

Метод прогнозирования выработки электроэнергии сверхкраткосрочным кластером ветроэнергетики на основе алгоритма конвергентного перекрестного отображения

Yuzhe Yang^a, Weiye Song^a, Shuang Han^{a,*}, Jie Yan^a, Han Wang^a, Qiangsheng Dai^b, Xuesong Huo^b, Yongqian Liu^a

^a Северо-Китайский электроэнергетический университет, Пекин 102206, КНР

^b Государственная электроэнергетическая компания Цзянсу, Цзянсу 210018, КНР

Аннотация Развитие кластеров ветроэнергетики значительно расширилось как по масштабу, так и по охвату, а влияние колебаний погоды на изменения выработки кластера стало все более сложным. Точное определение прогнозной информации о ключевых ветропарках кластера при различных погодных условиях является эффективным методом повышения точности сверхкраткосрочного прогнозирования выработки электроэнергии кластером. С этой целью в данной статье предлагается усовершенствованный метод моделирования для сверхкраткосрочного прогнозирования выработки электроэнергии кластером ветроэнергетики на основе алгоритма конвергентного перекрестного отображения. С точки зрения причинно-следственной связи были отобраны ключевые метеорологические факторы прогнозирования при различных процессах колебаний выработки электроэнергии кластером, и было выполнено усовершенствованное моделирование для различных процессов колебаний. Во-первых, создана система индексов описания ветровых процессов и классификационная модель на уровне кластера ветроэнергетики для реализации классификации типичных процессов колебаний. Предложена модель оценки причинно-следственной связи между мощностью метеорологического кластера и его влиянием, основанная на алгоритме конвергентного перекрестного отображения, для анализа факторов метеорологического прогнозирования в условиях различных типов типичных флуктуационных процессов. В заключение предложен усовершенствованный метод моделирования для различных типичных флуктуационных процессов, и сильные причинно-следственные факторы метеорологического прогнозирования каждого сценария используются в качестве входных данных для реализации высокоточного моделирования и прогнозирования сверхкраткосрочной мощности ветрового кластера. Пример анализа показывает, что точность краткосрочного прогнозирования мощности ветрового кластера с помощью предложенного метода может достигать 88,55%, что на 1,57-7,32% выше, чем у традиционных методов. ©2025 Global Energy Interconnection Group Co.Ltd. Издательские услуги предоставлены Elsevier B.V. от имени KeAi Communications Co.Ltd. Эта статья имеет открытый доступ, по лицензии CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Ключевые слова: Сверхкраткосрочное прогнозирование выработки ветровой энергии; Кластер ветроэнергетики; Анализ причинно-следственных связей; Алгоритм конвергентного перекрестного отображения

0 Введение

В рамках общей концепции «углеродного пика и углеродной нейтральности» ветроэнергетика стала важнейшим направлением развития благодаря своим преимуществам, включая устранение необходимости в передаче электроэнергии на большие расстояния, снижение стоимости земли и высокую социальную приемлемость.

Прогнозирование выработки ветровой энергии (WPF) является важным аспектом подключения ветропарков к сети, и его точность имеет решающее значение для повышения безопасности и экономической эффективности работы энергосетей [1]. Однако по мере расширения пространственного охвата кластеров ветроэнергетики входные данные, необходимые для прогнозирования на уровне площадки, не только динамически изменяются в зависимости от

колебаний скорости ветра, но и демонстрируют повышенную изменчивость в своей эффективности для общей мощности кластера, что усложняет сверхкраткосрочное прогнозирование мощности кластера ветроэнергетики [2,3].

Динамические изменения и растущее несоответствие эффективности входных данных прогнозирования на уровне площадки, в сочетании с изменяющимися процессами колебаний скорости ветра, представляют собой значительные ограничения для дальнейшего повышения точности прогнозирования. Следовательно, существует острая необходимость в высокоточном прогнозировании выработки электроэнергии для предоставления надежной информации о колебаниях выходной мощности [4]. Методы прогнозирования, основанные на статистических модельных структурах, таких как те, которые используют глубокое обучение, приобрели известность в последние годы благодаря постоянному расширению хранилища данных об эксплуатации ветроэнергетических установок и быстрому развитию технологий искусственного интеллекта [5-8].

Современные исследования методов прогнозирования на основе глубокого обучения можно условно разделить на два подхода. Первый фокусируется на улучшении механизмов сетей глубокого обучения для повышения их способности извлекать временные ряды из данных. Второй подход направлен на выявление характеристик процессов флуктуаций и сценариев выходных данных, их соответствующую классификацию, а затем создание моделей прогнозирования для каждой категории погодных процессов или индивидуальную коррекцию ошибок прогнозирования [9,10].

Первый метод [11] интегрирует метод взаимной информации с методом ложного ближайшего соседа для восстановления фазового пространства последовательностей ветровой энергии, предлагая инновационную структуру прогнозирования хаотических временных рядов с использованием механизма переключения ядерной функции. В [12] были объединены сверточные нейронные сети (CNN) и модели долговременной кратковременной памяти (LSTM), где CNN извлекали пространственные характеристики долговременных скоростей ветра, а LSTM фокусировались на краткосрочных характеристиках временных рядов, что позволяло прогнозировать скорость ветра в нескольких местах. В [13] точность прогнозирования скорости ветра была улучшена за счет уточнения модели LSTM-TCN с учетом эффектов турбулентного следа турбины, состояния агрегата и пространственного распределения на ветровых электростанциях. Второй тип метода [14] использует вектор скорости ветра и ежедневное изменение давления в численном прогнозировании погоды (NWP) в качестве основы для кластерного анализа для классификации типов погоды и создания моделей прогнозирования для различных типов погоды с целью повышения точности прогнозирования. В [15] был предложен метод краткосрочного прогнозирования ветровой энергии, сочетающий временные окна переменного масштаба и извлечение признаков флуктуаций. Он использует алгоритм скользящего окна переменного масштаба для динамического извлечения признаков на основе текущих характеристик скорости ветра, классификации исторических данных о ветровой энергии и выбора конкретных параметров для построения моделей прогнозирования. В [16] был представлен сегментированный метод краткосрочного прогнозирования для кластеров ветровой энергии путем адаптивного разделения переходных погодных периодов. Для повышения адаптивности прогнозирования ветровой энергии в экстремальных погодных условиях были применены различные методы прогнозирования для переходных и непереходных периодов.

В литературе рассматриваются различные подходы к повышению точности прогнозирования мощности в сверхкраткосрочной перспективе. Однако анализ взаимосвязей между метеорологическими факторами и изменениями мощности кластеров преимущественно основан на линейном или нелинейном корреляционном анализе, который дает симметричные, но не направленные результаты. Следовательно, модели сверхкраткосрочного прогнозирования,

основанные на сильно коррелированных метеорологических факторах, часто демонстрируют низкую интерпретируемость и ограниченную обобщающую способность [17-20].

В данном исследовании предлагается применять анализ причинно-следственных связей на основе алгоритма конвергентного кросс-картирования (CCM) для количественной оценки причинно-следственной связи между мощностью кластера и множеством метеорологических факторов. По сравнению с корреляционным анализом, причинно-следственный анализ глубже изучает статистические и физические взаимосвязи между изменениями мощности и метеорологическими факторами, подчеркивая направленность извлечения признаков временных рядов. Такой подход значительно снижает влияние избыточных факторов во время обучения и моделирования, тем самым существенно повышая точность прогнозирования мощности. Кроме того, учитывая различные причинно-следственные связи между множеством метеорологических факторов и колебаниями общей мощности кластера при различных процессах флуктуаций, разработана кластерная модель для проведения кластерного анализа при различных процессах флуктуаций. Эта модель точно определяет факторы прогнозирования при различных процессах флуктуаций. На основе алгоритма CCM в исследовании представлена усовершенствованная методика моделирования для сверхкраткосрочного прогнозирования мощности кластера ветровой энергии.

Основные преимущества этого метода заключаются в следующем:

- 1) Предлагается система индексов описания ветровых процессов на уровне кластера и метод классификации, учитывающие характеристики скорости и направления ветра в нескольких точках внутри кластера ветроэнергетических установок. Эта система характеризует изменение выходной мощности кластера при различных ветровых процессах, закладывая основу для усовершенствованного моделирования прогнозирования выработки электроэнергии кластером ветропарков.
- 2) Предлагается метод анализа причинно-следственных связей «многомерные колебания суммарной мощности кластеров метеорологических факторов прогнозирования», интегрирующий алгоритм CCM. Этот метод количественно оценивает причинно-следственную связь между различными метеорологическими факторами прогнозирования и мощностью кластеров, подчеркивая направленность извлечения признаков временных рядов. Он преобразует традиционное моделирование на основе корреляций в глубокое моделирование причинно-следственных связей, тем самым повышая устойчивость модели.
- 3) Предлагается усовершенствованный подход к моделированию прогнозирования выработки электроэнергии в кластерах ветроэнергетики, учитывающий процессы флуктуаций. Этот подход использует сильные причинно-следственные метеорологические факторы, выявленные в рамках различных типичных процессов флуктуаций, в качестве входных данных для разработки целевых моделей прогнозирования для этих процессов, тем самым повышая уровень точности моделей прогнозирования.
- 4) Используя данные кластера морской ветроэнергетики в Китае, была проведена сравнительная оценка и тестирование точности прогнозирования различных моделей глубокого обучения, применяющих предложенный метод. Результаты анализа примеров показывают, что предложенный метод повышает точность прогнозирования на 1,57% по сравнению с традиционными методами.

1 Методология

Основное содержание данного исследования сосредоточено на четырех исследовательских задачах: кластерный анализ исторических метеорологических данных, причинно-следственный анализ на основе алгоритма CCM, создание модели сверхкраткосрочного прогнозирования

энергопотребления и анализ ошибок результатов прогнозирования. Общий процесс данного исследования проиллюстрирован на рис. 1.

В данном исследовании использовались данные за прошедшие периоды по кластеру из пяти ВЭС за последние три месяца, включая скорость ветра, направление ветра, мощность каждой ВЭС и суммарную мощность кластера.

1.1 Метод классификации стандартных флуктуационных процессов

В этом разделе рассматривается метод классификации стандартных флуктуационных процессов с многовременными и пространственными характеристиками. Сначала предлагается система индексов описания флуктуационных процессов для колебаний мощности. Для кластеризации колебаний мощности в рамках различных флуктуационных процессов использовались пять показателей. Наконец, получены несколько типичных флуктуационных процессов со схожими характеристиками колебаний мощности, и обобщены несколько типичных характеристик изменения выходного сигнала.

В этом разделе в первую очередь представлено исследование по установлению типичных процессов колебаний. Кластерный анализ был выполнен на основе набора данных о характеристиках колебаний, и из большого числа процессов колебаний было выделено несколько типичных процессов колебаний. На рис. 2 показана блок-схема построения набора данных о характеристиках колебаний погоды.

