирование требует параллельного тестирования детализации, выбирая между множеством небольших задач и несколькими большими задачами, чтобы найти критическую точку максимальной вычислительной эффективности.

*2.4. Прогнозирование плотности вероятности ветровой энергии на основе МНСПКР и функции ядра Епанечникова*

*2.4.1. Параллельная модель «главный-подчиненный» для многоядерной нейронной сети с параллельной квантильной регрессией (МНСПКР)*

В этой работе предлагается параллельная модель «начальник-подчиненный» метода НСКР. Алгоритм в основном использует параллельную модель «начальник-подчиненный» и сочетает в себе структурные особенности и механизм работы НСКР для параллельного проектирования. Метод НСКР прогнозирует соответствующие условные квантили энергии ветра при различных квантилях случайных величин. Таким образом, можно более тщательно отобразить влияние объясняющих переменных на объясняемые переменные в разных квантилях и взаимосвязь между ними. Задачи прогнозирования в разных точках квантиля независимы друг от друга, поэтому операции задач прогнозирования параллельны. Интуитивно, используя характеристики нескольких процессоров, МНСПКР преобразует процесс, который изначально выполнял последовательное прогнозирование в разных квантилях, в процесс, который может запускать несколько процессов одновременно для параллельного прогнозирования. Поэтому теоретически время работы можно сократить в разы.



**Рис. 1. Модель параллельного выполнения фреймворка начальник-подчиненный.**

В процессе расчета НСКР цикл итерационного процесса занимает огромное количество вычислительного времени. Аналогично, если предположить, что модель НСКР включает в себя  квантилей, а количество итераций в каждом квантиле равно , общее количество итераций будет равно . Очевидно, что по мере увеличения количества итераций и расширения масштаба ветроэнергетической системы общее количество вычислений будет расти в геометрической прогрессии. Параллельная модель МНСПКР использует метод декомпозиции задач для достижения параллельного проектирования с помощью параллельной структуры «начальник-подчиненный». Основываясь на независимости задачи прогнозирования в каждом отдельном квантиле, при выполнении НСКР в многоядерной среде  ядер могут одновременно выполнять задачи обучения нейронной сети. В результате теоретически количество вычислений на итерацию в ядре последовательной программы равно , а временная сложность алгоритма МНСПКР равна . Легко обнаружить, что МНСПКР может значительно повысить эффективность вычислений по сравнению с НСКР.

В МНСПКР разделение задач тесно связано с масштабом решаемой задачи и количеством процессоров. Как координировать узлы для равномерного распределения задач между вычислительными узлами, является основным фактором, влияющим на то, сможет ли параллельный компьютер эффективно обеспечивать свою производительность. Существует множество стратегий распределения задач, таких как распределение блоков и распределение статоров [45]. Поиск разумного метода назначения задач позволит параллельным компьютерам более эффективно использовать свою производительность. Вычисления в квантилях  независимы друг от друга, и между задачами не требуется никакой зависимости от данных, никакой связи или синхронизации. Распределение подзадач в этой задаче использует традиционный метод распределения блоков, то есть, если имеется  процессоров, задачи делятся на  блоков, а затем каждый процессор или группа процессов отвечает за выполнение одной из задач. Метод расчета количества задач прогнозирования, распределенных по каждому процессору, заключается в следующем: сначала задачи прогнозирования равномерно распределяются по каждому процессору. Если число процессоров не может быть кратным, оговариваем, что избыточные задачи прогнозирования последовательно назначаются процессорам с небольшим числом. На рис. 2 показана принципиальная схема стратегии распределения.

*Описание алгоритма МНСПКР:*

*Алгоритм работы главного процесса:*

(1) главный процесс загружает локальную задачу, разделяет задачи прогнозирования на разные квантили в соответствии с количеством подчиненных процессов (количеством ядер) и отправляет подзадачи и параметры сетевой структуры НСКР назначенным подчиненным процессам;

(2) основной процесс загружает локальный файл и отправляет образец каждому подчиненному процессу;

(3) основной процесс получает результаты обучения от каждого процесса и суммирует их;

*Алгоритм подчиненного процесса:*

(1) каждый подчиненный процесс получает задачи прогнозирования и параметры НСКР, отправленные главным процессом в разных квантилях, и устанавливает структуру сети в соответствии с принятыми параметрами сети, отправленными главным процессом, затем инициализирует сеть;

(2) получить указанную выборку из процесса, разделить пространство и сохранить выборку в локальном пространстве;

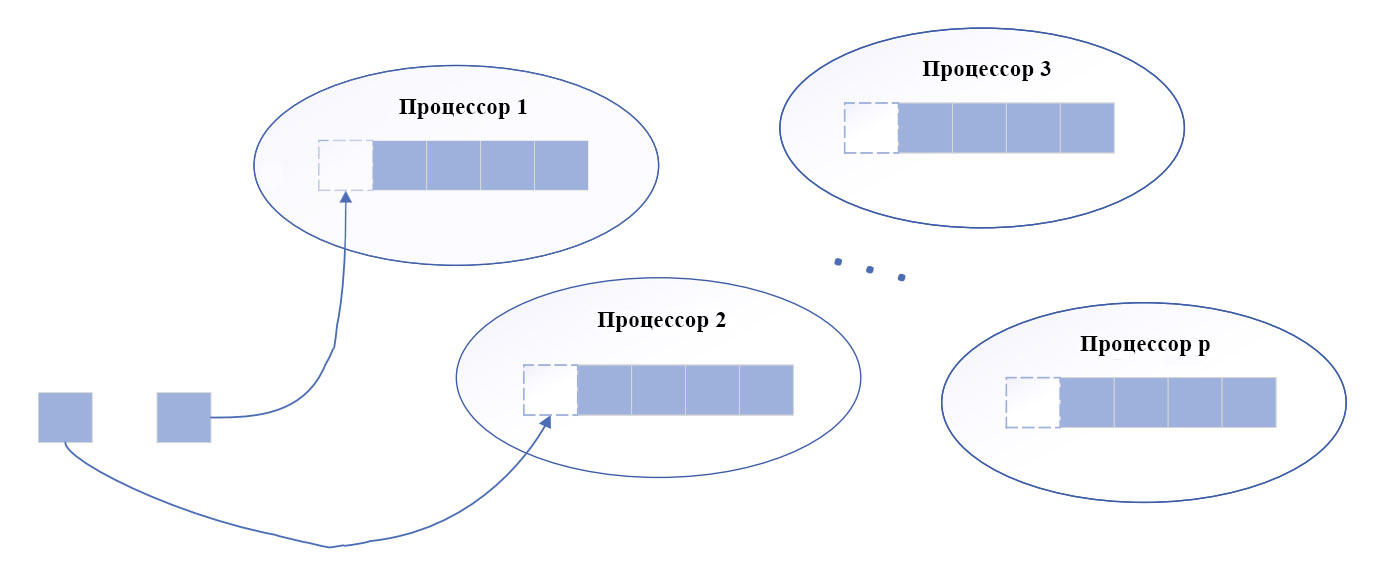
(3) каждый рабочий обучает все образцы подчиненного процесса в соответствии с алгоритмом НСКР, а затем отправляет результаты обучения главному процессу.

В течение этого периода каждый подчиненный процесс выполняет вычисления синхронно, не связываясь друг с другом. Описание алгоритма МНСПКР приведено в таблице 1.

*2.4.2. Оценка плотности ядра с использованием функции ядра Епанечникова*

В качестве хорошо известного непараметрического метода оценки функции плотности вероятности случайных величин в этой работе используется метод оценки плотности ядра для создания кривой плотности вероятности ветровой энергии. В отличие от интервальной оценки, преимущество метода оценки плотности ядра заключается в том, что основная функция плотности вероятности может быть построена непосредственно на основе наблюдаемых данных без гипотезы распределения [42,43]. Согласно теории подобия условный квантиль равен условной плотности, а затем функция условного квантиля, полученная в приведенном выше исследовании, используется в качестве входных данных функции метода оценки плотности ядра для построения кривой плотности вероятности ветровой энергии [46]. Полагая набор независимых переменных  , полученный из распределения оцененных функций плотности, ядерная оценка плотности квантиля описывается следующим образом:

|  |  |
| --- | --- |
|  | (8) |



**Рис. 2. Принципиальная схема стратегии распределения**

Таблица 1 Описание алгоритма МНСПКР

|  |
| --- |
| **Входные данные**: задача прогнозирования T, количество квантилей M, количество подчиненных p.  **Выход**: обучение на модели НСКР. |
| 1. Распределите T на p узлов с квантилями.  2. Главный процесс инициализирует сетевую структуру и параметры ИНС.  3.**While**(Mmod p = 0)  Главный процесс отправляет подзадачи M/p и параметры сетевой структуры каждому подчиненному процессу.  4.**End While**  5.**While**((Mmod p ≠ 0)  Главный процесс отправляет подзадачи M/p и параметры сетевой структуры каждому подчиненному процессу.  Ожидание, пока ведомое устройство завершит работу, затем отправьте оставшуюся подзадачу в простаивающий ведомый процесс с небольшим порядковым номером.  6.**End While**  7. Главный узел получает результаты обучения M от каждого процесса и суммирует их. |

где  — полоса пропускания, используемая для сглаживания параметров, и выбор полосы пропускания имеет решающее значение для метода оценки плотности ядра.  представляет количество квантилей, а  является неотрицательной ядерной функцией [47]. Результаты оценки метода оценки плотности ядра различаются для разных функций ядра. В общем, мы используем ядро Епанечникова с оптимальной среднеквадратической ошибкой [48]. Формула ядерной функции Епанечникова определяется уравнением:

|  |  |
| --- | --- |
|  | (9) |

где  равно .

