

Построение гибридных моделей временных рядов на основе вейвлет-преобразования

Д. В. Бакулин*, Н. В. Белецкая*,**, Н. Н. Горшков*, М. И. Джигоева*,
У. С. Мохнаткина*, Д. А. Петрусевиич*, А. А. Поляков*

*Российский Технологический Университет (МИРЭА), Москва, Россия

**Институт проблем передачи информации им. А.А. Харкевича, Москва, Россия

Поступила в редколлегию 04.08.2025 г. Принята 10.10.2025 г.

Аннотация—В представленной работе произведено сравнение применения вейвлетов в виде преобразований EWT и VMD, а также трансформера, принимающего на вход результаты этих преобразований, с работой стандартных моделей при прогнозировании временных рядов. Основная идея преобразований сводится к применению преобразования Фурье и извлечению информации о поведении ряда из полученного результата. При этом, применять преобразования можно по-разному: целиком к ряду, по отдельности к тренду и сезонности и т.п. Более того, компоненты, полученные после разложения, можно как собрать обратно, так и использовать в других моделях для определения взаимозависимостей в данных ряда. Для описания тренда и сезонности существует большое количество подходов, в том числе преобразование STL, работа статистических моделей ARIMA/ETS, обработка тренда с помощью полиномов, разложение сезонности в ряд Фурье и т.п. Эти подходы можно совмещать, используя разные модели для описания разных частей ряда (тренда, сезонности и остатка) в связи с тем, что обычно каждая модель лучшим образом описывает какую-то одну часть. Например, статистические модели ARIMA/ETS прекрасно описывают сезонность и остаток, обладая средствами для описания тренда. Нейросетевые модели лучше описывают тренд, имея проблемы с сезонностью. Предполагается, что совмещение моделей улучшит их прогнозы. В представленном исследовании компоненты EWT используются для описания тренда и сезонности. Их результаты вместе с исходным сигналом подаются на вход для нейросетевых моделей временных рядов (трансформеров), таким образом формируя гибридные модели. Прогнозы таких моделей сравниваются с прогнозами стандартных моделей.

КЛЮЧЕВЫЕ СЛОВА: EWT, VMD, преобразование Фурье, вейвлет-анализ, трансформер, временные ряды.

DOI: 10.53921/18195822_2025_25_3_249

1. ВВЕДЕНИЕ

В работе произведено построение гибридных моделей временных рядов, совмещающих при обучении информацию о тренде и сезонности в виде результата вейвлет-преобразования и данных в исходном виде. Одна из важных проблем нейросетевых моделей временных рядов – описание сезонности. Если тренд относительно неплохо описывается ими, то с описанием периодических колебаний почти у всех моделей возникают трудности [1]. В целом, для описания сезонности больше подходят либо модели стандартные ARIMA/ETS, либо модели, учитывающие сезонность в форме результата преобразования Фурье (например, EWT/VMD), которые и используются в работе при построении гибридных моделей.

Подробное теоретическое описание и исследование преобразования EWT произведено в работе [2]. Применение подобных преобразований было успешно испытано на сигналах, соответствующих графическим изображениям, и затем были этот опыт был перенесён на обработку

сигналов другой природы – временных процессов. Причём, в обработке изображений преобразование EWT применяется как для повышения чёткости [3, 4], так и для распознавания изображений [5, 6]. Основная особенность преобразования EWT по сравнению с быстрым преобразованием Фурье или, в целом, вейвлет-анализом состоит в формировании вейвлет-функций на основе обрабатываемого сигнала, а не применении построенного заранее набора функций. Процесс обработки сигналов с помощью EWT можно свести к следующим этапам [2].

- Строится амплитудный спектр сигнала для выявления ключевых особенностей в частотной области.
- Определение границ частотных диапазонов: спектр делится на несколько сегментов, границы которых выбираются на основе характерных особенностей спектра (например, резкие перепады или пиковые значения).
- Построение адаптивных фильтров: для каждого сегмента спектра создаются полосовые фильтры, определяющие соответствующие вейвлет-функции.
- Разложение сигнала: исходный сигнал разбивается на набор компонент (мод), каждая из которых соответствует определённому частотному диапазону.

Интересно, что EWT-преобразование способно как самостоятельно полностью обработать временной сигнал, так и обработать только часть временного процесса, например, сезонность, выделив компоненты, обладающие различной периодичностью. Это открывает широкое поле для внедрения преобразования в различные этапы обработки данных временного ряда. Очень часто преобразование используется для обработки сложной сезонности вместе с другими моделями, обрабатывающими остаток и тренд [7–10].