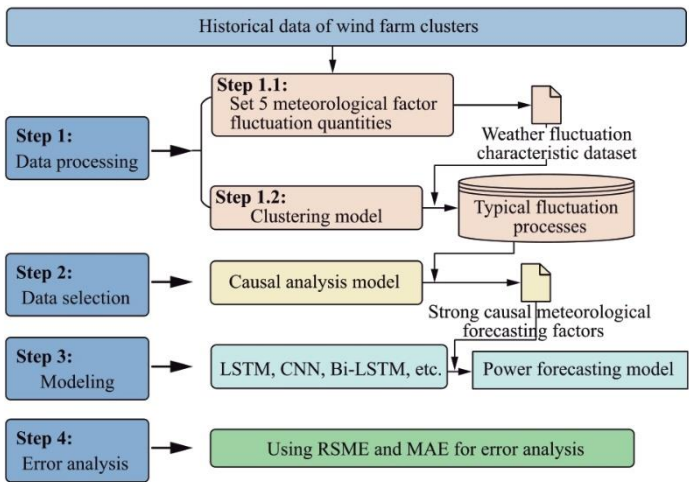


Рис. 1. Общая блок-схема данного исследования.

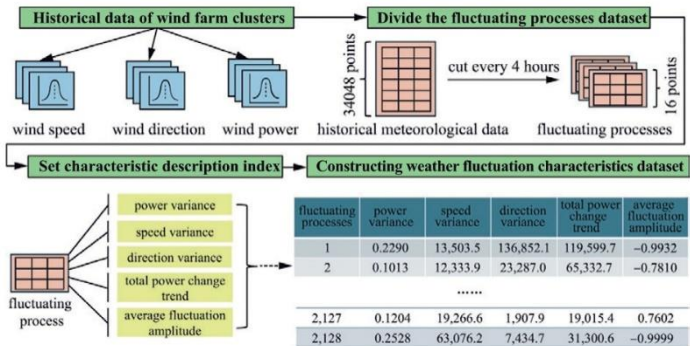


Рис. 2. Блок-схема построения набора данных характеристик колебаний погоды.

Стандарт для сверхкраткосрочного прогнозирования мощности предполагает прогнозирование мощности кластера с 15-минутными интервалами в течение следующих 4 часов [21]. Колебания метеорологических данных и мощности в течение этого 4-часового периода считаются процессом

флуктуаций. Данные временных рядов были разделены на 4-часовые сегменты, причем каждый сегмент представлял собой процесс флуктуаций, включающий исторические метеорологические данные.

После завершения разделения процессов флуктуаций необходимо количественно проанализировать характерные величины каждого процесса флуктуаций. Цель данного исследования — описать характеристики флуктуаций ветроэнергетических кластеров во времени и пространстве для каждого процесса флуктуаций на основе причинно-следственной связи между множеством метеорологических факторов и общей мощностью кластера. Пять ветропарков, входящих в ветроэнергетический кластер, имеют коррелированные скорости и направления ветра. Степень флуктуаций в пространстве характеризуется дисперсией скорости и направления ветра, а степень флуктуаций во времени — дисперсией общей мощности ветроэнергетического кластера. Тенденция изменения общей мощности и средняя амплитуда колебаний используются для характеристики тенденции и амплитуды выходных данных, а также для обобщения временных динамических изменений характеристик выходных данных. Пять величин колебаний метеорологических факторов, а именно «дисперсия скорости ветра, дисперсия направления ветра, дисперсия общей мощности, тенденция изменения общей мощности и средняя амплитуда колебаний [22]», устанавливаются для количественной характеристики пространственных и временных динамических изменений множества метеорологических факторов и общей мощности кластера в процессе колебаний. Эти пять величин колебаний использовались в качестве характеристических величин для кластерного анализа каждого процесса колебаний. Формула выглядит следующим образом:

Можно рассчитать скорость ветра, дисперсию направления ветра и значение дисперсии суммарной мощности кластера в каждом процессе флуктуаций, а также три значения дисперсии

σ_s^2, σ_d^2 , и σ_p^2 , соответственно, их можно получить следующим образом:

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N} \quad (1)$$

где N — общее количество данных о скорости/направлении/мощности ветра за время колебания; $i = 1, 2, \dots, N$ — временная точка колебания; x_i — измеренные данные о скорости/направлении/мощности ветра в i -й момент времени; μ — среднее значение данных о скорости/направлении/мощности ветра за время колебания.

Тенденция изменения суммарной мощности P_x общей мощности кластера в каждом процессе флуктуаций рассчитывается следующим образом:

$$P_x = \text{Max}(P) - \text{Min}(P) \quad (2)$$

где P — суммарная мощность кластера ветропарков во время колебаний; Max — функция максимума; Min — функция минимума.

Средняя амплитуда флуктуаций $\Delta P(t+\Delta t)$ в каждом процессе флуктуаций может быть количественно выражена следующим образом:

$$\Delta P(t + \Delta t) = P(t + \Delta t) - P(t) \quad (3)$$

где t — момент времени записи измеренных данных; Δt — выбранное временное окно (в данном исследовании выбрано временное окно в 4 часа); $P(t)$ — суммарная мощность кластера ветропарков во время колебаний в момент времени t .

Для каждого процесса колебаний рассматриваются пять величин, характеризующих колебания метеорологических факторов, а именно: σ_s^2 , σ_d^2 , σ_p^2 , P_x , и $\Delta P(t+\Delta t)$, соответствующие n процессам флуктуаций рассчитываются с использованием приведенной выше формулы. После удаления аномальных данных в качестве меток столбцов были использованы пять величин флуктуаций метеорологических факторов, и был получен набор данных характеристик флуктуаций погоды с пятью столбцами и n строками.

Кластеризация используется для определения внутренних связей между данными [23]. Используя кластерный анализ, различные данные могут быть разделены на несколько кластеров в соответствии с характерными значениями каждого набора данных. Хотя сопоставимые данные между различными кластерами значительно различаются, данные внутри кластера схожи.

Алгоритм кластеризации К-средних — это типичный алгоритм разбиения, используемый в кластерном анализе [24,25]. Принцип работы алгоритма следующий: изначально k центральных точек выбираются случайным образом из набора признаков. Затем вычисляются евклидовы расстояния между всеми точками признаков и этими центральными точками, после чего каждая точка признака присваивается кластеру ближайшей центральной точки, и в каждом кластере снова вычисляется новая центральная точка. Этот процесс повторяется многократно до тех пор, пока вычисленная центральная точка после каждого разбиения перестанет изменяться в допустимом диапазоне, и кластерный анализ можно считать завершенным [26].

Целевая функция E в процессе разделения данных и координаты новой центральной точки c_i рассчитывались следующим образом:

$$E = \sum_{i=1}^k \sum_{s_x \in S_i} \|s_x - c_i\|^2 \quad (4)$$

$$c_i = \frac{1}{|S_i|} \sum_{s_x \in S_i} s_x \quad (5)$$

где i — номер кластера; k — общее количество кластеров; S_i — множество, содержащее наборы всех характерных точек i -го кластера; s_x — координаты одной характерной точки.

При работе с большими объемами данных подход К-средних предлагает такие преимущества, как простота и высокая скорость вычислений, и обычно считается предпочтительным алгоритмом кластеризации [27]. В данном исследовании флуктуации были классифицированы с использованием метода k -средних.

1.2 Анализ причинно-следственных связей на основе алгоритма CCM

Учитывая причинно-следственную связь между множеством метеорологических факторов и колебаниями общей мощности кластера, в этом разделе предлагается метод анализа причинно-следственных связей, основанный на алгоритме CCM. С учетом вышеописанных типичных процессов колебаний, была проведена тщательная оценка интенсивности причинно-следственной связи между метеорологическими факторами и общей мощностью кластера в каждом типичном процессе колебаний, и с помощью алгоритма CCM была выявлена сильная причинно-следственная связь между метеорологическими факторами.

В отличие от метода корреляционного анализа [28], метод анализа причинно-следственных связей направлен на определение причины результата и последующее построение направленной причинно-следственной сетевой структуры между несколькими переменными системы. На рис. 3 показана причинно-следственная диаграмма между многомерными переменными временных рядов; то есть, посредством анализа причинно-следственных связей между несколькими переменными можно получить корреляционно-причинно-следственную диаграмму. С помощью

анализа причинно-следственных связей можно эффективно исключить избыточные и результирующие переменные; то есть, метеорологические факторы, не связанные с общей мощностью кластера, могут быть эффективно исключены для отсеивания причинных переменных в процессе колебаний мощности ветра и проведения более качественной проверки данных для последующего прогнозирования мощности [29]. В настоящее время анализ причинно-следственных связей применяется во многих областях и дает хорошие результаты.

В анализе причинно-следственных связей обычно используются алгоритмы, основанные на информационной энтропии, метод причинности Грейнджера, алгоритм ССМ и т. д. В данном исследовании, исходя из больших временных рядов данных с множеством метеорологических факторов и общей мощностью кластера, был применен алгоритм ССМ. Методы анализа причинно-следственных связей на основе причинности Грейнджера и информационной энтропии требуют большого объема высококачественных данных для эффективного анализа причинно-следственных связей. Более того, метод причинности Грейнджера может давать только качественные результаты. Однако алгоритм ССМ лучше адаптируется к искажённым данным [30] и обладает сильной способностью к поиску причинно-следственных связей [31]; поэтому он больше подходит для анализа причинно-следственных связей соответствующих данных в представленном исследовании.

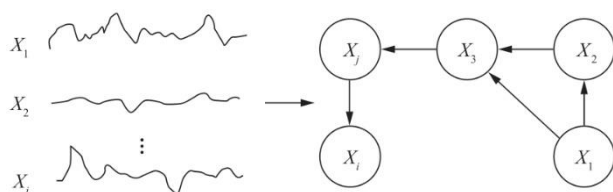


Рис. 3. Причинно-следственная диаграмма переменных многомерных временных рядов.

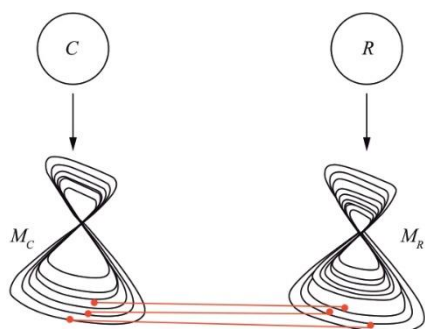


Рис. 4. Схема алгоритма ССМ.

Алгоритм ССМ основан на всей теореме вложения и теновом многообразии. Как показано на рис. 4, в качестве примера используются две переменные, C и R . Согласно теореме вложения Такенса, вектор со значением задержки может быть использован для генерации теневого многообразия истинного многообразия. Точки в теновом многообразии демонстрируют однозначное соответствие с точками в истинном многообразии [32]. Следовательно, R может быть использован для построения теневого многообразия M_R , а C может построить теновое многообразие M_C . Если C и R имеют причинно-следственную связь, то соседние точки, окружающие точку в M_R , будут более точно идентифицировать соседние точки соответствующей точки в M_C . По мере увеличения длины временного ряда теновые многообразия M_C и M_R становятся «плотнее», что приводит к большему количеству «компактных» соседних точек при идентификации точки. Следовательно, ошибка идентификации уменьшается, а коэффициент корреляции между истинным значением переменной C и оценочным значением, полученным с помощью кросс-отображения, выше.