Для более наглядного отображения блок-схема реализации метода прогнозирования плотности вероятности ветровой энергии на основе МНСПКР показана на рис. 3.

(1) Нормализация и инициализация: исторические данные о ветровой энергии должны быть стандартизированы до [0,1]. Формула:

|  |  |
| --- | --- |
| , | (10) |

где  — обработанные данные,  — необработанные данные, а ,  — максимум и минимум исходного ряда мощности ветра соответственно. Обработанные наборы данных делятся на обучающий набор и тестовый набор. Начальные веса  и  назначаются случайным образом с равномерно распределенными случайными числами из . Модель МНСПКР обучается с помощью набора обучающих данных для определения оптимальных весов. Наборы тестовых данных используются для проверки эффективности метода.

(2) Построить модель МНСПКР: стандартизированные ряды энергии ветра  используются в качестве входных данных МНСПКР в соответствии с уравнением. (5).

(3) Выбрать параметры модели: параметры структуры нейронной сети влияют на прогнозирующую эффективность МНСПКР, и мы определяем параметры модели методом проб и ошибок.

(4) Условная квантильная оценка энергии ветра: Условная квантильная оценка энергии ветра  может быть получена с помощью уравнения (7).

(5) Прогнозирование плотности вероятности энергии ветра: Функция плотности вероятности энергии ветра оценивается с помощью метода оценки плотности ядра, определяемого как уравнение (8). В данной работе в качестве функции ядра выбрано ядро Епанечникова, поскольку оно имеет наилучшее среднеквадратическое отклонение. Ядерная функция Епанечникова представлена в виде уравнения. (9). Условный квантиль  принимается в качестве входной переменной ядерной функции Епанечникова. В сочетании с методом KDE получается кривая плотности вероятности мощности ветра.

(6) Прекращение обучения: как только результаты прогнозирования будут соответствовать требованиям точности, прекратите обучение; в противном случае вернитесь к шагу 3.

(7) Анализ результатов прогнозирования: после процесса обучения оптимальные параметры применяются для проверки предлагаемого метода. Метод МНСПКР используется для получения условного квантиля, интервалов прогнозирования и кривых плотности вероятности ветровой энергии. Индексы оценки в разделе 3 используются для оценки эффективности метода плотности вероятности МНСПКР.

**3. Меры оценки**

**3.1. Меры оценки параллельной производительности**

Для оценки параллельной производительности алгоритма МНСПКР, разработанного в этом исследовании, здесь используются два важнейших показателя: ускорение и эффективность параллельного выполнения, которые определяются как  и  [49,50] соответственно.

(1) Ускорение

 — это время, когда задача выполняется последовательно на одном ядре,  — это время выполнения задач в параллельном режиме  — ядер, а  указывает количество задействованных процессоров. Ускорение параллельного алгоритма определяется как

|  |  |
| --- | --- |
| , | (11) |

где . Если параллельная программа выполняется в  раз быстрее, чем последовательная программа, а именно , параллельный алгоритм достигает максимальной производительности. Тогда это называется линейным ускорением параллельного алгоритма. На самом деле добиться оптимального ускорения сложно, поскольку несколько процессов всегда требуют определенных затрат.

(2) Параллельная эффективность

Если заменить формулу, то параллельную эффективность можно выразить как:

|  |  |
| --- | --- |
| , | (12) |

здесь , что отражает загрузку процессоров при выполнении алгоритма. Если эффективность параллельного алгоритма равна 1, это эквивалентно ускорению параллельного алгоритма, достигающему . Как уже говорилось выше, это вообще невозможно.

*3.2. Прогнозирование показателей оценки эффективности*

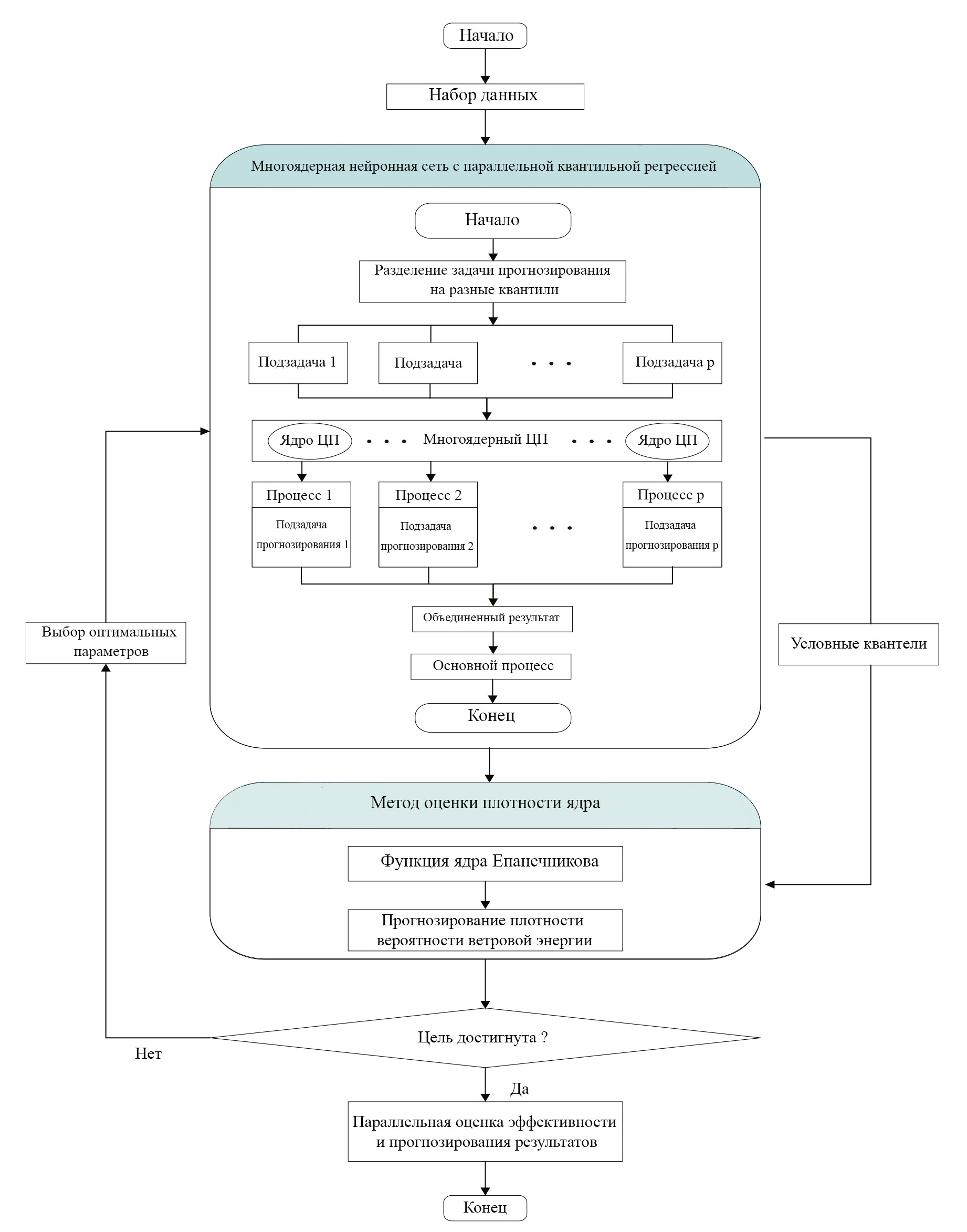
*(1) Точность результатов прогнозирования точек*

С целью проверки доступности метода прогнозирования плотности вероятности ветровой энергии на основе МНСПКР и ядерной функции Епанечникова в статье используются два наиболее распространенных и достоверных статистических показателя для оценки результатов прогнозирования ветровой энергии: MAPE (средняя абсолютная процентная ошибка) [51] и RMSE (среднеквадратическая ошибка) [52]. Эти формулы показаны ниже:

|  |  |
| --- | --- |
| , | (13) |

|  |  |
| --- | --- |
| , | (14) |

где  — это фактический результат использования энергии ветра, а  — прогнозируемое значение. Поскольку энергия ветра является неопределенностью и изменчивостью, так что выходная мощность стремится к 0 или равна 0,  представляет собой максимальное значение исторических выборок энергии ветра, которое используется здесь в качестве знаменателя для расчета , избегая случая, когда ошибка значение приближается к бесконечности [51].