Отдельно разберём третий пункт сценария обработки сигнала. Для каждого частотного диапазона $[\omega_{k-1}, \omega_k]$ создаётся фильтр $\psi_k(f)$ с использованием пороговых функций. Низкочастотные фильтр ($k = 1$):

$$\psi_k(f) = \begin{cases} 1, |f| < \omega_1 - \gamma, \\ \cos\left(\frac{\pi}{2} \frac{|f| - (\omega_1 - \gamma)}{\gamma}\right), \omega_1 - \gamma < |f| \leq \omega_1, \\ 0, \text{иначе} \end{cases}$$

Здесь γ – параметр ширины переходной области фильтра [2]. Диапазонные фильтры ($k > 1$):

$$\psi_k(f) = \begin{cases} \sin\left(\frac{\pi}{2} \frac{|f| - (\omega_{k-1} - \gamma)}{\gamma}\right), \omega_{k-1} < |f| \leq \omega_{k-1} + \gamma, \\ 1, \omega_{k-1} + \gamma < |f| \leq \omega_k - \gamma, \\ \cos\left(\frac{\pi}{2} \frac{|f| - (\omega_k - \gamma)}{\gamma}\right), \omega_k - \gamma < |f| \leq \omega_k, \\ 0, \text{иначе} \end{cases}$$

Высокочастотный фильтр ($k = N$):

$$\psi_k(f) = \begin{cases} \sin\left(\frac{\pi}{2} \frac{|f| - (\omega_{N-1} - \gamma)}{\gamma}\right), \omega_{N-1} < |f| \leq \omega_{N-1} + \gamma, \\ 1, \omega_{N-1} + \gamma < |f|, \\ 0, \text{иначе} \end{cases}$$

В приложениях дан пример разложения ряда количества авиапассажиров [11] с помощью преобразования EWT: рисунки 1, 2, 3. Хотя согласно теоретическому подходу к построению

моделей временных рядов, модель, представляющая тренд и сезонность, должна совершенствоваться вплоть до того момента, когда остаток будет равномерно распределён вокруг нуля, на практике можно провести обратное преобразование по конечному числу компонент и проверить точность прогнозирования и близость восстановленного сигнала к исходному.

К преимуществам метода можно отнести:

- Адаптивность: EWT автоматически подстраивается под спектральные характеристики сигнала, что делает его особенно полезным для анализа нестационарных сигналов.
- Изоляция частотных компонент: метод эффективно выделяет частоты без наложений, улучшая разложение сложных сигналов.
- Гибкость: EWT применим к сигналам с произвольной структурой и спектральным распределением.
- Высокая точность: благодаря адаптивным фильтрам EWT обеспечивает более точное разложение, чем стандартные методы.

Вариационное разложение на моды (Variational Mode Decomposition, VMD) [12] — современный метод адаптивной декомпозиции сигнала на набор компонент (мод), каждая из которых рассматривается как узкополосный сигнал, локализованный в частотной области вокруг некоторой центральной частоты. В основе преобразования находится определённая вариационная задача. Каждая мода – это сигнал, который можно интерпретировать как АМ-ФМ компоненту, сосредоточенную вокруг собственной центральной частоты в спектральной области. Моды преобразуются в аналитический сигнал, умножение на $\exp(-i\omega_k t)$ (где ω_k – частота обрабатываемой компоненты) сдвигает спектр моды к низким частотам. К лагранжиану задачи зачастую добавляются множители для обеспечения точного или приблизительного выполнения условия реконструкции, а также для регуляризации:

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_{k=1}^K \left\| \partial_t \left(\left(\delta + \frac{i}{\pi t} \right) u_k e^{-i\omega_k t} \right) \right\|^2 + \|f - \sum_k u_k\|^2 + \langle \lambda, f - \sum_k u_k \rangle,$$

здесь $\lambda(t)$ – множитель Лагранжа, α задаёт уровень сжатия полосы (регуляризация), u_k – обрабатываемые моды сигнала $f(t) = \sum_{k=1}^K u_k(t)$, $\|\cdot\|$ – норма, $(\delta + \frac{i}{\pi t} u_k e^{-j\omega_k t})$ – результат преобразования Гильберта, применённого к моде u_k [12, 13]. Метод чередования направлений множителей (ADMM) позволяет решить эту задачу, последовательно оптимизируя лагранжиан на каждом шаге по моде u_k , затем по частоте ω_k . Для мод u_k решение в частотной области сводится к винеровской фильтрации (Wiener filtering), а для частот ω_k оказывается, что оптимальная центральная частота является центром масс спектра соответствующей моды [12, 13]:

$$\omega_k = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega}.$$

К основным особенностям преобразования можно отнести следующее.