В этом методе суммарная мощность кластера и различных метеорологических факторов представляет собой два временных ряда $x(t)$ и $y(t)$ длиной L , которые можно рассматривать как два

теневых многообразия, полученных путем проецирования системы М в одномерное пространство. Предполагая, что размерность D теневого многообразия равна 2, а интервал дискретизации τ равен 1, тогда два временных ряда могут быть реконструированы в два пространства состояний, а векторы состояний $X(t)$ и $Y(t)$ в момент времени t в двух пространствах состояний могут быть выражены следующим образом:

$$X(t) = [x(t), x(t - \tau), \dots, x(t - (D - 1)\tau)] \quad (6)$$

$$Y(t) = [y(t), y(t - \tau), \dots, y(t - (D - 1)\tau)] \quad (7)$$

где D — размерность вложения; τ — интервал дискретизации; $x(t)$ — значение метеорологического временного ряда $x(t)$ в момент времени t ; $y(t)$ — значение временного ряда полной мощности $y(t)$ в момент времени t .

Тогда два вектора состояния $X(t)$ и $Y(t)$ во все моменты времени образуют два пространства состояний, которые можно рассматривать как два теневого многообразия M_x и M_y проекции системы М в одномерное пространство.

Согласно теореме Такенса о вложении, для любой точки на теновом многообразии в каждом измерении можно найти число D смежных точек. Множество $G_{x,i,D}$ всех смежных точек для точки $X(t_i)$ на теновом многообразии M_x в момент времени t_i можно выразить следующим образом:

$$G_{x,i,D} = \{X(t_i)_1, X(t_i)_2, \dots, X(t_i)_D\} \quad (8)$$

где t_i — i -я точка временного ряда; $X(t_i)_D$ — смежная точка $X(t_i)$ на измерении с номером D .

Из смежных точек в $G_{x,i,D}$, указанных выше, мы можем отобразить эти смежные точки на теновое многообразие M_y и получить множество точек $G_{y,i,D}$ следующим образом:

$$G_{y,i,D} = \{Y(t_i)_1, Y(t_i)_2, \dots, Y(t_i)_D\} \quad (9)$$

где $Y(t_i)_1, Y(t_i)_2, \dots, Y(t_i)_D$ это точки, полученные в результате построения карты $X(t_i)_1, X(t_i)_2, \dots, X(t_i)_D$ в M_y .

Для точки $Y(t_i)$ на теновом многообразии M_x в момент времени t_i , соответствующий $X(t_i)$, мы можем объединить весовой коэффициент w_i и набор точек $G_{y,i,D}$ для вычисления оценочного значения $^Y(t_i)|M_y$ следующим образом:

$$\hat{Y}(t_i)|M_y = \sum_{j=1}^D w_j Y(t_i)_j \quad (10)$$

$$m_j = \exp \left\{ - \frac{\|X(t_i) - X(t_i)_j\|}{\|X(t_i) - X(t_i)_1\|} \right\} \quad (11)$$

$$w_j = \frac{m_j}{\sum_{j=1}^D m_j} \quad (12)$$

где j — номер смежной точки, причем $j=1, 2, \dots, D$; L — длина временного ряда; i — номер временной точки, причем $i=1, 2, \dots, L$; t_i — i -я временная точка временного ряда; w_i — весовой коэффициент; m_i — промежуточная переменная; $X(t_i)_j$ — смежная точка $X(t_i)$ в измерении с номером j .

Согласно приведенному выше методу расчета, мы можем получить оценочное значение $^Y(t_i)|M_y$ в любой момент времени $t_i \in t$. Определим значение CCM $rx \rightarrow y$ от $x(t)$ до $y(t)$ следующим образом:

$$r_{x \rightarrow y} = \frac{\sum_{i=1}^L (Y(t_i) - \bar{Y}) (\hat{Y}(t_i) - \bar{\hat{Y}})}{\sqrt{\sum_{i=1}^L (Y(t_i) - \bar{Y})^2 \sum_{i=1}^L (\hat{Y}(t_i) - \bar{\hat{Y}})^2}} \quad (13)$$

где L — длина временного ряда; i — номер временной точки, и

$$\hat{Y}(t_i) | M_y \quad \hat{Y}(t_i); \bar{Y}(t_i) \quad \text{сокращенно до} \quad \hat{Y}(t_i) \quad \text{является}$$

$i=1, 2, \dots, L;$

средним значением $\bar{Y}(t_i)$; $\hat{Y}(t_i)$ является средним значением $\hat{Y}(t_i)$.

Если существует причинно-следственная связь от $x(t)$ к $y(t)$, то будет получено пересечение $g_{x \rightarrow y}$. Диапазон значений этого параметра — $(0, 1)$, и чем больше значение $g_{x \rightarrow y}$, тем сильнее причинно-следственная связь.

В соответствии с этим принципом, на основе количественного значения можно выбрать метеорологические факторы с сильной причинно-следственной связью в различных типах флуктуационных процессов в качестве входных данных для последующего моделирования сверхкраткосрочного прогнозирования мощности.

1.3 Метод сверхкраткосрочного прогнозирования выработки ветровой энергии для многомерных процессов колебаний погоды

Учитывая различия в взаимосвязях отображения метеорологической мощности при различных процессах флуктуаций [33], этот метод сначала классифицирует данные по процессам флуктуаций с многовременными и пространственными масштабами, а затем проводит анализ причинно-следственных связей между метеорологическими факторами в нескольких точках и общей мощностью кластера на основе ССМ. Обучающие и тестовые наборы были созданы на основе данных о метеорологических факторах с сильными причинно-следственными связями, а для моделирования различных процессов флуктуаций использовались модели глубокого обучения. Наконец, для оценки результатов прогнозирования использовались среднеквадратичная ошибка (RMSE) и средняя абсолютная ошибка (MAE). На рис. 5 и в таблице 1 показан весь процесс уточненного моделирования сверхкраткосрочного прогнозирования мощности.

Поскольку данные, используемые в этом исследовании, включают 6 типов типичных флуктуационных процессов, соответствующие данные о сильных причинно-следственных метеорологических факторах являются первым входным параметром для различных типичных флуктуационных процессов, и входные данные разделены на два набора; впоследствии модель углублённого обучения используется для обучения и тестирования модели для 6 типичных флуктуационных процессов, что включает в себя настройку количества шагов обучения, функции потерь и других шагов; после завершения процесса моделирования для 6 типов моделей, в конечном итоге получается уточненная модель сверхкраткосрочного прогнозирования мощности.

В данном исследовании в качестве входных данных используются скорость и направление сильного причинного ветра, а также историческая суммарная мощность, а в качестве выходных данных — прогнозируемая мощность. Таким образом, количество входных слоев установлено равным 3, а количество выходных слоев — 1. Во время обучения модели мы можем установить функцию потерь J , а процесс обратного распространения ошибки через J выглядит следующим образом:

$$J = \frac{1}{2} (p_{\text{pred}} - p)^2 \quad (14)$$

$$p_{\text{pred},k} = f \left[\sum_{i=1}^m \omega_i f \left(\sum_{j=1}^n v_{ji} x_j \right) \right] \quad (15)$$

$$\begin{cases} \Delta \omega_i = -\eta \frac{\partial J}{\partial \omega_i} \\ \Delta v_{ij} = -\eta \frac{\partial J}{\partial v_{ij}} \end{cases} \quad (16)$$

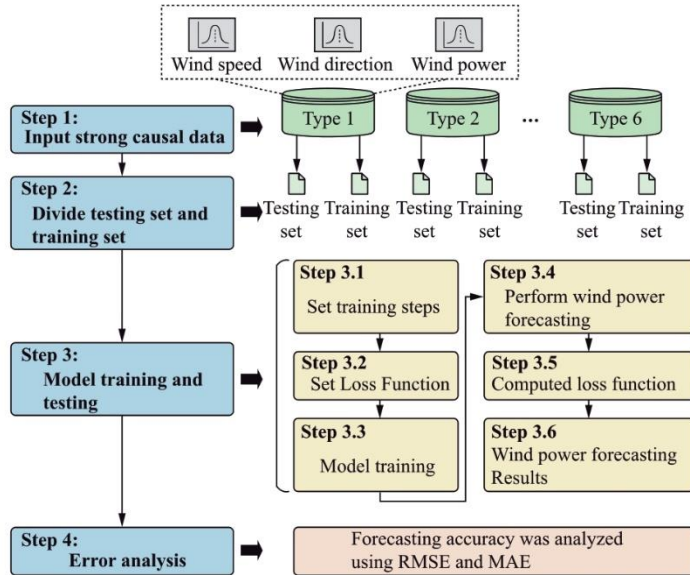


Рис. 5. Процесс уточненного моделирования сверхкраткосрочного прогнозирования энергопотребления.

Таблица 1. Методология уточненного моделирования сверхкраткосрочного прогнозирования энергопотребления.

Input:

X : All historical data of wind Farm clusters,
 $X = \{x(t)_1, x(t)_2, \dots, x(t)_5, x(t)_{all}\}$. Where $x(t)_1, x(t)_2, \dots, x(t)_5$ is the time series of all historical wind speed and wind direction of the wind farm numbered 1, 2, ..., 5 in the wind farm cluster; $y(t)_{all}$ is the time series of all total power of the wind farm cluster;

k : The number of the fluctuation process;

X_{rd} : Randomly selected historical meteorological factors,
 $X_{rd} = \{x(t_r)_1, x(t_r)_2, \dots, x(t_r)_5, y(t)_{rd}\}$. Where $x(t_r)_1, x(t_r)_2, \dots, x(t_r)_5$ is the time series of historical wind speed and wind direction of the wind farm numbered 1, 2, ..., 5 in the wind farm cluster every 15 minutes within 4 hours before the forecast period; $y(t)_{rd}$ is the time series of total power of the wind farm cluster.

Output:

$model_{ccm}$: Power forecasting refinement model based on CCM,
 $model_{ccm} = \{model_1, model_2, \dots, model_k\}$;

p_{pred} : Power forecast results.

```
0  Strat
1  Input:  $X, k, X_{rd}$ 
2   $S_{index} = \{S_1, S_2, S_3, \dots, S_k\}$  is the fluctuation process characteristic
   data set based on five indicators
3   $X \rightarrow \{X_1, X_2, X_3, \dots, X_k\}$  is divided into  $k$  categories by clustering
   based on  $S_{index}$ 
4  For  $i: 1$  to  $k$ 
5       $x(t)_{i,CCM} = CCM(x(t) = x(t)_1, x(t)_2, \dots, x(t)_5, y(t) = y(t)_{all})$ 
       is the strong causal meteorological factors
6      Divide  $x(t)_{i,CCM}$  into training set:  $x(t)_{i,train}$  and testing set:
        $x(t)_{i,test}$ 
7       $model_i = \text{train model}(x(t)_{i,train}, x(t)_{i,test})$ 
8       $s = [\sigma_s^2, \sigma_d^2, \sigma_p^2, P_x, \Delta P(t + \Delta t)]$  is the fluctuation process
       characteristic from  $X_{rd}$ 
9      For  $i: 1$  to  $k$ 
10         If  $s \in S_i$ 
11              $p_{pred} = model_i(x(t)_{i,CCM})$ 
12 Output:  $model_{ccm}, p_{pred}$ 
13 End
```

где p — измеренная суммарная мощность; p_{pred} — результат прогнозирования мощности, рассчитанный моделью; f — функция активации; m — количество скрытых слоев; n — количество входных слоев; w_i — матрица весов от i -го скрытого слоя к выходному слою; u_j — матрица весов от j -го входного слоя к i -му скрытому слою; x_j — временной ряд входных метеорологических факторов j -го входного слоя; Δw_i и Δu_j — модифицированные значения w_i и u_j ; η — скорость обучения; $\partial J / \partial w_i$ и $\partial J / \partial u_j$ — производная функции потерь.