**Рис. 3. Подробная блок-схема структуры МНСПКР**

*(2) Неопределенность интервалов прогнозирования*

В то же время, как и в случае точечного прогнозирования, также крайне важно количественно оценить неопределенность интервалов прогнозирования с точки зрения прогнозирования плотности вероятности. Поэтому в этой работе также используются вероятность покрытия (PICP) и нормализованная средняя ширина (PINAW) [53] в качестве двух важных показателей для оценки производительности. PICP демонстрирует процент целей. Большее значение PICP означает, что в прогнозируемые интервалы попадает больше реальных данных, тогда как слишком маленькое значение PICP указывает на неприемлемое поведение покрытия. Формула расчета PICP:

|  |  |
| --- | --- |
| , | (15) |

в которой  — булева переменная, а переменная , если , в противном случае . Здесь ,  — верхняя и нижняя границы i-го PI. Чтобы получить отличные прогнозируемые интервалы, PICP должен быть выше уровня достоверности. Однако бесполезно просто обращать внимание на более высокий PICP, если игнорировать размер PINAW. Фактически, с точки зрения прогнозирования неопределенности, ширина прогнозного интервала (PINAW) также должна быть как можно более узкой при выборе более высокой вероятности покрытия (большего PICP). PINAW является жизненно важной мерой оценки качества прогнозного интервала [53,54]. Его математическая формула выглядит следующим образом:

|  |  |
| --- | --- |
| , | (16) |

в вышеупомянутом уравнении  — это разница между верхней и нижней границами базового целевого диапазона, которая используется для нормализации средней ширины прогнозный интервал в процентах.

Таблица 2. Подробная информация о четырех наборах данных.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Набор данных | Время | Т | Tr | Te |
| Единицы | 1 период = 1 час | Период |  |  |
| Набор данных 1 | 06/2021–08/2021 | 2208 | 1405 | 803 |
| Набор данных 2 | 06/2021–09/2021 | 2928 | 2308 | 620 |
| Набор данных 3 | 06/2021–10/2021 | 3672 | 2950 | 722 |
| Набор данных 4 | 06/2021–11/2021 | 4932 | 3682 | 1250 |

Таблица 3. Основные параметры, используемые в НСКР и МНСПКР

|  |  |  |
| --- | --- | --- |
| Метод | Параметр (уровень достоверности 95 %) |  |
| МНСПКР & НСКР | Максимальное количество итераций  Штраф на вес за регуляризацию снижения веса  Количество скрытых узлов  Количество квантилей  Количество повторных испытаний | 1000  3  10  20  10 |

**4. Анализ результатов эксперимента**

*4.1. Источники данных и анализ данных*

В экспериментах использовалась нечеткая нейронная сеть Такаги–Сугено в Matlab 2018b. Другие эксперименты были реализованы в версии R 3.6.0 с применением языка R, который в основном включает пакеты «AMORE», «RSNNS», «quantreg», «parallel» и «QRNN». Операционная система компьютера — версия Windows 10, а процессор состоит из 6-ядерного процессора Intel (R) Core i7-8700 с частотой 3,20 ГГц. Кроме того, используются случайные начальные числа, чтобы гарантировать возможность повторения эксперимента. Чтобы проверить достоверность и осуществимость модели МНСПКР, для проверки и тестирования использовался набор данных о ветровой энергии из работы [55]. Онтарио — вторая по величине провинция Канады. В энергетической структуре Онтарио более высокая доля приходится на возобновляемые источники энергии, включая гидроэнергетику, солнечную энергию, энергию ветра и т. д. Среди них технология производства ветровой энергии является зрелой, экономически целесообразной и имеет широкие перспективы развития. Пример основан на исторических данных о ветровой энергии с 1 июня 2021 года по 30 ноября 2021 года с разрешением 1 час. Производительность модели тестируется с применением четырех наборов данных разной длиной, чтобы проанализировать адекватность параллельных методов для различных сценариев. Каждый набор данных разделен на две части: обучающий набор и тестовый набор. В наборе данных 1 мы применяем данные о ветровой энергии с 1 июня по 31 июля 2021 года в качестве обучающего набора для прогнозирования почасовой мощности ветра на август 2021 года. Обучающие данные набора данных 2 представляют собой значения мощности ветра, выбранные с 1 июня по 31 августа. 2021. С 1 по 30 сентября 2021 года прогнозируется сила ветра в общей сложности 720 баллов. На этой основе расширено разделение обучающего набора и тестового набора в наборах данных 3 и 4. Подробную информацию об этих выборках можно увидеть в Таблице 2. T, Tr и Te представляют общее количество выборок, обучающего набора и тестового набора соответственно. Временные ряды четырех наборов данных представлены на рис. 4.

*4.2. Сравнение последовательного алгоритма НСКР и параллельного алгоритма МНСПКР*

В этом разделе традиционный алгоритм последовательного НСКР и алгоритм МНСПКР экспериментально сравниваются в четырех разных наборах данных в разных многоядерных средах. Прежде чем обучать модель, мы сначала нормализуем все обучающие и тестовые выборки. Кроме того, задав количество работающих ядер, мы оценили алгоритм МНСПКР в двухъядерных, трехъядерных, четырехъядерных, пятиядерных и шестиядерных средах соответственно. При этом время работы последовательной программы мы принимаем за время работы одноядерной в параллельной системе. Чтобы сделать экспериментальные результаты более точными и достоверными, все экспериментальные параметры моделей МНСПКР и НСКР установлены абсолютно одинаковыми. На основе анализа наборов обучающих данных методом проб и ошибок определяются параметры модели перед обучением. Подробности моделей НСКР и МНСПКР обобщены в таблице 3. В конечном итоге структура двух моделей нейронных сетей подтверждена как 4-10-1; то есть 4 входных узла, 10 скрытых слоев и один выходной слой. В качестве входных данных моделей используется энергия ветра за предыдущие 4 часа. Благодаря структуре обученной нейронной сети получаются последовательные условные квантили под разными квантилями. Квантили, выбранные в данной работе, варьируются от 0,01 до 0,99 с интервалом 0,05. Затем прогнозируемая выработка энергии ветра, принадлежащая разным квантилям, используется в качестве входной переменной ядерной функции Епанечникова. В сочетании с методом оценки плотности ядра получается кривая плотности вероятности мощности ветра.

Кроме того, в тех же условиях каждый алгоритм независимо повторяется 10 раз под заданным количеством ядер для каждого набора данных для статистического анализа среднего значения экспериментальных результатов. В таблице 4 показаны результаты времени выполнения, ускорения и параллельной эффективности для двух алгоритмов с разным количеством ядер. Среднее время выполнения обозначает среднее статистическое время эксперимента для 10 независимых экспериментов.

Как показано в таблице 4, временная сложность НСКР быстро растет с увеличением размера обучающих выборок. Например, в наборе данных 4 время обучения последовательного НСКР достигает 785,70 с, тогда как для распараллеленного МНСПКР оно занимает всего 247,99 с в 5-ядерной среде. Ускорение достигает 3,168, а эффективность параллельного выполнения составляет 63,36%. Как и в наборе данных 4, время работы 5-ядерного МНСПКР по-прежнему самое короткое в наборах данных 1–3. Таким образом, алгоритм МНСПКР позволяет максимально эффективно использовать параллельные ресурсы ЦП, значительно сокращая время вычислений и повышая эффективность работы.

В то же время из рис. 5 видно, что при одинаковом количестве ядер ЦП по мере увеличения размера задачи время обучения параллельного алгоритма меньше, чем у последовательного алгоритма, что практически приведет к большему ускорению.