- Теоретическая обоснованность: В отличие от предшественника этого алгоритма метода EMD [14], VMD имеет чётко сформулированный функционал, что упрощает теоретический анализ сходимости и устойчивости.
- Устойчивость к шуму. VMD может интерпретироваться как обобщение винеровской фильтрации, что обеспечивает более стабильные решения в присутствии аддитивных шумов по сравнению с EMD.

- Адаптивность. VMD не требует предопределённой базы. Моды определяются из данных, адаптивно выделяя доминирующие квазигармонические компоненты с различными центральными частотами.
- Численная эффективность. Оптимизационная схема в частотной области обычно сходится быстрее и более надёжно, чем эвристические процедуры EMD.

Так же, как и в случае преобразования EWT, здесь можно моделировать целиком временной процесс или его отдельные части, формируя гибридные модели [15–21]. На практике чаще всего первая мода представляет собой тренд или тренд с некоторой сезонной частью, а следующие компоненты уменьшаются по амплитуде и улавливают основную часть сезонных и циклических колебаний. В приложениях дан пример разложения ряда D3597 из набора данных [22] с помощью преобразования VMD: рисунки 4, 5, 6.

EWT и VMD – два мощных адаптивных способа разложения нестационарных сигналов. Оба метода выделяют узкополосные компоненты, однако делают это по-разному:

- EWT строит набор полосовых вейвлет-фильтров, адаптируя границы полос к локальным экстремумам амплитудного спектра.
- VMD формулирует задачу как вариационную оптимизацию: одновременно ищет моды и их центральные частоты, минимизируя суммарную ширину полос

Сочетание этих разложений с моделями на основе самовнимания (трансформерами) полезно там, где требуется и частотный анализ, и выявление долгосрочных зависимостей в последовательностях. Причём, строить трансформеры, принимающие сигналы различной природы, можно несколькими способами. Приведём несколько примеров (сохраняется не очень удачное название терминов при описании работы трансформера на основе запросов Q , ключей K и значений V [23, 24]).

1. Многоканальный трансформер. После разложения все компоненты складывают в трёхмерный тензор. На каждый временной шаг поступает вектор, где элементы-каналы – это значения отдельных мод. К входным векторам прибавляют позиционные кодировки и обучаемые векторные представления мод, чтобы модель различала, какая мода перед ней. Энкодер из attention-блоков сразу учит как временные корреляции, так и взаимодействия между каналами. Итоговые предсказания каналов просто суммируются, давая прогноз исходного сигнала. Особенность этого подхода – одинаковая архитектура для всех масштабов при относительно небольшом количестве параметров.
2. Разделение и сборка. У каждой моды свой небольшой энкодер (2-3 слоя). Каждый энкодер работает индивидуально, настраиваясь исключительно на информацию своего канала. Скрытые представления всех энкодеров затем объединяют и пропускают через обобщающий настраиваемый модуль: например, многослойный персептрон. В отличие от первого способа память и параметры масштабируются.
3. Слияние через перекрёстное внимание. Конструируют два параллельных энкодера: один получает исходный сигнал, другой – набор мод, полученных в результате преобразований EWT/VMD. В декодере организуется кросс-внимание: запросы Q берутся из скрытых состояний потока, обрабатывающего исходный сигнал, – а ключи-значения K, V – из представлений мод. Так каждое предсказание использует только нужную информацию из спектральных компонент. Настройка перекрёстного внимания сложнее, вычислений больше, чем в предыдущих случаях, т.к. присутствует два направления для обработки сигналов.
4. Композиция трансформеров. Каждый трансформер видит только свою моду и прогнозирует её будущее; итоговое предсказание – сумма предсказаний на всех моделях. С одной

стороны, возможна гибкая настройка каждого трансформера. Но плата за это – существенное увеличение объёма вычислений. Кроме этого, при необходимости учёта взаимодействия между модами нужны дополнительные слои сети.