С помощью указанной выше функции потерь матрица весов многократно обновляется, и в конечном итоге получается оптимальный результат прогнозирования, после чего завершается обучение модели.

В данном исследовании для формирования модели сверхкраткосрочного прогнозирования выработки ветровой энергии для шести типичных процессов колебаний были использованы шесть моделей глубокого обучения: LSTM, Bi-LSTM, CNN, N-Beats, DLinear и Трансформер. Каждая модель имеет следующие основные концепции:

1) Сеть краткосрочной и долгосрочной памяти (LSTM)

LSTM — важный метод для DL LSTM, представляющий собой особый тип RNN; LSTM использует временную последовательность входных данных для вычислений и анализа [3,36-38]. Подобно

RNN, LSTM также имеет цепную структуру, как показано на рис. 6, в качестве блока нейронной сети LSTM.

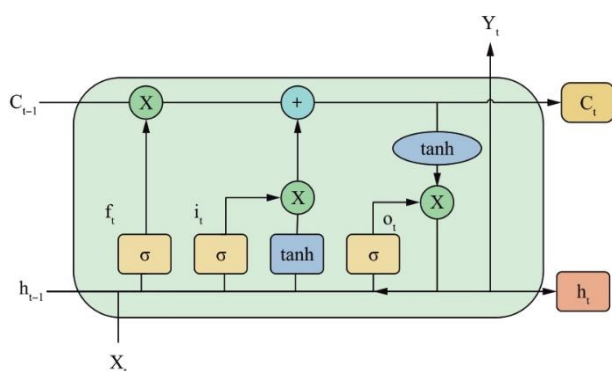


Рис. 6. Блок нейронной сети LSTM.

Компонентами блока LSTM являются вентиль забывания, входной вентиль и выходной вентиль. В дополнение к этим четырем структурам вентилей, LSTM также добавляет состояние нейрона, которое может быть выражено с помощью вектора, отображающего «память» на входную информацию после момента времени t . Этот вектор включает в себя «сумму» всей предыдущей входной информации нейронной сети до этого момента. Это позволяет предотвратить долговременную зависимость, вызванную длинными временными рядами, что приводит к колоссальным ошибкам в результатах прогнозирования.

2) Двухнаправленная долговременная кратковременная память (Bi-LSTM)

Однослойная Bi-LSTM состоит из прямой и обратной LSTM. На рис. 7 показана работа модели Bi-LSTM. По сравнению с моделью LSTM, Bi-LSTM может одновременно получать предыдущую и последующую информацию и интегрировать информацию после всех временных шагов, что делает результаты более полными.

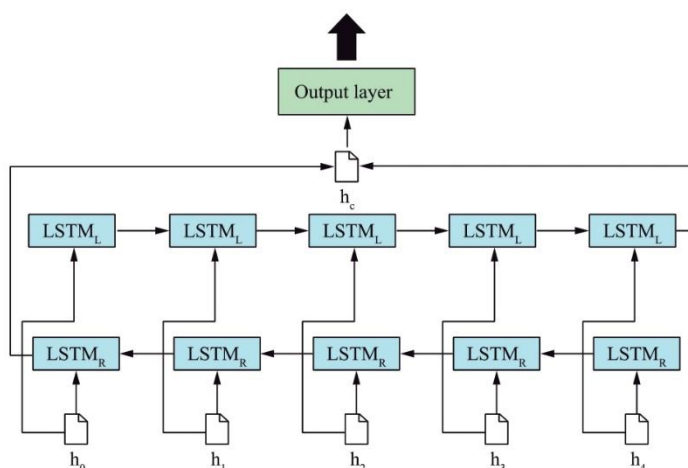


Рис. 7. Схема работы Bi-LSTM.

3) Сверточные нейронные сети (CNN)

По сравнению с традиционными полносвязанными нейронными сетями, сверточные нейронные сети обладают более мощными возможностями извлечения признаков и пространственной инвариантностью. Сверточные слои, функции активации, слои пулинга, полносвязанные слои и нормализующие слои являются одними из наиболее распространенных компонентов. Основным компонентом сверточной нейронной сети является сверточный слой. Как показано на рис. 8,

сверточный слой генерирует карты признаков путем скольжения по входным данным и выполнения операций скалярного произведения. Эти карты признаков сохраняют пространственные взаимосвязи и важные характеристики входных данных.

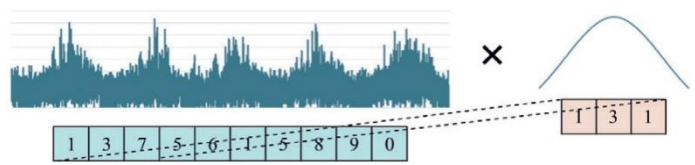


Рис. 8. Одномерная свертка.

4) Анализ расширения нейронных базисов (N-Beats)

Модель N-Beats состоит из нескольких последовательно соединенных нейронных сетей прямого распространения, каждая из которых называется «блоком». Каждый блок содержит четыре полносвязных слоя. Входной сигнал модели x_ℓ проходит через четыре полносвязных слоя и

линейный слой LINEAR для получения двух выходных сигналов θ_ℓ^b и θ_ℓ^f . Из них, θ_ℓ^f — это прогнозируемое значение, которое используется для получения окончательного результата

прогнозирования, и θ_ℓ^b — это значение обратного прогнозирования, которое используется для

представления части временного ряда, не объясняемой моделью. θ_ℓ^b будет передано в качестве входных данных следующему блоку, что позволит постепенно корректировать модель и повышать точность прогнозирования. Формула, используемая для расчета каждого полносвязного слоя, выглядит следующим образом.:

$$h_{\ell,i} = FC_{\ell,i}(w_{\ell,i}h_{\ell,i-1} + b_{\ell,i}) \quad (17)$$

где $FC_{\ell,i}$ — функция активации; ℓ — номер блока; i — количество полносвязных слоев в каждом блоке, $i=1,2,3,4$; $w_{\ell,i}$ — вектор матрицы весов; $h_{\ell,i}$ — результат вычислений для i -го блока; $b_{\ell,i}$ — вектор отклонений.

После четырех полносвязных слоев конечный выходной сигнал равен $h_{\ell,4}$. В качестве входных данных для линейного слоя используется следующий алгоритм вычислений:

$$\begin{cases} \theta_\ell^b = w_\ell^b h_{\ell,4} \\ \theta_\ell^f = w_\ell^f h_{\ell,4} \end{cases} \quad (18)$$

где w_ℓ^b и w_ℓ^f это вектор матрицы весов; θ_ℓ^b и θ_ℓ^f - коэффициент разложения.

θ_ℓ^b и θ_ℓ^f являются входными данными для двух функций активации базового слоя

q_{ℓ}^b и q_{ℓ}^f для расчета. Формула следующая:

$$\begin{cases} \hat{x}_{\ell} = q_{\ell}^b(\theta_{\ell}^b) \\ \hat{y}_{\ell} = q_{\ell}^f(\theta_{\ell}^f) \end{cases} \quad (19)$$

где $\hat{x}_{\ell}$ является результатом обратного прогнозирования; $\hat{y}_{\ell}$ является результатом прогнозирования будущего значения.

После обработки нескольких блоков конечный результат работы N-Beats представляет собой

сумму каждого из них $\hat{y}_{\ell}$. Формула представлена ниже:

$$\hat{y} = \sum_{\ell} \hat{y}_{\ell} \quad (20)$$

где $\hat{y}$ это конечный результат N-Beats; ℓ — количество блоков.

5) DLinear

Структура DLinear основана на линейной регрессии и состоит из нескольких линейных слоев, расположенных друг над другом. Каждый слой выполняет линейное преобразование входных данных, что позволяет модели улавливать сложные линейные зависимости во временном ряду, сохраняя при этом интерпретируемость линейных моделей.

По сравнению с обычными глубокими нейронными сетями, структура DLinear обеспечивает большую прозрачность, поскольку веса каждого линейного слоя можно проанализировать, чтобы понять выходные данные модели.

6) Трансформер

Архитектура трансформера полностью построена на механизме самовнимания, исключая зависимость от порядка последовательности. Такая конструкция позволяет осуществлять параллельную обработку данных последовательности, что значительно повышает вычислительную эффективность. Трансформер состоит из двух основных компонентов: кодировщика и расшифровщика. Кодировщик преобразует входную последовательность в альтернативное представление, которое расшифровщик впоследствии использует для генерации целевой последовательности. Оба компонента состоят из нескольких идентичных слоев, каждый из которых включает в себя многоголовочный подслой самовнимания и подслой прямой нейронной сети. Многопоточный механизм внимания позволяет модели вычислять внимание в разных подпространствах одновременно, тем самым извлекая широкий спектр признаков из данных.

2 Анализ ситуации

2.1 Пакет данных

В данном исследовании использованы исторические данные по кластеру ветропарков в Китае, включающему пять ветропарков. Имеются временные исторические данные о скорости ветра, направлении ветра, мощности и суммарной мощности кластера для каждого ветропарка за

последние три месяца. Для спорных данных в исследовании используются традиционные методы очистки данных и метод квартилей для выявления выбросов, их восстановление осуществляется с помощью методов линейной регрессии, и в итоге получается набор данных каждые 15 минут в пяти точках.

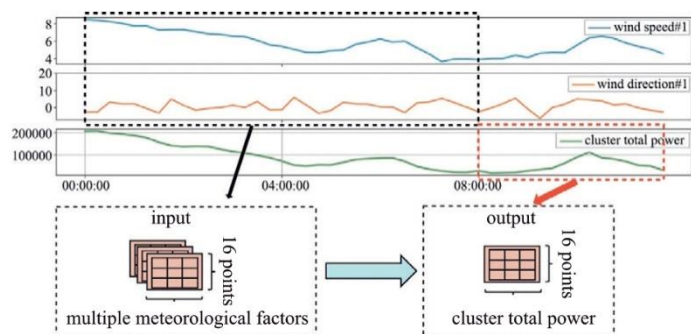


Рис. 9. Входные и выходные данные модели.

Набор данных был разделен на несколько флуктуационных процессов с интервалом в 4 часа (16 точек), а входные и выходные данные модели прогнозирования были представлены в единицах этих флуктуационных процессов, как показано на рис. 9. В качестве входных данных использовались метеорологические значения, содержащиеся в нескольких исторических флуктуационных процессах, а в качестве выходных данных — суммарное значение мощности кластера, содержащегося в будущем флуктуационном процессе.