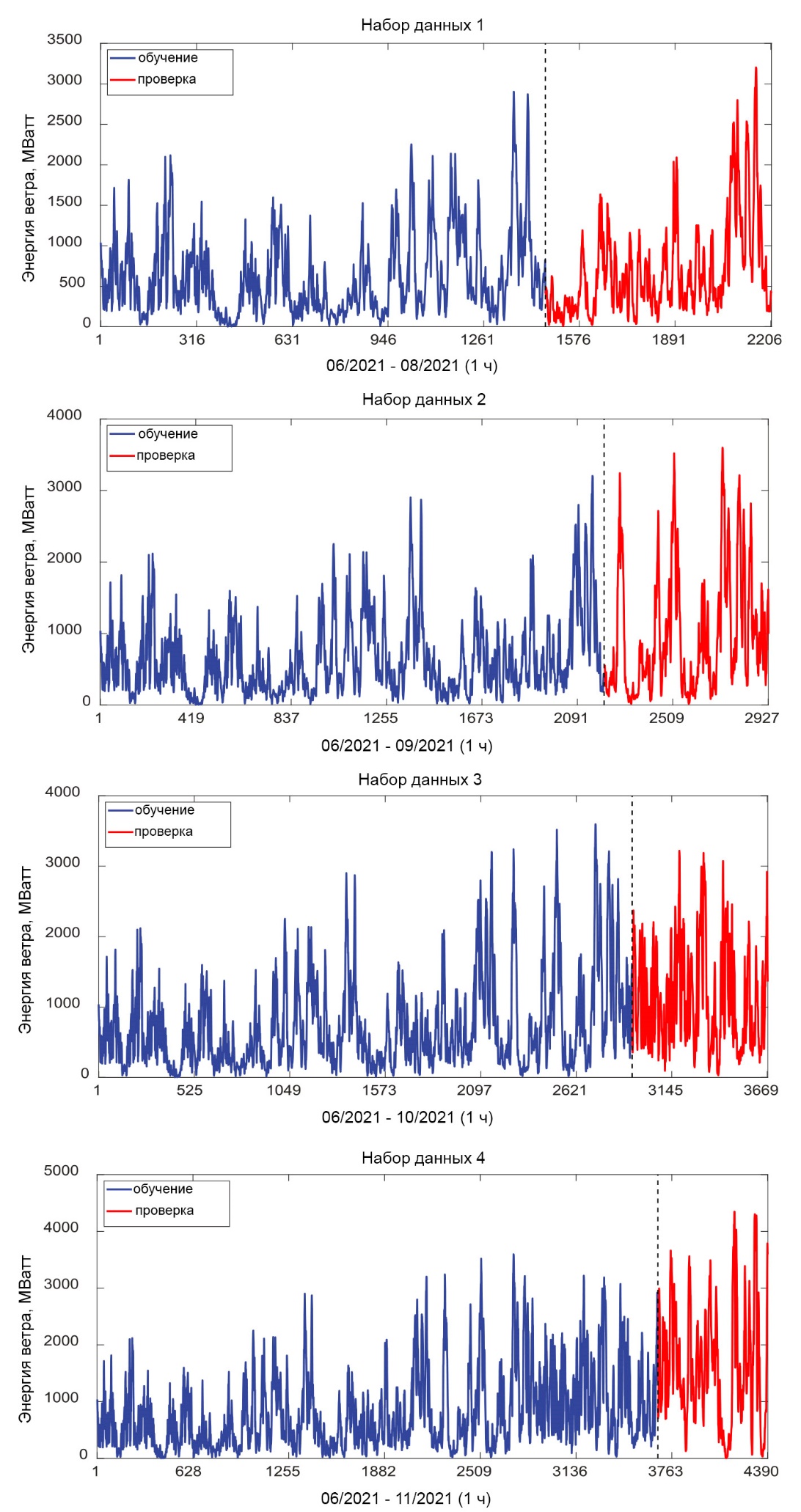
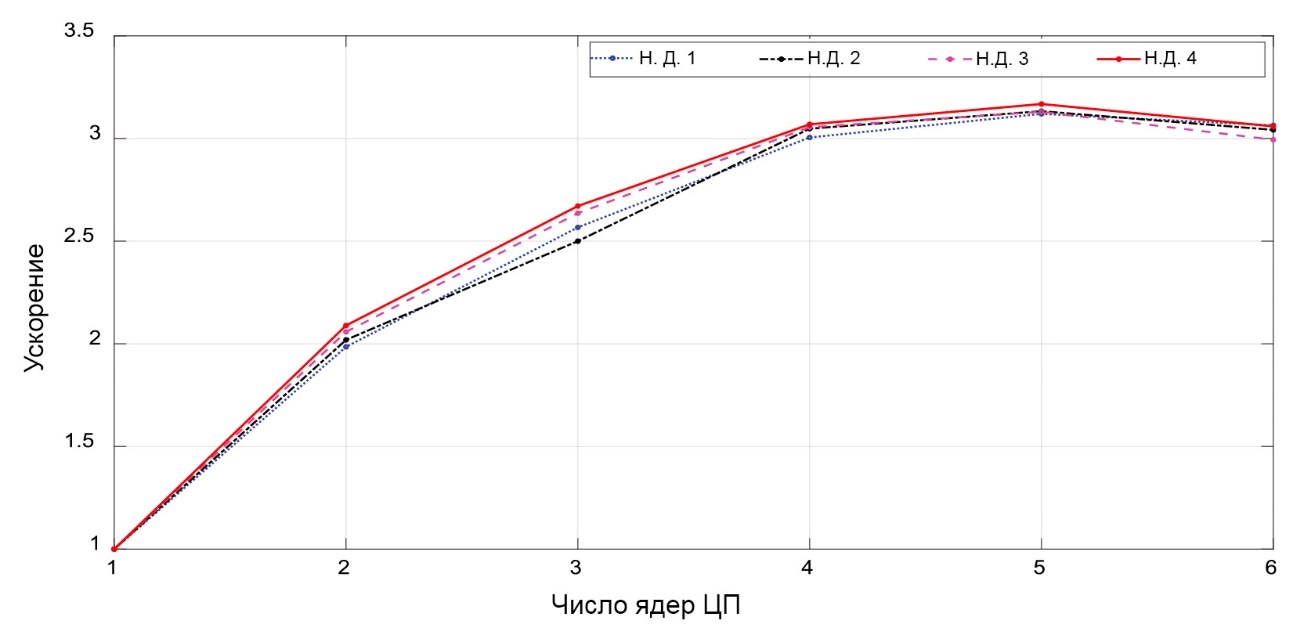


Рис. 4. Ветровая энергия четырех наборов данных.

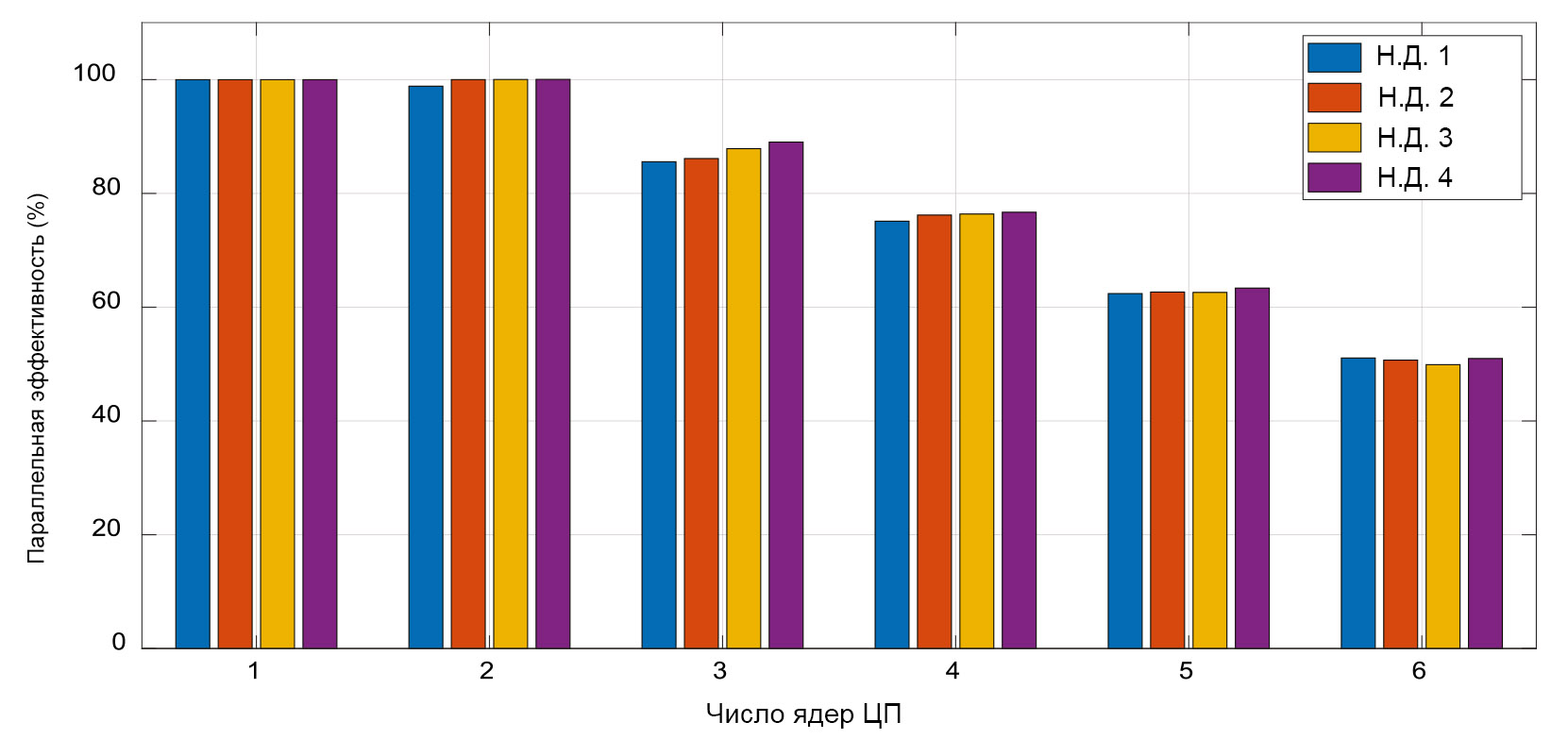
Таблица 4. Время выполнения, ускорение и эффективность при различном количестве процессов.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Набор данных | (НСКР) серийные запуски | (МНСПКР) параллельные запуски | | Sp | Ep |
| Среднее время выполнения (с) | Количество ядер | Среднее время выполнения (с) |
| 1 | 332.58 | 2 | 167.55 | 1.985 | 98.85% |
| 3 | 129.55 | 2.567 | 85.57% |
| 4 | 110.66 | 3.005 | 75.12% |
| 5 | 106.58 | 3.120 | 62.40% |
| 6 | 108.54 | 3.064 | 51.07% |
| 2 | 481.66 | 2 | 238.56 | 2.019 | 100.01% |
| 3 | 186.39 | 2.584 | 86.13% |
| 4 | 158.02 | 3.048 | 76.20% |
| 5 | 153.73 | 3.133 | 62.66% |
| 6 | 158.36 | 3.042 | 50.70% |
| 3 | 630.29 | 2 | 306.32 | 2.058 | 100.03% |
| 3 | 239.11 | 2.636 | 87.87% |
| 4 | 206.31 | 3.055 | 76.38% |
| 5 | 201.37 | 3.130 | 62.60% |
| 6 | 210.52 | 2.994 | 49.90% |
| 4 | 785.70 | 2 | 376.17 | 2.089 | 100.04% |
| 3 | 294.19 | 2.671 | 89.03% |
| 4 | 256.05 | 3.068 | 76.70% |
| 5 | 247.99 | 3.168 | 63.36% |
| 6 | 256.85 | 3.059 | 50.98% |



**Рис. 5. Ускорение четырех образцов для заданного количества ядер.**

Кроме того, для того же объема данных на рис. 5 также видно, что ускорение увеличивается с увеличением ядра ЦП, когда количество ядер ЦП не превышает 5 . Но это непропорционально увеличению количества ядер процессора. В частности, в параллельной среде работы 5-ядерника ускорение достигает пика. Ускорение уменьшится, когда количество ядер процессора станет равным 6.