В рамках представленной статьи используется последний подход. В дальнейшем планируется провести эксперименты по всем четырём способам построения системы.

Сейчас подход, при котором результат частотного преобразования тренда или тренда и сезонности передаётся на вход нейросетевой модели: стандартной (LSTM, GRU и т.п.) [1] или трансформера [25, 26], – является очень популярным. Большое количество работ посвящено обработке определённых временных процессов. Основным вопросом остаётся оптимальный способ применения нейронных сетей. На вход нейросетевому модулю можно передать исходные данные временного ряда, результат обработки с помощью вейвлет-преобразования или их можно комбинировать. Вычислительный эксперимент показывает, что зачастую выигрывает подход, при котором удаётся выделить существенные характеристики для описания тренда и сезонности (чаще всего, на основе вейвлет-преобразований) и передать их в обработанном виде на вход трансформера (или другой сети, выполняющей функции выявления связей во временном процессе). В работах [25, 26] параметры преобразований подбираются специально для описания единственного временного процесса, и такой подход выглядит очень перспективным. В представленной работе гибридные модели используются для обработки целого набора данных, поэтому выводы, которые можно сделать по результатам исследования, обладают обобщающей способностью. Кроме этого, характеристики ряда, которые извлекаются из его данных и подаются на вход нейронной сети, можно получить и другим способом (не только на основе вейвлет-преобразования), например, на основе пакета FFORMA [27]. В целом, принцип обработки не меняется. В перспективе предполагается использовать этот подход и сравнить результаты с применением вейвлет-преобразования.

2. ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ

Модели, реализующие различные подходы к моделированию тренда и сезонности, были проверены на наборе данных M4 [22]. Таблица 1 содержит точность (RMSE) лучших моделей в сравнении со стандартными моделями ARIMA и нейросетевыми моделями LSTM [28, 29]. В исследовании были построены гибридные модели, где сезонность описывается с помощью вейвлет-преобразований EWT/VMD, а нейронная сеть получает данные в исходном виде (имеет информацию о тренде) и результат обработки сезонности. Схожий подход использован на основе EWT, например, в [25], а в работе [26] использован VMD. Второй тип гибридных моделей, участвующих в эксперименте, состоит из описания тренда с помощью вейвлет-преобразований, а сезонности – на основе нейросетевых моделей (LSTM). Результаты прогнозирования, проведённого с помощью гибридных моделей, сравниваются с результатами нейросетевой модели LSTM и стандартной статистической модели ARIMA. Преобразование EWT выполнено на основе ортогональных вейвлетов Мейера. К особенностям набора данных следует отнести следующее.

1. Разнообразие временных масштабов — от колебаний на часовом масштабе до многолетних трендов. Это заставляет модель быть одновременно чувствительной к короткой памяти и устойчивой на длинном горизонте.
2. Сильный дисбаланс классов: более 7/10 рядов – длиной в месяц.
3. Чисто унивариативный корпус: никаких экзогенных переменных не дано, присутствует только один столбец значений.

По результатам построения моделей можно сделать следующие выводы по обработке рядов с различной частотностью.

Таблица 1. Точность прогнозов моделей временных рядов (функция RMSE) на примере набора данных M4

Модель, частота	Часовая	Дневная	Недельная	Годовая
Трансформер + VMD	771	76	307	287
LSTM + VMD	1076	128	668	480
Трансформер + EWT	1126	106	364	274
LSTM + EWT	1460	122	399	387
LSTM	441	243	1417	548
ARIMA	326	68	715	467

- Для рядов с часовой частотностью лучшие прогнозы получены с помощью нейронной сети LSTM. У этих рядов ярко выражен тренд. С описанием таких рядов прекрасно справляются нейросетевые модели. Можно ожидать, что здесь могли бы дать хорошие результаты и другие нейронные сети, не применявшиеся в эксперименте (например, GRU). LSTM справляется лучше тяжёлых композиций моделей, вейвлет-разложение дробит сигнал: получается много моделей, каждая обучается на меньшем количестве точек, качество падает из-за переобобщения.
- Для рядов с дневной частотностью трансформер + VMD и ARIMA дали приблизительно одинаковые результаты по качеству прогноза. Тренд почти линейен и имеет стабильную сезонность, поэтому стандартная модель ARIMA прекрасно подходит для описания таких данных.
- У рядов с недельной частотностью ярко выражена сезонность, гибридные модели трансформер + EWT/VMD дают наилучшие прогнозы.
- У рядов с месячной частотностью «длинная» мода колебаний описывается с помощью вейвлет-преобразований самым эффективным образом по сравнению с другими моделями, участвующими в сравнении. Гибридные модели трансформер + EWT/VMD дают наилучшие прогнозы. Самовнимание прекрасно находит длинные зависимости в данных.