В процессе обучения и тестирования модели данные численного прогнозирования погоды не использовались. Вводились только метеорологические данные n флуктуационных процессов до предполагаемого периода прогнозирования, полученные из данных с ветроэнергетических установок каждой ветроэлектростанции.

В качестве входных данных в данном исследовании проводился отбор на основе силы причинно-следственной связи между метеорологическими факторами различных ветропарков и общей мощностью кластера. Учитывая, что различные типичные процессы колебаний имеют разные входные и выходные данные, сначала был создан набор данных характеристик колебаний в соответствии с характеристиками колебаний общей мощности кластера для кластерного анализа с целью дальнейшего отбора метеорологических факторов с сильными причинно-следственными связями в качестве входных данных.

Для каждого процесса флуктуаций в данном исследовании были рассчитаны значения флуктуаций пяти метеорологических факторов для n процессов флуктуаций, и значения флуктуаций пяти метеорологических факторов были использованы в качестве меток столбцов. В итоге был получен набор данных характеристик погодных флуктуаций с n строками и пятью столбцами. В таблице 2 представлена часть набора данных характеристик погодных флуктуаций.

После сортировки набора данных характеристик флуктуаций все данные характеристик флуктуаций были нормализованы с использованием нормализации $\max\text{-min}$ для предотвращения ошибок, вызванных различными размерностями в кластерном анализе.

2.2 Характерный анализ стандартного процесса флуктуаций

В данном исследовании для создания кластерной модели использовался метод k -средних, а множественные процессы флуктуаций в кластере ветроэнергетики были кластеризованы и проанализированы на основе набора данных о характеристиках флуктуаций погоды. n процессов флуктуаций были разделены на несколько типичных процессов флуктуаций, и, как показано в

таблице 3, количество процессов флуктуаций, содержащихся в каждом типе процесса флуктуации, использовалось в качестве выходных данных. Одновременно использовалась частотная гистограмма для визуализации характеристик каждого типа процесса флуктуации, и были изучены характеристики флуктуаций мощности каждого типа типичного процесса флуктуации ветроэнергетики.

Таблица 2. Часть набора данных о характеристиках колебаний погоды.

Fluctuation Process	Wind Speed Variance	Wind Direction Variance	Wind Power Variance	Total Power Variation Trend	Mean Fluctuation Amplitude
1	0.229	13,503.5	136,852.1	119,599.7	−0.993
2	0.101	12,333.9	23,287.0	65,332.7	−0.781
...	...	...	...	...	...
2,127	0.120	19,266.6	1,907.9	19,015.4	0.760
2,128	0.253	63,076.2	7,434.7	31,300.6	−0.999

В таблице 3 указано количество флуктуационных процессов, включенных в типичные флуктуационные процессы.

Fluctuation Process Category	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6
The number of Fluctuation Processes	470	363	412	161	295	427

Результаты кластеризации были визуализированы, и для анализа характеристик мощности каждого процесса флуктуации в каждом классе была построена гистограмма распределения частоты собственных значений каждого процесса флуктуации. Были проанализированы характеристики колебаний скорости и направления ветра в шести типах процессов флуктуации.

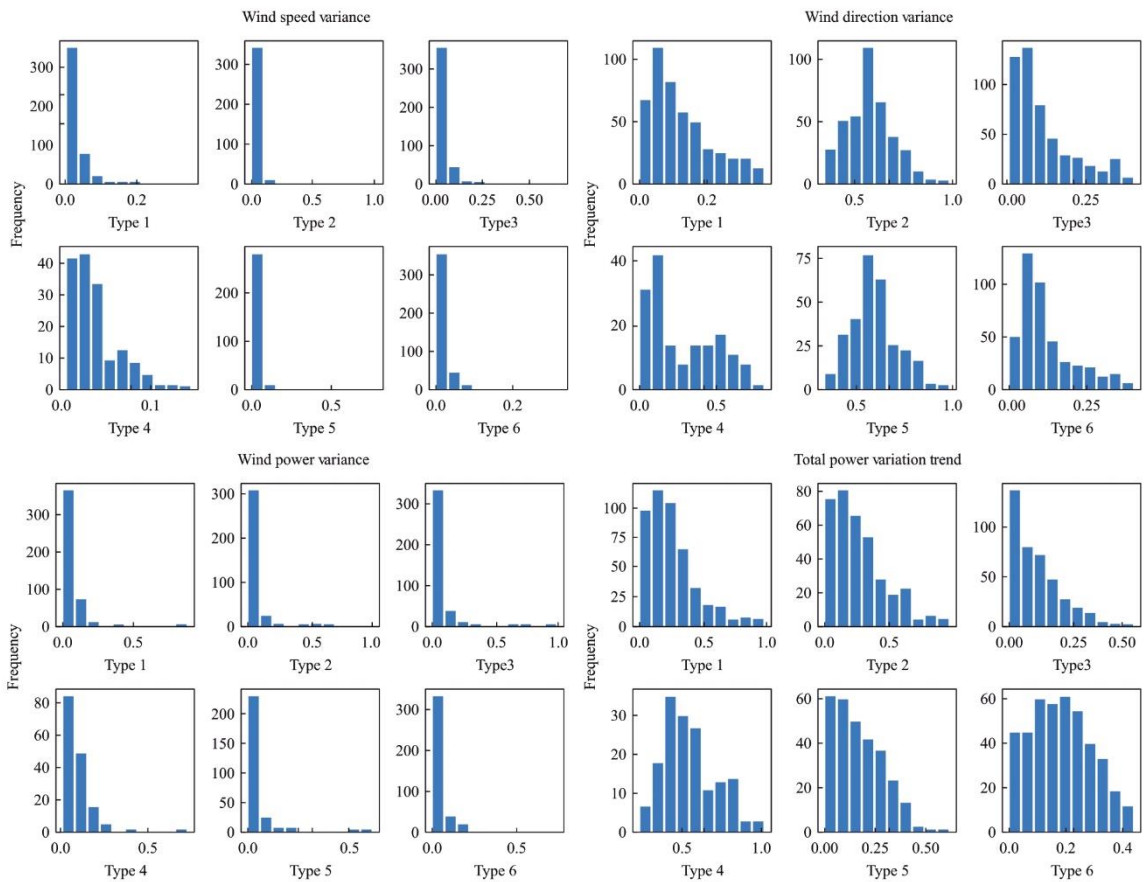


Рис. 10. Гистограмма частоты флуктуационных характеристик типичных флуктуационных процессов.

Как показано на рис. 10, между типичными процессами флуктуаций наблюдается минимальная разница с точки зрения дисперсии скорости ветра и дисперсии мощности. С точки зрения дисперсии

направления ветра, дисперсия направления ветра для процессов флуктуаций типов 1, 3, 4 и 6 в основном распределена в интервале (0, 0,2), что указывает на относительно небольшие колебания направления ветра для каждого типа процесса флуктуаций. Для процессов флуктуаций типов 2 и 5 дисперсия направления ветра в основном распределена в интервале (0,5, 1,0), а значение левого интервала максимальной частоты превышает 0,5, что указывает на относительно большие колебания направления ветра для каждого типа процесса флуктуаций. С точки зрения тренда изменения общей мощности, значение тренда изменения мощности для процесса флуктуаций типа 6 является наименьшим и в основном распределено в интервале (0, 0,25), тогда как значение тренда изменения мощности для типа 4 Процесс флуктуаций в основном распределен в интервале (0,5, 1,0), что представляет собой наибольший тренд изменений среди всех процессов флуктуаций.

Средние амплитуды флуктуаций каждого флуктуационного процесса демонстрировали значительные различия. Как показано на рис. 11, первый и второй типы флуктуационных процессов имеют распределение максимальной частоты (0, 0,05). Четвертый, пятый и шестой типы флуктуационных процессов имеют распределение (0,95, 1,00). Третий тип флуктуационного процесса имеет распределение (0,45, 0,50). С точки зрения формы распределения, первый, второй, четвертый и шестой типы флуктуационных процессов демонстрируют L-образное распределение, причем частота интервала с наибольшей частотой значительно выше, чем у других интервалов, и находится на самом краю всего интервала распределения. Однако после исключения интервала с наибольшей частотой частоты других интервалов существенно не различались. Третий тип флуктуационного процесса имеет приблизительно перевернутую V-образную форму распределения. Пятый тип флуктуационного процесса равномерно распределен по всему интервалу распределения, и частота Наибольший частотный интервал лишь немного превышал остальные.

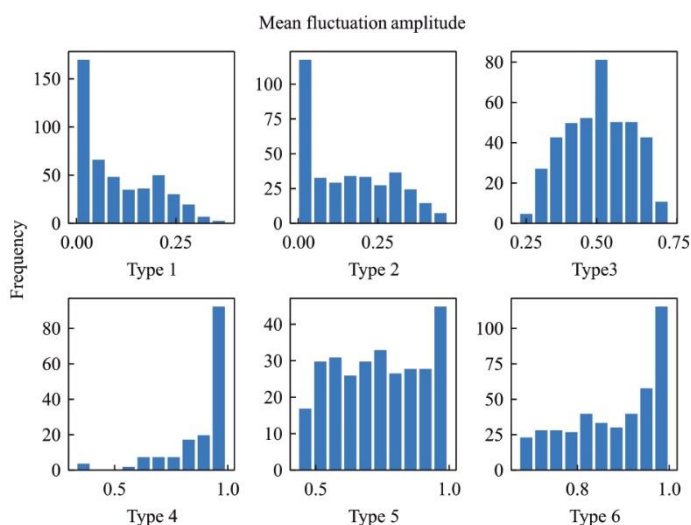


Рис. 11. Гистограмма средней амплитуды и частоты типичных флуктуационных процессов.

В целом, рассматривая в качестве примеров типичные флуктуационные процессы типа 1 и типа 4, можно отметить, что дисперсия скорости и направления ветра для обоих типов типичных флуктуационных процессов невелика, то есть степень флуктуации этих двух типов типичных флуктуационных процессов невелика во времени и пространстве; Однако степень концентрации дисперсий скорости и направления ветра в типичных процессах флуктуаций типа 1 значительно выше, чем в типичных процессах флуктуаций типа 4 в низком диапазоне. Поэтому разницу между двумя типичными процессами флуктуаций можно увидеть по колебаниям выходных данных: тенденция изменения общей мощности кластеров ветровой энергии в типичном процессе флуктуаций типа 1 в основном сосредоточена в периоде (0, 0,5), тогда как интервал (0,5, 1) в

основном является местом сосредоточения типичных процессов флуктуаций типа 4, и его средний диапазон колебаний более очевиден. Область между 0 и 0,25 в основном является местом возникновения типичных процессов флуктуаций типа 1. Интервал (0,5, 1) содержит наиболее типичные процессы флуктуаций типа 4. Этот метод классификации позволяет эффективно различать типичные процессы флуктуаций.

2.3 Анализ причинно-следственных связей на основе CCM

При анализе причинно-следственных связей сначала необходимо установить параметры. Как показано в таблице 4, основные параметры были установлены до начала анализа причинно-следственных связей на основе алгоритма CCM.