**Рис. 6. Параллельная эффективность четырех образцов для заданного количества ядер.**

Таблица 5 MAPE, RMSE, PICP и PINAW в МНСПКР и НСКР.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Набор данных | Методы | MAPE (%) | RMSE (MВатт) | PICP (%) | PINAW (%) |
| 1 | НСКР | 3.05 | 185.74 | 99.19 | 26.98 |
| МНСПКР 2 | 3.05 | 185.37 | 99.34 | 26.90 |
| МНСПКР 3 | 3.04 | 185.21 | 99.19 | 27.47 |
| МНСПКР 4 | 3.01 | 183.81 | 99.06 | 27.17 |
| МНСПКР 5 | 3.05 | 185.44 | 99.06 | 27.28 |
| МНСПКР 6 | 3.04 | 185.69 | 99.19 | 27.45 |
| 2 | НСКР | 3.51 | 217.38 | 98.33 | 25.19 |
| МНСПКР 2 | 3.53 | 218.76 | 98.61 | 25.62 |
| МНСПКР 3 | 3.49 | 216.27 | 98.06 | 25.69 |
| МНСПКР 4 | 3.53 | 217.78 | 98.19 | 25.37 |
| МНСПКР 5 | 3.52 | 218.50 | 98.75 | 25.78 |
| МНСПКР 6 | 3.52 | 217.91 | 98.33 | 25.08 |
| 3 | НСКР | 4.08 | 245.87 | 98.79 | 30.63 |
| МНСПКР 2 | 4.20 | 251.60 | 99.06 | 30.45 |
| МНСПКР 3 | 4.24 | 253.67 | 99.06 | 30.25 |
| МНСПКР 4 | 4.23 | 253.21 | 98.79 | 30.48 |
| МНСПКР 5 | 4.21 | 252.39 | 98.92 | 30.15 |
| МНСПКР 6 | 4.23 | 253.42 | 98.92 | 29.54 |
| 4 | НСКР | 4.56 | 267.00 | 99.86 | 29.18 |
| МНСПКР 2 | 4.47 | 261.66 | 99.72 | 29.01 |
| МНСПКР 3 | 4.50 | 263.83 | 99.86 | 28.47 |
| МНСПКР 4 | 4.52 | 264.08 | 99.86 | 28.94 |
| МНСПКР 5 | 4.47 | 262.18 | 99.72 | 28.66 |
| МНСПКР 6 | 4.51 | 263.36 | 99.86 | 28.94 |

Основная причина заключается в том, что при соответствующей параллельной детализации увеличение количества процессов уменьшает задачи обучения в наборе данных, совместно используемом каждым работником, тем самым сокращая время обучения; но накладные расходы на связь увеличиваются с увеличением количества процессов. В кластерной среде стоимость одного общения намного превышает стоимость одного расчета. Хотя большее количество процессов будет генерировать меньше вычислений для каждого работника, время связи между процессами значительно увеличится. Когда доля времени связи превышает определенный предел, эффективность параллельных вычислений снижается [30]. С другой стороны, увеличение количества процессов также усугубит потребление системных ресурсов для запуска программ, например системной памяти, что также приведет к снижению эффективности вычислений. Следовательно, параллельная степень детализации должна выбираться соответствующим образом в соответствии с реальной ситуацией. Как показано на рис. 6, легко обнаружить, что при одинаковом количестве ядер ЦП эффективность небольших наборов данных ниже, чем у крупномасштабных наборов данных. Кроме того, в той же ситуации увеличение количества процессов приводит к постепенному снижению параллельной эффективности.

Очевидно, что ,  и  зависят не только от количества процессоров, но и от размера проблемы. Многие параллельные программы разделяют задачи последовательных программ, распределяют их между процессами и добавляют необходимые «параллельные издержки», такие как взаимное исключение или обмен данными. Следовательно, если  используется для представления параллельных служебных данных, то

|  |  |
| --- | --- |
| , | (17) |

По мере увеличения масштаба проблемы  растет медленнее, чем , у каждого процесса появляется больше задач, а время, необходимое для координации процесса, сравнительно меньше.

Кроме того, для дальнейшей проверки эффекта оптимизации многоядерного параллельного алгоритма ошибки прогнозирования точек двух алгоритмов сравниваются с помощью MAPE и RMSE, как показано в таблице 5. Согласно результатам таблицы 5, прогнозирование точки результаты с использованием метода МНСПКР имеют небольшие колебания в зависимости от результатов прогнозирования традиционного метода последовательных НСКР. Это связано со случайностью при обучении методов НСКР, но общее качество прогнозирования не ухудшилось. При этом для оценки эффективности выбираются оптимальные PICP и PINAW в 10 повторных экспериментах. Как видно из таблицы 6, уровень достоверности прогнозного интервала, построенного методами МНСПКР и НСКР, может быть удовлетворен. Для набора данных 1 PICP и PINAW НСКР составляют 99,46% и 23,04% соответственно. Кроме того, для сравнения мы выбрали 5-ядерный МНСПКР с наибольшим ускорением. В то время как в 5-ядерной среде параллельных вычислений с наибольшим ускорением значение PICP МНСПКР немного ниже, чем у НСКР, а PINAW МНСПКР шире, чем у метода НСКР. Поскольку ошибка прогнозирования точки 5-ядерного МНСПКР ниже, чем у НСКР в наборе данных 1, МНСПКР может только увеличить ширину интервала, чтобы PICP соответствовал 95%. Аналогично, для набора данных 4 значение PICP, полученное с использованием НСКР, составляет 99,86%, что на 0,14% выше, чем результаты, полученные с помощью 5-ядерного МНСПКР, но PINAW МНСПКР уже. Из таблицы 5 видно, что в наборе данных 3 значение PICP, полученное МНСПКР с 5 ядрами, на 0,13% выше, чем результат НСКР, а значение PINAW 5-ядерного МНСПКР уже. В наборе данных 2 результаты интервального прогнозирования 5-ядерного МНСПКР немного хуже, чем у НСКР. Результаты интервального прогнозирования 5-ядерных МНСПКР и НСКР на четырех наборах данных представлены на рис. 7. Из рисунков видно, что результаты прогнозирования, полученные методом МНСПКР, также находятся в непосредственной близости от кривой фактических значений, а фактические точки почти все лежат в пределах узкого прогнозного интервала.

Кривые плотности вероятности, полученные с помощью МНСПКР с 5 ядрами и НСКР для всех наборов данных посредством выборки с равными интервалами, показаны на рис. 8 соответственно. Как видно из рисунков, кривые плотности вероятности на основе МНСПКР практически идентичны кривым на основе НСКР. Кроме того, на рис. 8(а), за исключением 373-го и 621-го часов, точки наблюдения почти все находятся в окрестности пика кривых, что указывает на совпадение прогнозируемых значений с реальными с высокой вероятностью. Это видно из рис. 8(б), только в 1-й, 121-й и 601-й часы реальные значения отклоняются от центра кривых плотности вероятности. Аналогично, фактические значения почти для всех периодов, кроме 621-го часа, попадают на пик кривой на рис. 8 (в). Его можно найти по рис. 8 (г) видно, что наблюдение расположено вблизи центра кривой в 241-й, 481-й, 601-й часы. Таким образом, результаты прогнозирования плотности вероятности на основе МНСПКР убедительны.

*4.3. Сравнение разных методов*

В этом разделе результаты точечного прогнозирования последовательных и параллельных моделей НСКР сравниваются с существующими моделями нечеткой нейронной сети [56], нейронной сети обратного распространения ошибки и нейронной сети радиальной базисной функции [57].