Для построения модели причинно-следственного анализа использовался алгоритм CCM. В итоге, для каждого типа колебаний скорость и направление ветра каждой ветровой электростанции принимались в качестве входной переменной X , а общая мощность кластера — в качестве входной переменной Y . Причинно-следственный анализ проводился с использованием модели причинно-следственного анализа. Алгоритм CCM позволяет количественно определить силу причинно-следственной связи. В процессе различных колебаний, чем сильнее причинно-следственная связь между метеорологическими данными и общей мощностью кластера, тем больше значение силы причинно-следственной связи; таким образом, можно соответствующим образом отбирать метеорологические факторы с сильной причинно-следственной связью. После анализа причинно-следственной связи на основе алгоритма CCM результаты были выведены и сопоставлены. На рис. 12 представлена линейная диаграмма интенсивности причинно-следственной связи между метеорологическими данными и общей мощностью кластера в каждом процессе колебаний.

Как показано на рис. 12, за исключением флуктуационного процесса типа 4, причинное влияние скорости ветра на общую мощность кластера, как правило, сильнее, чем влияние направления ветра, при других типах флуктуационных процессов. Кроме того, направление ветра на ветропарке № 1 демонстрирует значительную причинно-следственную связь с общей мощностью кластера при различных сценариях флуктуационных процессов.

Таблица 4. Настройка параметров модели анализа причинно-следственных связей.

Model Parameter Name	Parameter Setting
Variable $x(t)$	Wind speed and direction of each wind farm
Variable $y(t)$	Total cluster power
The length of the time series L	The total time series length of a typical fluctuation process

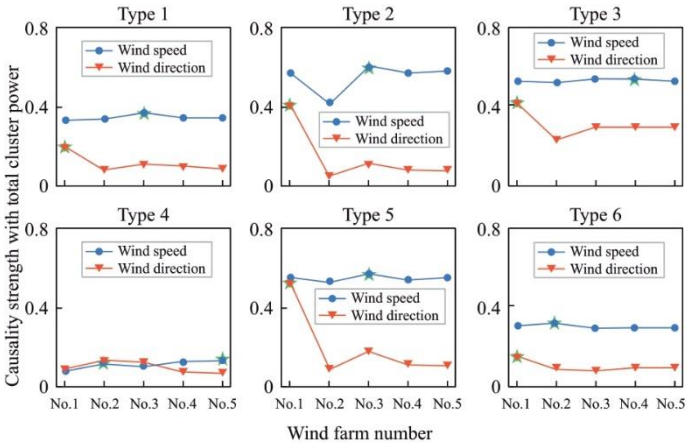


Рис. 12. Линейный график силы причинно-следственной связи.

2.4 Анализ результатов прогнозирования энергопотребления

В данном исследовании результаты прогнозирования мощности анализировались с двух точек зрения. Во-первых, была оценена эффективность причинно-следственного отбора предикторов. Результаты показали, что этот подход повышает точность прогнозирования мощности по сравнению с традиционным методом отбора на основе корреляции и методом без какого-либо отбора. Для подтверждения этого были созданы две контрольные группы: одна с использованием отбора на основе корреляции, а другая без отбора. Для исключения влияния конкретных моделей глубокого обучения для синхронной валидации было использовано шесть моделей глубокого обучения.

Во-вторых, была изучена эффективность усовершенствованного подхода к моделированию. В этом исследовании модели прогнозирования мощности были адаптированы для шести типичных процессов флуктуаций. Чтобы продемонстрировать превосходство усовершенствованного подхода к моделированию над традиционным прямым моделированием, были проведены контрольные эксперименты с идентичными наборами данных для обоих подходов. Результаты прогнозирования оценивались с использованием показателей RMSE и MAE [10,34,35].

2.4.1 Показатели оценки

В данном исследовании были выявлены и включены в модель глубокого обучения сильные причинные метеорологические факторы. Эти факторы были разделены на обучающую и тестовую выборки, специфичные для каждого типа флуктуационного процесса. В процессе обучения модели для сильных причинных метеорологических факторов для различных ветровых электростанций были назначены разные скорости обучения, чтобы оптимизировать модель прогнозирования и повысить точность прогнозирования. После получения результатов прогнозирования для количественной оценки ошибок прогнозирования использовались показатели RMSE и MAE. RMSE и MAE рассчитываются следующим образом:

$$RMSE = \sqrt{\frac{1}{L} \sum_{i=1}^L (P_{\text{ture},i} - P_{\text{pred},i})^2} \quad (21)$$

$$MAE = \frac{1}{L} \sum_{i=1}^L |P_{\text{ture},i} - P_{\text{pred},i}| \quad (22)$$

где $P_{\text{ture},i}$ — измеренная суммарная мощность временного ряда; $P_{\text{pred},i}$ — прогнозируемое значение суммарной мощности временного ряда; L — длина $P_{\text{ture},i}$; i — номер временной точки, $i=1,2,\dots,L$.

Для результатов прогнозирования суммарной мощности кластера каждые 15 минут в течение следующих 4 часов проводилась оценка погрешности, и вычислялось среднее значение шести групп значений RMSE или MAE, полученных для шести типов типичных флуктуационных процессов в каждой временной точке, чтобы получить комплексное значение RMSE или MAE для всего временного ряда; то есть, оценка погрешности результатов прогнозирования суммарной мощности кластера для нескольких типичных флуктуационных процессов проводилась за весь период времени.

2.4.2 Анализ достоверности причинно-следственного скрининга предикторов

Для оценки эффективности причинно-следственного отбора предикторов (сильная причинность) были созданы две контрольные группы для сравнения. В первой группе был применен традиционный метод корреляционного анализа, где в качестве входных данных использовались сильно коррелированные метеорологические факторы (сильная корреляция), при этом остальные параметры оставались неизменными. Во второй группе в качестве входных данных использовались

все метеорологические данные без какого-либо отбора (все факторы). На рис. 13 показаны значения RMSE ошибки прогнозирования при использовании метода CCM и в двух контрольных группах для шести различных моделей глубокого обучения.

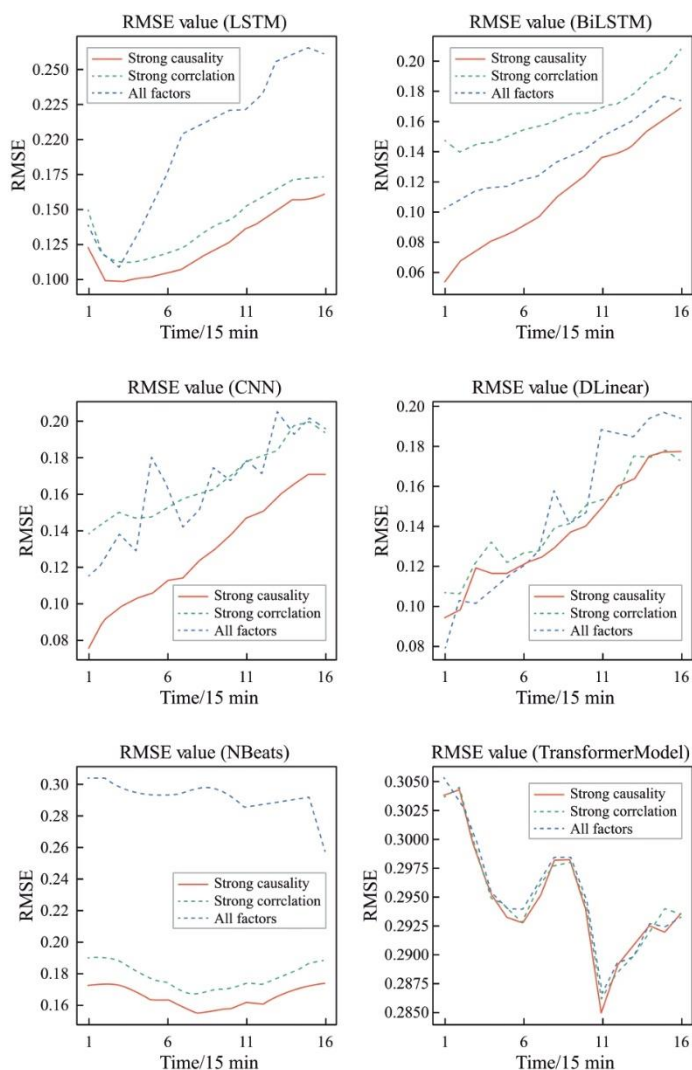


Рис. 13. Суммарное значение RMSE (включая две контрольные группы).

Как показано на рис. 13, ошибка прогнозирования RMSE модели сверхкраткосрочного кластерного прогнозирования ветровой энергии, основанной на методе CCM, была неизменно меньше, чем у двух контрольных групп во всех моделях глубокого обучения. На примере модели LSTM значение RMSE метода без учета факторов прогнозирования (все факторы) значительно увеличивалось со временем, при этом общие колебания превышали 0,15. В отличие от этого, ошибки прогнозирования для метода причинно-следственного анализа (сильная причинность) и метода корреляционного анализа (сильная корреляция) увеличивались более постепенно, при этом общие колебания оставались ниже 0,07. С точки зрения точности прогнозирования, метод причинно-следственного анализа (сильная причинность) превзошел контрольные методы. По сравнению с методом без учета факторов прогнозирования (все факторы) он достиг повышения точности на 7,32%. Кроме того, по сравнению с методом корреляционного анализа (сильная корреляция) он достигнуто улучшение на 1,57%. Эти результаты подтверждают, что предложенный метод причинно-следственного анализа на основе CCM значительно повышает точность сверхкраткосрочного прогнозирования кластеров ветровой энергии.

Точность прогнозирования варьировалась в зависимости от модели глубокого обучения, используемой в методе причинно-следственного анализа. В таблице 5 представлены средние

значения MAE, полученные с использованием уточненного подхода к моделированию с различными моделями глубокого обучения.

Из таблицы видно, что точность прогнозирования этого метода была выше при использовании моделей LSTM, Bi-LSTM и DLinear, а время обучения составляло приблизительно 40 секунд. В качестве примера рассмотрим ошибку прогнозирования, смоделированную с использованием модели LSTM. В таблице 6 показаны значения RMSE каждые 15 минут в течение следующих 4 часов, спрогнозированные методом причинно-следственного анализа и двумя контрольными группами.

На основе комплексных значений RMSE максимальное значение RMSE результата прогнозирования, полученного с помощью метода причинно-следственного анализа (сильная причинно-следственная связь), составило 0,160, а минимальное — 0,099, что указывает на хорошую эффективность прогнозирования.

2.4.3 Анализ эффективности методов моделирования

Чтобы продемонстрировать, что предложенный в данном исследовании усовершенствованный подход к моделированию для сверхкраткосрочного прогнозирования мощности — классифицированный на основе шести различных типичных процессов колебаний — повышает точность прогнозирования, была создана контрольная группа, использующая прямое моделирование (неклассифицированное). Результаты классифицированного прогнозирования мощности были сопоставлены с результатами контрольной группы. На рис. 14 представлено сравнение фактической суммарной мощности кластера и значений прогнозируемой мощности для двух подходов.