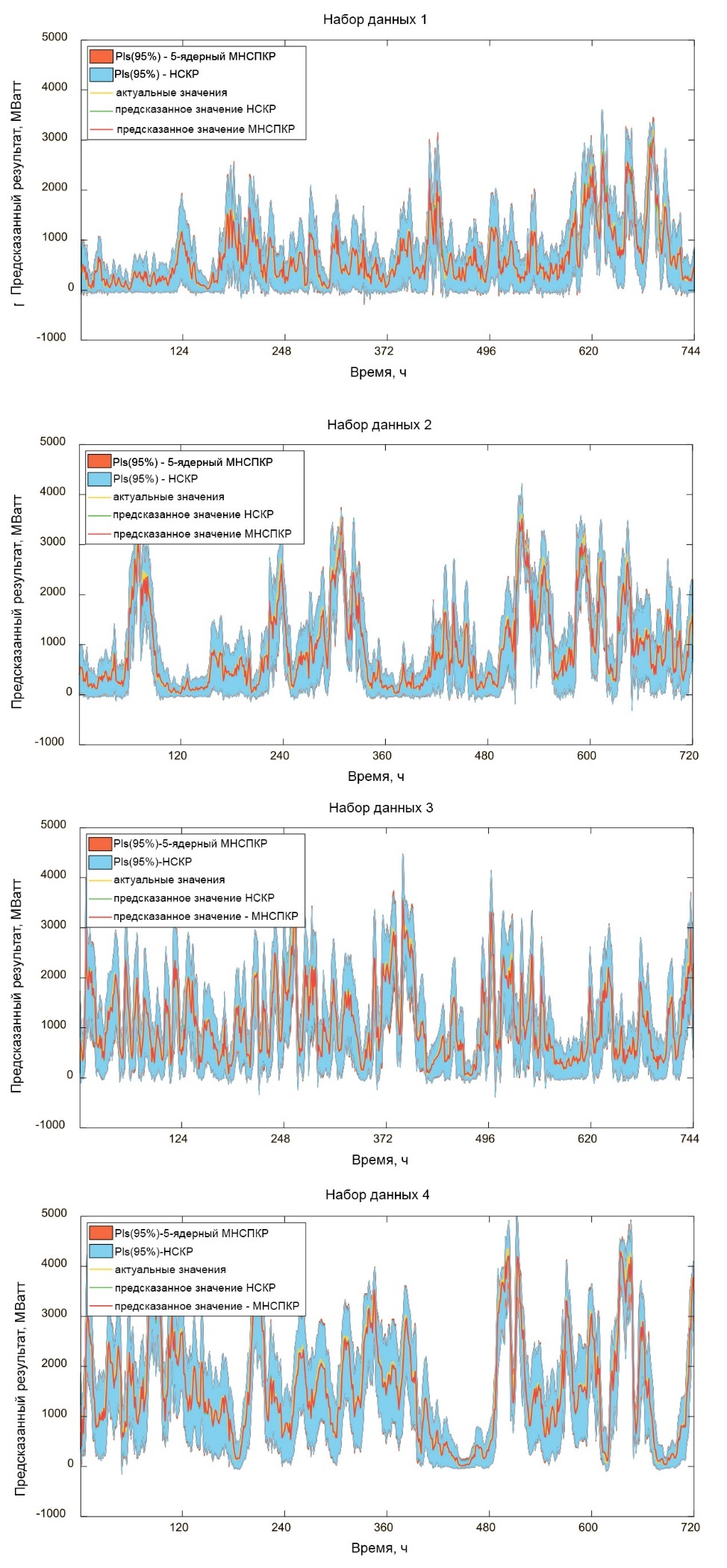
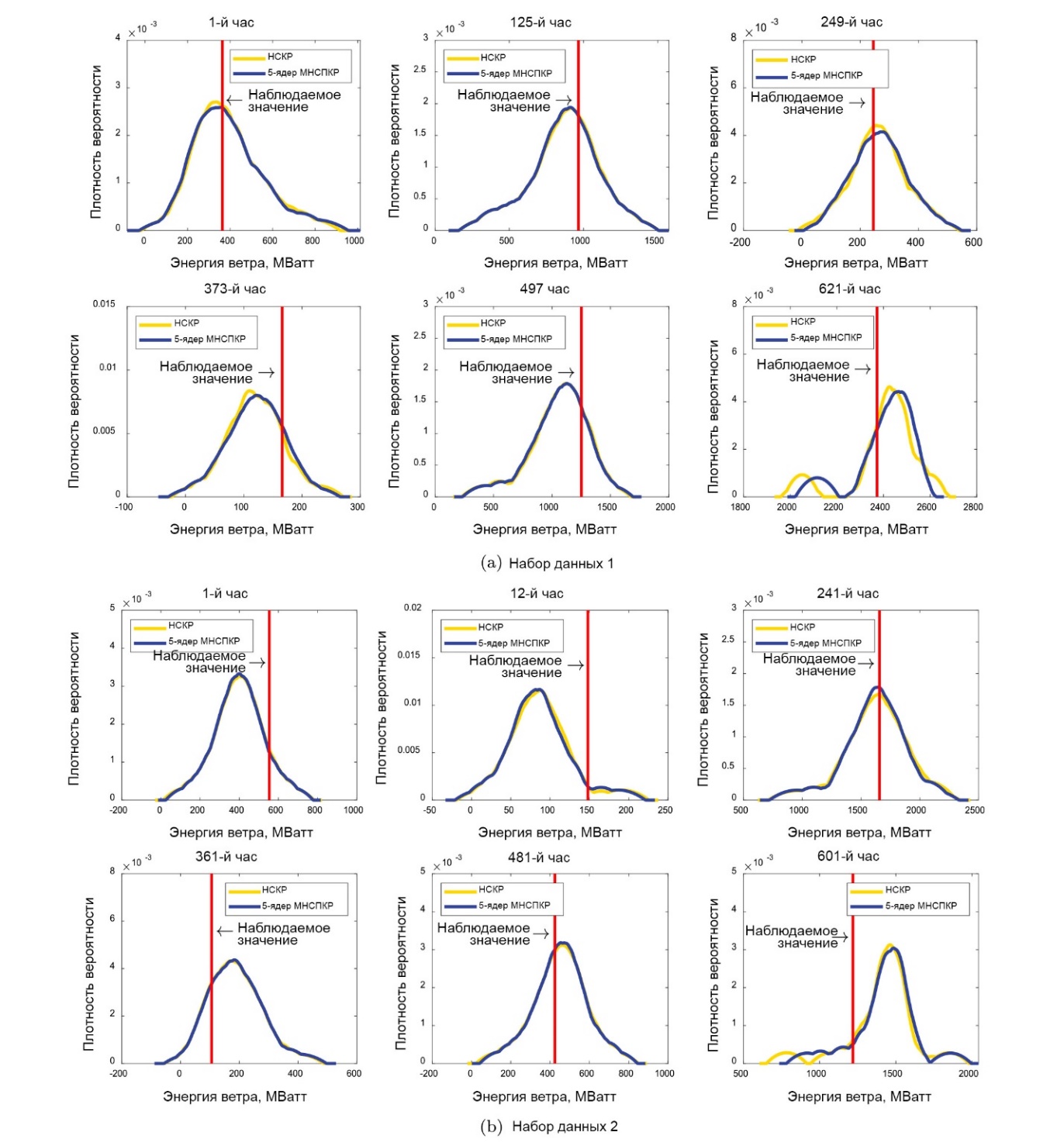
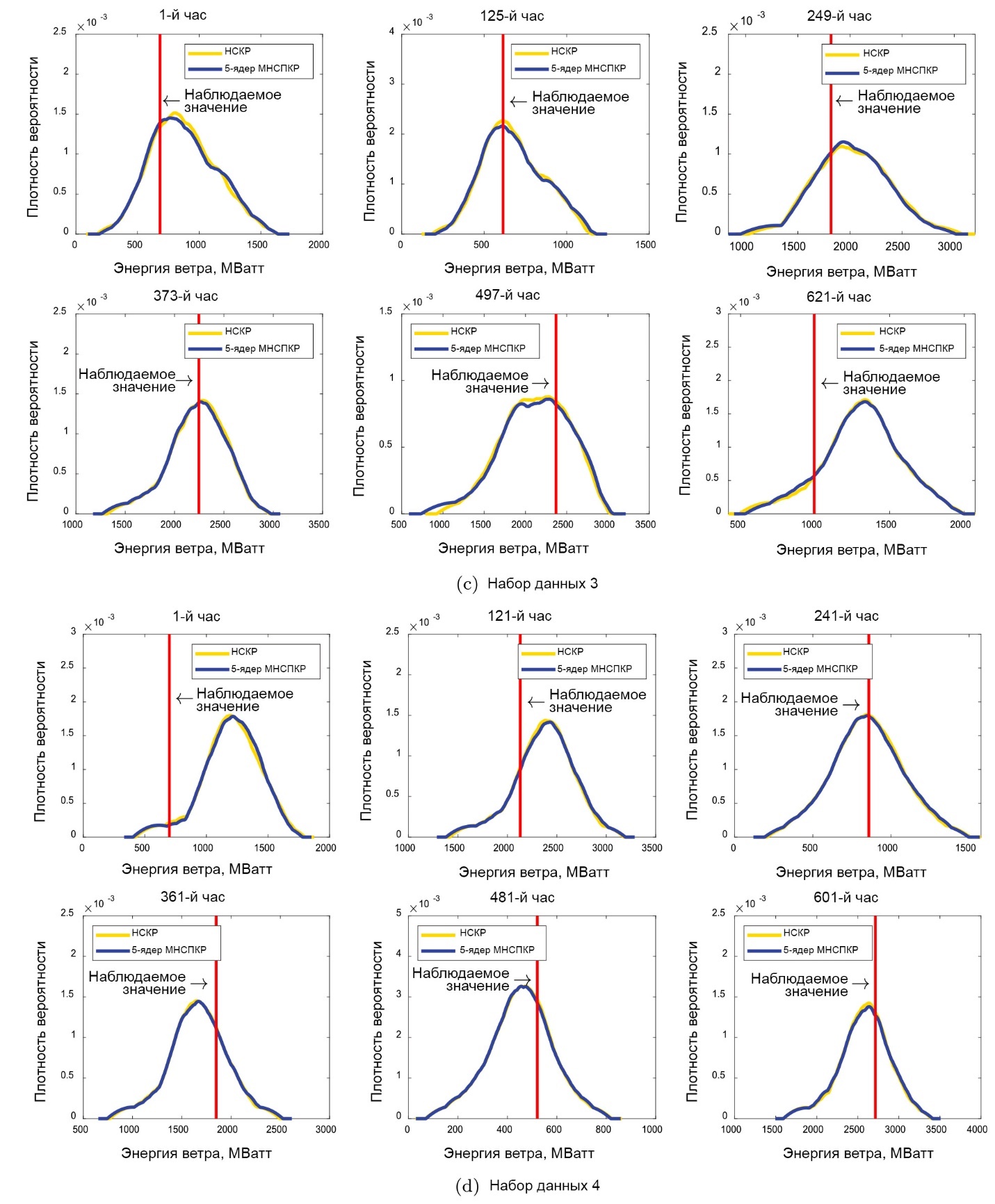