Реальная мощность представлена черной сплошной линией, прогностическая мощность уточненной модели для шести различных типичных флуктуационных процессов представлена красной сплошной линией, а прогнозируемая мощность прямого моделирования представлена синей пунктирной линией. Результаты прогнозирования уточненной модели (категоризированной), предложенной в данном исследовании, наиболее близки к фактической мощности результатов прогнозирования, полученных для шести типов флуктуационных процессов. В отличие от контрольной группы (некатегоризированной), на основе причинно-следственного анализа, прогностическая модель, выполняющая сверхкраткосрочное прогнозирование мощности, уточненная для шести различных типичных флуктуационных процессов (категоризированных), имеет лучший прогностический эффект.

Показатель MAE использовался для оценки ошибки в значениях прогнозируемой суммарной мощности кластера с 15-минутными интервалами в течение последующих 4 часов, как это было предсказано моделями сверхкраткосрочного прогнозирования мощности с использованием двух методов. В таблице 7 приведены сводные значения MAE для двух методов моделирования по 16 будущим временным точкам.

Данные в таблице 7 показывают, что усовершенствованный метод моделирования (категоризированный), адаптированный к процессу флуктуаций, неизменно давал меньшие ошибки прогнозирования общей мощности кластера в течение следующих 4 часов по сравнению с прямым методом моделирования (некатегоризированный). Кроме того, с течением времени темп увеличения значений MAE для категоризированного метода был значительно медленнее, чем наблюдаемый в контрольной группе (некатегоризированный).

Таблица 5. Среднее значение MAE уточненного метода моделирования при использовании различных моделей глубокого обучения.

Model	LSTM	Bi-LSTM	CNN	N-Beats	Dlinear	Transformer
Average MAE	0.089	0.088	0.099	0.129	0.087	0.259

Таблица 6. Суммарное значение RMSE в каждой временной точке.

Time Point	Strong Causality RMSE	Strong Correlation RMSE	All Factors RMSE	Time Point	Strong Causality RMSE	Strong Correlation RMSE	All Factors RMSE
1	0.122	0.149	0.139	9	0.121	0.138	0.216
2	0.099	0.116	0.118	10	0.127	0.143	0.220
3	0.097	0.112	0.108	11	0.136	0.152	0.221
4	0.100	0.112	0.127	12	0.141	0.157	0.230
5	0.101	0.115	0.153	13	0.149	0.164	0.255
6	0.105	0.118	0.175	14	0.156	0.171	0.260
7	0.107	0.122	0.204	15	0.157	0.172	0.265
8	0.114	0.130	0.21	16	0.160	0.173	0.260

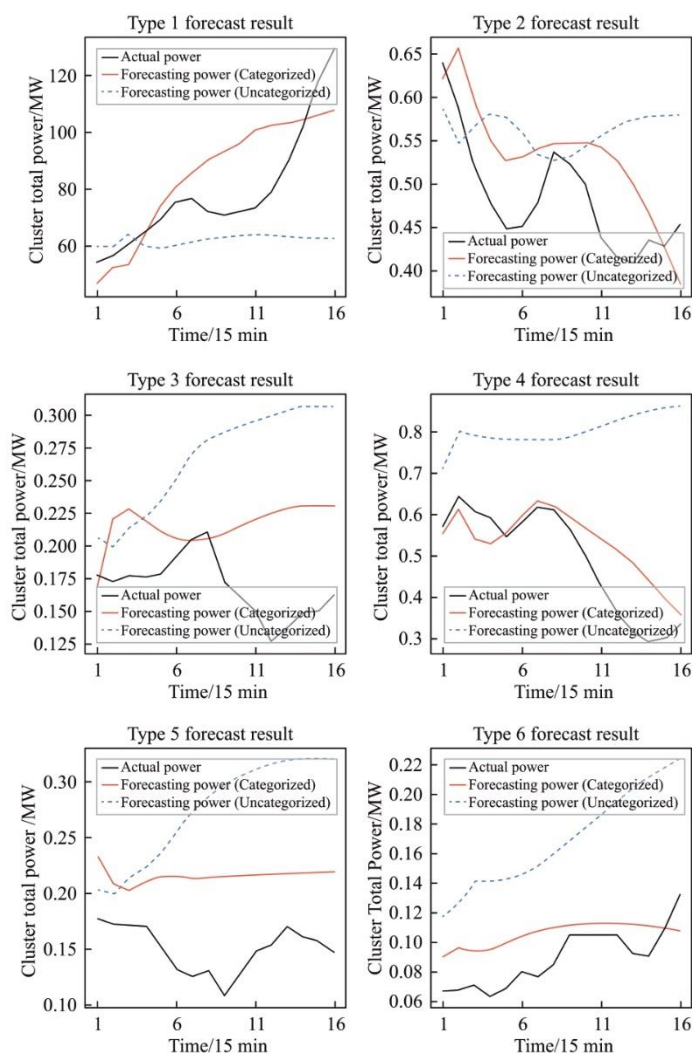


Рис. 14. Кривые результатов прогнозирования мощности для двух методов моделирования.

Таблица 7. Суммарное значение MAE для двух методов моделирования в каждой временной точке.

Time Point	Categorized MAE	Uncategorized MAE	Time Point	Categorized MAE	Uncategorized MAE
1	0.086	0.071	9	0.094	0.124
2	0.071	0.062	10	0.099	0.133
3	0.071	0.070	11	0.107	0.139
4	0.071	0.079	12	0.112	0.145
5	0.075	0.092	13	0.117	0.152
6	0.079	0.104	14	0.125	0.157
7	0.082	0.112	15	0.126	0.164
8	0.087	0.116	16	0.127	0.170

3 Заключение

В данном исследовании рассматриваются причинно-следственные связи между многомерными метеорологическими факторами и колебаниями суммарной мощности кластера ветровой энергии для разработки метода сверхкраткосрочного прогнозирования кластеров ветровой энергии на основе модели кластерной модели. Сделаны следующие выводы:

1) Используя подход анализа причинно-следственных связей алгоритма ССМ, этот метод выявляет метеорологические факторы прогноза с интенсивными причинно-следственными связями и количественно оценивает причинно-следственные связи между этими факторами и общей мощностью кластера. Кроме того, он исследует лежащие в основе физические взаимосвязи, значительно уменьшая влияние избыточных метеорологических факторов прогноза на прогнозирование общей мощности кластера. По сравнению с традиционными методами, этот подход улучшает направленность и интерпретируемость модели прогнозирования.

2) Учитывая вариации в причинно-следственных связях между многомерными метеорологическими факторами и суммарной мощностью кластера при различных типичных флуктуационных процессах, разработана кластерная модель для типичных флуктуационных процессов. Эта модель эффективно объединяет большое количество флуктуационных процессов в несколько различных типов, каждый из которых характеризуется уникальными закономерностями изменения выходных данных.

3) Используя различные сильные причинно-следственные метеорологические факторы прогнозирования, соответствующие различным типичным процессам колебаний, в качестве входных данных для модели прогнозирования мощности, достигается уточненное моделирование прогнозирования типичных процессов колебаний. Сравнительная проверка показывает, что этот метод обеспечивает измеримое улучшение точности прогнозирования по сравнению с традиционными подходами. В частности, точность прогнозирования улучшается в среднем на 1,57–7,32 % по сравнению с моделями сверхкраткосрочного прогнозирования мощности, основанными на корреляционном анализе.

В заключение, предложенный метод сверхкраткосрочного кластерного прогнозирования ветровой энергии на основе ССМ может эффективно повысить точность прогнозирования за счет учета причинно-следственных связей между метеорологическими факторами и колебаниями мощности с учетом различных типичных процессов колебаний. Это исследование не только повышает точность прогнозирования, но и предоставляет ценную информацию о лежащих в основе физических процессах, управляющих колебаниями мощности. Результаты анализа подчеркивают потенциал этого метода для более широкого применения в прогнозировании энергоснабжения и указывают на важность причинно-следственного анализа для повышения интерпретируемости и надежности модели.

О вкладе авторов в проект CRediT

Ючже Ян: Написание — первоначальный вариант, визуализация, методология, формальный анализ, концептуализация. Вэйе Сун: Написание — рецензирование и редактирование. Шуан Хань: Написание — рецензирование и редактирование. Цзе Янь: Написание — рецензирование и редактирование. Хань Ван: Написание — рецензирование и редактирование. Цяншэн Дай: Написание — рецензирование и редактирование. Сюэсун Хо: Написание — рецензирование и редактирование. Юнцян Лю: Написание — рецензирование и редактирование.

О конфликте интересов

Авторы заявляют, что у них нет известных конкурирующих финансовых интересов или личных связей, которые могли бы повлиять на результаты работы, представленной в данной статье.

Благодарности

Данное исследование было профинансировано в рамках научно-технического проекта Государственной электросетевой компании «Исследование ключевых технологий для прогнозирования и раннего предупреждения о масштабных скачках в режиме работы морских ветроэлектростанций на основе улучшения метеорологических данных» (4000-202318098A-1-1-ZN).

Список используемой литературы

- [1] Дж. Ян, Ч.Ю. Гао, Г.Б. Ю. Сравнительный анализ ошибок прогнозирования на основе характеристик морской ветроэнергетики, *Comput.Integr.Manuf.Syst.* 26 (3) (2020) 648-654.
- [2] Х. Демолли, А.С. Докуз, А. Эджемис и др., Прогнозирование ветровой энергии на основе ежедневных данных о скорости ветра с использованием алгоритмов машинного обучения, *Energ.Conver.Manage.* 198 (2019) 111823.
- [3] А.М. Фоли, П.Г. Лихи, А. Марвуглиа и др., Современные методы и достижения в прогнозировании выработки ветровой энергии, *Renew.Energy* 37 (1) (2012) 1-8.
- [4] Z.Y.Jiang, Q.S.Jia, X.H.Guan, Обзор методов прогнозирования ветровой энергии в многовременном и пространственном масштабе, *Acta Autom.Sin.* 45(1) (2019) 51-71. [
- 5] C.Ge, J.Yan, Y.Q.Liu и др., Обзор ключевых технологий управления и обслуживания морских ветропарков, *Proc.CSEE* 42(12) (2022) 4278-4291.
- [6] X.S.Peng, L.Xiong, J.Y.Wen и др., Обзор современного состояния дел в области краткосрочного и сверхкраткосрочного прогнозирования выработки ветровой энергии в регионах, *Proc.CSEE* 36 (23) (2016) 6315-6326, 6596.
- [7] Y.W.Cao, S.H.Yan, L.Haitao и др., Метод краткосрочного прогнозирования выработки ветровой энергии на основе сети глубокого обучения с временными рядами шумоподавления, *Proc.CSU-EPSC* 32 (01) (2019) 145-150.
- [8] J.X.Wang, B.Deng, J.Wang, Краткосрочное прогнозирование выработки ветровой энергии на основе эмпирического разложения на моды и нейронной сети RBF, *Proc.CSU-EPSC* 32 (11) (2020) 109-115.
- [9] Л. Ёе, М. Пей, П. Лу и др., Комбинированный метод прогнозирования краткосрочной выработки фотоэлектрической энергии на основе классификации погоды, *Autom.Electric Power Syst.* 45 (01) (2021) 44-54.
- [10] Ф. Шахид, А. Замир, М. Муниб, Новая генетическая модель LSTM для прогнозирования выработки ветровой энергии, *Energy* 223 (2021) 120069.
- [11] Y.Fu, Z.X.Ren, S.R.Wei и др., Сверхкраткосрочное прогнозирование мощности морских ветроэнергетических установок на основе улучшенной модели LSTMTCN, *Proc.CSEE* 42 (12) (2021) 4292-4302.
- [12] C.Liang, Y.Q.Liu, J.K.Zhou и др., Прогнозирование скорости ветра в нескольких точках на основе комбинации рекуррентных и сверточных нейронных сетей, *Power Syst.Technol.* 45 (2) (2021) 534-542.