Рис. 7. Результаты прогнозирования НСКР и 5-ядерного метода МНСПКР для всех наборов данных.



**Рис. 8. Диаграммы кривых плотности вероятности на основе НСКР и 5-ядерного МНСПКР на четырех наборах данных**

Для сравнения: основные параметры трёх моделей нейросетей имеют те же параметры, что и НСКР. Между тем для сравнения с ним также используется метод предсказания плотности вероятности квантельной регрессии [58]. Чтобы проиллюстрировать преимущества модели МНСПКР с точки зрения эффективности и производительности прогнозирования, ошибки прогнозирования каждой модели сравниваются с использованием различных наборов данных, как показано в таблице 6. Из двух показателей оценки ошибок в таблице 6 видно, что модели НСКР и МНСПКР демонстрируют лучшая производительность прогнозирования, в то время как МНСПКР имеет более высокую скорость вычислений. Это показывает превосходство модели МНСПКР в прогнозировании ветровой энергии.



**Рис. 8. (продолжение).**

**5. Вывод**

В этой работе предлагается совершенно новый метод МНСПКР для прогнозирования плотности вероятности для решения проблемы эффективности традиционной НСКР, чтобы прогнозирование энергии ветра могло быть более надежным и быстрым в среде больших данных. Мы тестируем алгоритм, используя архивные данные о энергии ветра.

Таблица 6 MAPE, RMSE различных методов.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Модель | Набор данных 1 | | Набор данных 2 | | Набор данных 3 | | Набор данных 4 | |
|  | MAPE  (%) | RMSE  (МВатт) | MAPE  (%) | RMSE  (МВатт) | MAPE  (%) | RMSE  (МВатт) | MAPE  (%) | RMSE  (МВатт) |
| НСКР | 3.05 | 185.74 | 3.51 | 217.38 | 4.08 | 245.87 | 4.56 | 267.00 |
| TS-FNN | 3.98 | 239.12 | 5.24 | 320.49 | 5.08 | 287.26 | 6.10 | 530.24 |
| BPNN | 3.96 | 238.37 | 4.66 | 284.86 | 6.02 | 351.93 | 5.07 | 305.41 |
| RBFNN | 3.69 | 225.67 | 3.67 | 230.83 | 4.54 | 262.33 | 6.01 | 354.07 |
| КР | 3.27 | 207.61 | 4.27 | 265.77 | 4.63 | 276.41 | 5.76 | 333.42 |
| МНСПКР 5 | 3.05 | 185.44 | 3.52 | 218.50 | 4.21 | 252.39 | 4.47 | 262.18 |

Параллельная производительность экспериментальных результатов анализируется с использованием ускорения и параллельной эффективности. В то же время четыре метрики проверки MAPE, RMSE, PICP и PINAW используются для оценки точности прогнозирования точек и пригодности интервального прогнозирования между МНСПКР и существующим НСКР. Кроме того, предложенный метод и другие пять моделей реализованы на четырех наборах данных для контрастных экспериментов, а всестороннее обсуждение проводится с точки зрения результатов прогнозирования точек. Результаты исследования и выводы можно резюмировать следующим образом:

(1) Учитывая большой объем данных о ветровой энергии, метод МНСПКР использует независимые характеристики задач прогнозирования НСКР в разных квантилях для параллельного расчета нескольких задач прогнозирования. По сравнению с традиционным последовательным алгоритмом алгоритм МНСПКР, предложенный в этой статье, может полностью обеспечить более высокую скорость вычислений и эффективность работы.

(2) С увеличением размера вычислительной задачи время алгоритма МНСПКР сокращается более явно. Кроме того, по мере увеличения количества ядер ЦП преимущества метода параллельной обработки станут более существенными; однако параллельное ускорение не продолжает увеличиваться с увеличением количества процессов, а снижается после достижения пика. Снижение ускорения объясняется тем, что увеличение количества процессов приводит к большему потреблению системных ресурсов. В четырех экспериментальных случаях, описанных в этой статье, производительность параллельного выполнения оптимальна в пяти ядерной параллельной операционной среде.

(3) Прогноз плотности вероятности мощности ветра на основе модели МНСПКР также дает результаты точечного прогнозирования, аналогичные НСКР, точность значительно выше, чем у других четырех сравнительных моделей. В то же время полученный интервал прогнозирования является подходящим, он охватывает как можно больше значений наблюдения с минимально возможной шириной. Таким образом, МНСПКР получает высокоточное предсказание точек и подходящие прогнозные интервалы с наименьшим временем обучения. Метод данной работы дополнительно оптимизирует модель прогнозирования ветровой энергии, что станет достойным путем к повышению эффективности работы гигантских ветроэнергетических систем в будущем, особенно с учетом роста ветроэнергетики и постоянного совершенствования компьютерного программного обеспечения. и аппаратное обеспечение.

**Литература**

1. Li L. et al. An adaptive time-resolution method for ultra-short-term wind power prediction //International Journal of Electrical Power & Energy Systems. – 2020. – Т. 118. – С. 105814.

2. Lin Z., Liu X., Collu M. Wind power prediction based on high-frequency SCADA data along with isolation forest and deep learning neural networks //International Journal of Electrical Power & Energy Systems. – 2020. – Т. 118. – С. 105835.

3. Hong D. Y. et al. Ultra-short-term forecast of wind speed and wind power based on morphological high frequency filter and double similarity search algorithm //International Journal of Electrical Power & Energy Systems. – 2019. – Т. 104. – С. 868-879.

4. Newell R., Raimi D., Aldana G. Global energy outlook 2019: the next generation of energy //Resources for the Future. – 2019. – Т. 1. – С. 8-19.

5. Li Y. Z. et al. Wind-thermal power system dispatch using MLSAD model and GSOICLW algorithm //Knowledge-Based Systems. – 2017. – Т. 116. – С. 94-101.

6. Tasnim S. et al. Wind power prediction in new stations based on knowledge of existing Stations: A cluster based multi source domain adaptation approach //Knowledge-Based Systems. – 2018. – Т. 145. – С. 15-24.

7. Yan L. et al. The performance prediction of ground source heat pump system based on monitoring data and data mining technology //Energy and Buildings. – 2016. – Т. 127. – С. 1085-1095.

8. Rabieyan R., Pohl P. Improving a fuzzy neural network for predicting storage usage and calculating customer value //Journal of Revenue and Pricing Management. – 2020. – Т. 19. – С. 292-301.

9. Azimi R., Ghofrani M., Ghayekhloo M. A hybrid wind power forecasting model based on data mining and wavelets analysis //Energy conversion and management. – 2016. – Т. 127. – С. 208-225.

10. Zendehboudi A., Baseer M. A., Saidur R. Application of support vector machine models for forecasting solar and wind energy resources: A review //Journal of cleaner production. – 2018. – Т. 199. – С. 272-285.

11. Kotsiopoulos T. et al. Machine learning and deep learning in smart manufacturing: The smart grid paradigm //Computer Science Review. – 2021. – Т. 40. – С. 100341.

12. Gil R. P. A. et al. Surrogate model based optimization of traffic lights cycles and green period ratios using microscopic simulation and fuzzy rule interpolation //Int. J. Artif. Intell. – 2018. – Т. 16. – №. 1. – С. 20-40.

13. Zuo H. et al. Granular fuzzy regression domain adaptation in Takagi–Sugeno fuzzy models //IEEE Transactions on Fuzzy Systems. – 2017. – Т. 26. – №. 2. – С. 847-858.

14. Wang H. et al. Probabilistic wind power forecasting based on spiking neural network //Energy. – 2020. – Т. 196. – С. 117072.