- [13] G.Sideratos, N.D.Hatziargyriou, Усовершенствованный статистический метод прогнозирования ветровой энергии, *IEEE Trans.Power Syst.* 22 (1) (2007) 258-265.
- [14] Б. Ван, С. Л. Фэн, С. Лю, Исследование прогнозирования ветровой энергии на основе типов погоды, *Power Syst.Technol.*38 (01) (2014) 93-98.
- [15] З. Лин, С. Л. Лю, Прогнозирование ветровой энергии морской ветровой турбины на основе высокочастотных данных SCADA и нейронной сети глубокого обучения, *Energy* 201 (2020) 117693. 804
- [16] Г.З.Ю, Л.Лу, Б.Ян и др., Исследование метода прогнозирования сверхкраткосрочных субсекций морской ветроэнергетики с учетом переходных погодных условий, *Proc.CSEE* 42 (13) (2022) 4859-4871.
- [17] Ю.Цяо, З.Х.Лу, Ю.Мин, Исследование и применение методов повышения точности прогнозирования ветровой энергии, *Power Syst.Technol.* 41 (10) (2017) 3261-3269.
- [18] Дж.Дж.Чжу, Л.К.Су, Ю.В.Ли, Прогнозирование ветровой энергии на основе новой гибридной модели с модификацией остатков TCN, *Energy AI* 10 (2022) 100199.
- [19] С. Аль-Яхьяи, Й. Чараби, А. Гастли, Обзор использования моделей численного прогнозирования погоды (ЧПП) для оценки ветровой энергии, *Renew.Sustain.Energy Rev.* 14 (9) (2010) 3192-3198.
- [20] З. Х. Би, Дж. К. Ан, С. З. Чен, С. В. Хэ, Региональное прогнозирование ветровой энергии с учетом временных и пространственных характеристик, *J.Xi'an Jiaotong Univ. (Soc.Sci.)* 47 (10) (2013) 68-74.
- [21] С.М. Чжу, М. Ян, С.С. Хан и др., Совместное прогнозирование плотности вероятности краткосрочной выходной мощности нескольких ветропарков, *Автоматизированные электроэнергетические системы* 38 (19) (2014) 8-15.
- [22] Х. Ван, Исследование интегрированного метода прогнозирования ветровой энергии, солнечной энергии и системной нагрузки, *Электроэнергетический университет, Северный Китай*, 2021.
- [23] М. Ян, Л.Б. Чжан, Обзор сверхкраткосрочного прогнозирования ветровой энергии на основе подхода, основанного на данных, *Power Syst.Protect.Control* 47 (13) (2019) 171-186.
- [24] Y.N.Huang,K.Qu,C.Li и др., Исследование метода моделирования среднесрочных и долгосрочных временных рядов ветровой энергии на основе алгоритма К-средних MCMC, *Power Syst.Technol.*43 (7) (2019) 2469-2476.
- [25] M.Abdeyazdan, Кластерный анализ на основе гибридного алгоритма К-гармонических средних и модификатора империалистической конкуренции, *J.Supercomput.*68 (2) (2014) 574-598.
- [26] П. Ли, С. Х. Гуань, Дж. У и др., Анализ характеристик агрегированной генерации ветровой энергии на основе классификации погодных режимов, *Power Syst.Technol.* 39 (7) (2015) 1866-1872.
- [27] Ч. Х. Хунг, Х. М. Чиоу, В. Н. Ян, Поиск групп-кандидатов для кластеризации данных К-гармонических средних, *App.Math.Model.* 37 (24) (2013) 10123-10128.
- [28] С. Чжан, Л. Дун, Д. Ю. Цзи и др., Прогнозирование мощности сверхкраткосрочной фотоэлектрической станции на основе анализа сходства численного прогноза погоды, *Acta Energiae Solaris Sinica* 43 (4) (2022) 142-147.
- [29] Д. Мёнстер, Р. Фусароли, К. Тайлен и др., Причинно-следственный вывод из зашумленных данных временных рядов: тестирование алгоритма конвергентного перекрестного отображения в присутствии шума и внешнего влияния, *Futur.Gener.Comput.Syst.*73 (2017) 52-62.

[30] Дж. Дж. Лу, Т. Цай, Дж. Х. Ланг и др., Краткосрочное прогнозирование выходной мощности кластерных фотоэлектрических солнечных электростанций на основе кластерного разделения, *High Volt.Eng.*48 (5) (2022) 1943-1951.

[31] Дж.М.Маккракен, Р.С.Вайгель, Конвергентное перекрестное отображение и парный асимметричный вывод, *Phys.Rev.E* 90 (6) (2014) 062903.

[32] Г. Сугихара, Р. Мэй, Х. Йе и др., Обнаружение причинно-следственной связи в сложных экосистемах, *Science* 338 (6106) (2012) 496-500.

[33] W.Ма, Y.Qiao, L.R.Xie и др., Краткосрочное прогнозирование энергии ветра с двумя целями с учетом порога разделения колебаний метеорологических характеристик, *High Volt.Eng.*48 (10) (2022) 4154-4162.

[34] Н.В.Пенг, Ф.Р.Лиу, Х.Ф.Янг, Краткосрочное прогнозирование выработки ветровой энергии на основе искусственной нейронной сети, *Acta Energaie Solaris Sinica* 32 (8) (2011) 1245-1250.

[35] Q.Y.Wu, F.Guan, C.Lv и др., Сверхкраткосрочное многошаговое прогнозирование выработки ветровой энергии на основе CNN-LSTM, *IET Renew.Power Gener.*15 (5) (2021) 1019-1029.

[36] Х.Ванг, Метод прогнозирования оптимизации ветроэнергетики на основе нейронной сети. Шанхайский университет Цзяо Тонг, 2022 г.

[37] В.С.Ванг, З.Ванг, К.Донг и др., Статус и анализ ошибок технологии краткосрочного прогнозирования ветроэнергетики в Китае, *Autom.Electric Power Syst.*45 (01) (2021) 17-27.

[38] Мартин, Ш.Х.Э.Абдель Алим, А.Ф.Зобаа, О расчете среднеквадратической ошибки (RMSE) для оценки параметров фотоэлектрических моделей: новое точное аналитическое решение, основанное на W-функции Ламберта, *Energ.Conver.Manag.*210 (2020) 112716.

Поступила 17 мая 2024 г.;

отредактировано 11 сентября 2024 г.;

принято 4 ноября 2024 г.



Отсканируйте qr-код для получения более подробной информации

Экспертная оценка под ответственностью Global Energy Interconnection Group Co.Ltd. * Автор-корреспондент.

Адреса электронной почты:

nsepuyangyuzhe@163.com(Y.Yang),swy18813192984@163.com(W.Song),hanshuang@ncepu.edu.cn(S. Han),yanjie@ncepu.edu.cn (J.Yan), wangxiaohan0420@163.com (H.Wang), day_qs@163.com(Q.Dai),huoxuesong@js.sgcc.com.cn(X.Huo),yqliu@ncepu.edu.cn(Y.Liu).
<https://doi.org/10.1016/j.gloi.2024.11.014> 2096-5117/© 2025 Global Energy Interconnection Group Co.Ltd.

Издательские услуги Elsevier B.V. от имени KeAi Communications Co.Ltd. Это статья имеется в открытом доступе по лицензии CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



Ючжэ Ян получила степень бакалавра наук в области возобновляемых и экологически чистых источников энергии в Школе новой энергетики Северо-Китайского университета электроэнергетики, Пекин, Китай, в 2023 году. В настоящее время она осваивает докторскую степень в той же школе. Направление ее исследований — оценка ветровых ресурсов.



Вэйе Сун получил степень бакалавра наук в области возобновляемых и чистых источников энергии в Школе новой энергетики Северо-Китайского университета электроэнергетики, Пекин, Китай, в 2020 году. В настоящее время он осваивает докторскую степень в той же школе. Направление его исследований — прогнозирование мощности ветровой и солнечной энергии.



Шуан Хань получила степень доктора философии в Школе энергетической энергетики и машиностроения Северо-Китайского университета электроэнергетики (NCEPU), Пекин, Китай, в 2008 году. В настоящее время она является профессором и руководителем докторантуры в Школе новой энергетики NCEPU. Она уже давно занимается преподаванием и исследованиями в области технологий ветряных электростанций, включая прогнозирование мощности ветряных электростанций, анализ характеристик ветровых ресурсов и оценку эффективности ветряных электростанций.



Цзе Ян (старший член IEEE). В 2016 году получила совместную докторскую степень в области возобновляемых и экологически чистых источников энергии в NCEPU, Пекин, Китай, и Университете Бата, Бат, Великобритания. В настоящее время она является профессором Школы новой энергетики в NCEPU. Ее основные исследовательские интересы включают анализ погрешностей в производстве ветровой энергии, прогнозирование ветровой/солнечной энергии.



Хань Ван получила степень доктора философии в области возобновляемых и чистых источников энергии в Школе новой энергетики NCEPU, Пекин, Китай, в 2021 году. Сейчас она преподает в Школе новой энергетики Северо-Китайского университета электроэнергетики в Пекине, Китай. Ее основной исследовательский интерес включает прогнозирование энергии ветра и солнца.



Цяншэн Дай получил степень доктора философии в области электротехники на факультете электротехники Университета Цинхуа, Пекин, Китай, в 2020 году. Он работает в компании State Grid Jiangsu Electric Power Company Ltd., Нанкин, Цзянсу, Китай. Его исследовательские интересы включают интензивное обучение, прогнозирование мощности и интеллектуальные сети.



Сюэсун Хо получил степень магистра в области компьютерных технологий в Университете Хохай в 2002 году. Он работает в компании State Grid Jiangsu Electric Power Company Ltd., Нанкин, Цзянсу, Китай. Его исследовательские интересы включают интенсивное обучение, прогнозирование мощности и интеллектуальные сети.



Юнцзянь Лю получил докторскую степень по автоматизации производства в Университете Нанси-1 и гидроэнергетике в Хуачжунском университете науки и технологий в 2002 году соответственно. Он имеет 30-летний профессиональный опыт работы в области ветроэнергетики и гидроэнергетики. Сейчас он является профессором школы новой энергетики Северо-Китайского университета электроэнергетики, Пекин, Китай. Его основные исследовательские интересы сосредоточены на технологиях ветряных электростанций, включая оценку ветровых ресурсов и проектирование ветряных электростанций, моделирование следа, прогнозирование ветровой энергии, эксплуатацию и техническое обслуживание ветропарков.

(Редактор Цзэдун Чжан)