15. Cannon A. J. Quantile regression neural networks: Implementation in R and application to precipitation downscaling //Computers & geosciences. – 2011. – Т. 37. – №. 9. – С. 1277-1284.

16. He Y., Li H. Probability density forecasting of wind power using quantile regression neural network and kernel density estimation //Energy conversion and management. – 2018. – Т. 164. – С. 374-384.

17. Wang J. Bayesian quantile regression for parametric nonlinear mixed effects models //Statistical Methods & Applications. – 2012. – Т. 21. – С. 279-295.

18. Taylor J. W. A quantile regression neural network approach to estimating the conditional density of multiperiod returns //Journal of forecasting. – 2000. – Т. 19. – №. 4. – С. 299-311.

19. Zameer A. et al. Intelligent and robust prediction of short term wind power using genetic programming based ensemble of neural networks //Energy conversion and management. – 2017. – Т. 134. – С. 361-372.

20. Pradeepkumar D., Ravi V. Forecasting financial time series volatility using particle swarm optimization trained quantile regression neural network //Applied Soft Computing. – 2017. – Т. 58. – С. 35-52.

21. Wu X. et al. Data mining with big data //IEEE transactions on knowledge and data engineering. – 2013. – Т. 26. – №. 1. – С. 97-107.

22. Hwang H. T. et al. A parallel computational framework to solve flow and transport in integrated surface–subsurface hydrologic systems //Environmental modelling & software. – 2014. – Т. 61. – С. 39-58.

23. Ahmed R. E., Dhodhi M. K. Directory-based cache coherence protocol for power-aware chip-multiprocessors //2011 24th Canadian conference on electrical and computer engineering (CCECE). – IEEE, 2011. – С. 001036-001039.

24. Cheng C. et al. Parallel discrete differential dynamic programming for multireservoir operation //Environmental modelling & software. – 2014. – Т. 57. – С. 152-164.

25. Tu K. Y., Liang Z. C. Parallel computation models of particle swarm optimization implemented by multiple threads //Expert Systems with Applications. – 2011. – Т. 38. – №. 5. – С. 5858-5866.

26. Subotic M., Tuba M., Stanarevic N. Parallelization of the artificial bee colony (ABC) algorithm //Proceedings of the 11th WSEAS international conference on nural networks and 11th WSEAS international conference on evolutionary computing and 11th WSEAS international conference on Fuzzy systems, World Scientific and Engineering Academy and Society (WSEAS). Stevens Point, Wisconsin, USA, NN. – 2010. – Т. 10. – С. 191-196.

27. Li T. et al. Dynamic parallelization of hydrological model simulations //Environmental Modelling & Software. – 2011. – Т. 26. – №. 12. – С. 1736-1746.

28. Tiwari A., Alam M. A. Implementation of parallel artificial bee colony algorithm on vehicle routing problem //Int. J. Adv. Res. Sci. Eng.(IJARSE). – 2013. – Т. 2. – №. 5.

29. Bryan B. A. High-performance computing tools for the integrated assessment and modelling of social–ecological systems //Environmental Modelling & Software. – 2013. – Т. 39. – С. 295-303.

30. Zhao G. et al. Large-scale, high-resolution agricultural systems modeling using a hybrid approach combining grid computing and parallel processing //Environmental Modelling & Software. – 2013. – Т. 41. – С. 231-238.

31. Huo J., Liu L., Zhang Y. An improved multi-cores parallel artificial Bee colony optimization algorithm for parameters calibration of hydrological model //Future Generation Computer Systems. – 2018. – Т. 81. – С. 492-504.

32. Feng Z. et al. Peak operation of hydropower system with parallel technique and progressive optimality algorithm //International Journal of Electrical Power & Energy Systems. – 2018. – Т. 94. – С. 267-275.

33. Sha F., Zandavi S. M., Chung Y. Y. Fast deep parallel residual network for accurate super resolution image processing //Expert Systems with Applications. – 2019. – Т. 128. – С. 157-168.

34. Madsen A. L. et al. A parallel algorithm for Bayesian network structure learning from large data sets //Knowledge-Based Systems. – 2017. – Т. 117. – С. 46-55.

35. Masood T. B., Ray T., Natarajan V. Parallel computation of alpha complexes for biomolecules //Computational Geometry. – 2020. – Т. 90. – С. 101651.

36. Jin N., Rahmat-Samii Y. Parallel particle swarm optimization and finite-difference time-domain (PSO/FDTD) algorithm for multiband and wide-band patch antenna designs //IEEE Transactions on Antennas and Propagation. – 2005. – Т. 53. – №. 11. – С. 3459-3468.

37. Schutte J. F. et al. Parallel global optimization with the particle swarm algorithm //International journal for numerical methods in engineering. – 2004. – Т. 61. – №. 13. – С. 2296-2315.

38. Parsopoulos K. E. Parallel cooperative micro-particle swarm optimization: A master–slave model //Applied Soft Computing. – 2012. – Т. 12. – №. 11. – С. 3552-3579.

39. Tang Y., Reed P. M., Kollat J. B. Parallelization strategies for rapid and robust evolutionary multiobjective optimization in water resources applications //Advances in Water Resources. – 2007. – Т. 30. – №. 3. – С. 335-353.

40. Li X. et al. Knowledge-based approach for reservoir system optimization //Journal of Water Resources Planning and Management. – 2014. – Т. 140. – №. 6. – С. 04014001.

41. Zhao L. et al. Parallel computing method of deep belief networks and its application to traffic flow prediction //Knowledge-Based Systems. – 2019. – Т. 163. – С. 972-987.

42. He Y. et al. Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory //Applied energy. – 2017. – Т. 185. – С. 254-266.

43. He Y. et al. Electricity consumption probability density forecasting method based on LASSO-Quantile Regression Neural Network //Applied energy. – 2019. – Т. 233. – С. 565-575.

44. Donaldson R. G., Kamstra M. Forecast combining with neural networks //Journal of Forecasting. – 1996. – Т. 15. – №. 1. – С. 49-61.

45. Zeng Y. On MPI-based master-slave parallel task allocation and its implementation //Jisuanji Yingyong yu Ruanjian. – 2010. – Т. 27. – №. 6. – С. 139-141.

46. Altman N. S. An introduction to kernel and nearest-neighbor nonparametric regression //The American Statistician. – 1992. – Т. 46. – №. 3. – С. 175-185.

47. Xu Q. et al. Weighted quantile regression via support vector machine //Expert Systems with Applications. – 2015. – Т. 42. – №. 13. – С. 5441-5451.

48. Epanechnikov V. A. Non-parametric estimation of a multivariate probability density //Theory of Probability & Its Applications. – 1969. – Т. 14. – №. 1. – С. 153-158.

49. Zhang X. et al. Efficient multi-objective calibration of a computationally intensive hydrologic model with parallel computing software in Python //Environmental modelling & software. – 2013. – Т. 46. – С. 208-218.

50. Tesfa T. K. et al. Extraction of hydrological proximity measures from DEMs using parallel processing //Environmental Modelling & Software. – 2011. – Т. 26. – №. 12. – С. 1696-1709.

51. Zhou S., Mao M., Su J. Short-term forecasting of wind power and non-parametric confidence interval estimation //Zhongguo Dianji Gongcheng Xuebao(Proceedings of the Chinese Society of Electrical Engineering). – Chinese Society for Electrical Engineering, 2011. – Т. 31. – №. 25. – С. 10-16.

52. Zhang C. et al. A compound structure of ELM based on feature selection and parameter optimization using hybrid backtracking search algorithm for wind speed forecasting //Energy Conversion and Management. – 2017. – Т. 143. – С. 360-376.

53. Quan H., Srinivasan D., Khosravi A. Uncertainty handling using neural network-based prediction intervals for electrical load forecasting //Energy. – 2014. – Т. 73. – С. 916-925.

54. Khosravi A. et al. Lower upper bound estimation method for construction of neural network-based prediction intervals //IEEE transactions on neural networks. – 2010. – Т. 22. – №. 3. – С. 337-346.

55. Generator output by fuel type hourly report, 2021, Website https://reports.ieso.ca/public/GenOutputbyFuelHourly/PUB\_GenOutputbyFuelHourly\_2021\_v365.xml.

56. Takagi T., Sugeno M. Fuzzy identification of systems and its applications to modeling and control //IEEE transactions on systems, man, and cybernetics. – 1985. – №. 1. – С. 116-132.

57. Yang Z. et al. A novel competitive swarm optimized RBF neural network model for short-term solar power generation forecasting //Neurocomputing. – 2020. – Т. 397. – С. 415-421.

58. He F. et al. Day-ahead short-term load probability density forecasting method with a decomposition-based quantile regression forest //Applied Energy. – 2020. – Т. 262. – С. 114396.

59. Hedrea E. L., Precup R. E., Bojan-Dragos C. A. Results on tensor product-based model transformation of magnetic levitation systems //Acta Polytechnica Hungarica. – 2019. – Т. 16. – №. 9. – С. 93-111.

60. Zuo H. et al. Fuzzy regression transfer learning in Takagi–Sugeno fuzzy models //IEEE Transactions on Fuzzy Systems. – 2016. – Т. 25. – №. 6. – С. 1795-1807.