3. ЗАКЛЮЧЕНИЕ

В вычислительном эксперименте сравнивались прогнозы стандартной модели ARIMA и одной из простых нейросетевых моделей LSTM с гибридными моделями, где результат обработки тренда и сезонности с помощью вейвлет-преобразований EWT/VMD подавался на вход нейронной сети трансформера. Композиции, комбинирующие спектральное разложение и нейронные сети, продемонстрировали ощутимый выигрыш: лучшая конфигурация VMD + трансформер снизила ошибку прогноза RMSE на четверть по сравнению с ARIMA на недельных и месячных данных и удержала R^2 0.9 на данных с дневной частотой. Отмечено, что VMD-разложение формирует чище изолированные полосы, поэтому трансформер обучается стабильнее, чем при использовании EWT. В свою очередь композиции LSTM + VMD/EWT улучшили простую сеть LSTM, но без механизма внимания оказались не чувствительны к длинным зависимостям и уступили трансформерам на длинных рядах и рядах со сложной сезонностью.

Эксперимент показал, что на часовых рядах с монотонным трендом простые модели (ARIMA или одиночная LSTM) дают сопоставимое качество при значительно меньших затратах, тогда как при ярко выраженной сезонности (недельная) и долгих трендах и циклах (месячные) преимущество остаётся за спектрально-трансформерными композициями.

Дальнейшая работа заключается в применении усовершенствованных методов вейвлет-анализа, например: ансамблевого эмпирического модового разложения [30] и CEEMDAN [31,32] – в гибридных моделях. Кроме этого, предполагается добавление в модели описания остатка на

основе беггинга [33] и построение гибридных моделей, где вейвлет-преобразование используется для описания тренда и сезонности, либо только тренда с описанием сезонности с помощью стандартных моделей ARIMA/ETS, а остаток обрабатывается на основе беггинга.

4. ПРИЛОЖЕНИЯ

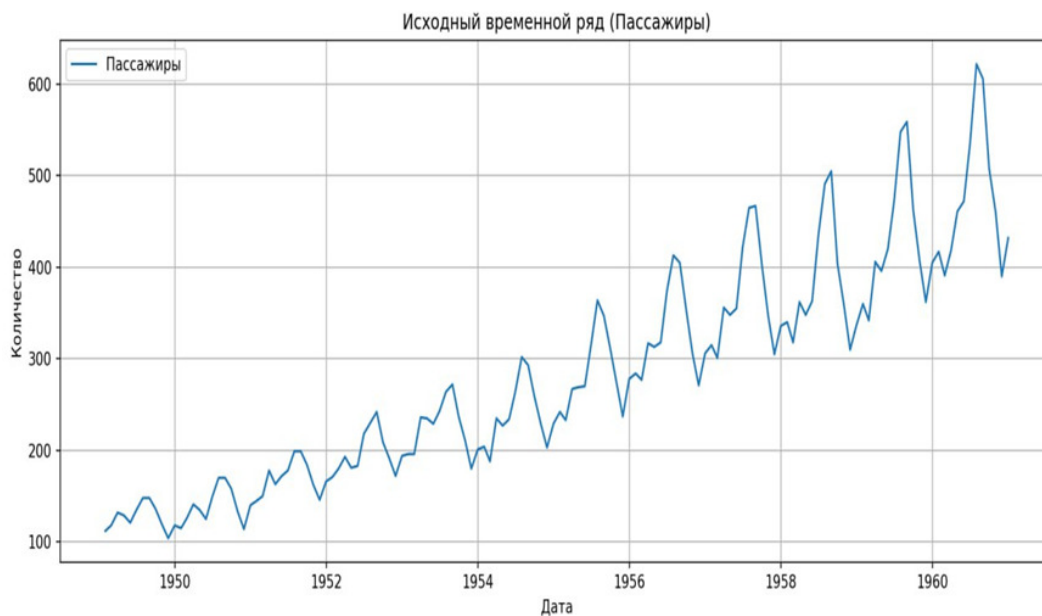


Рис. 1. График ряда пассажиропотока

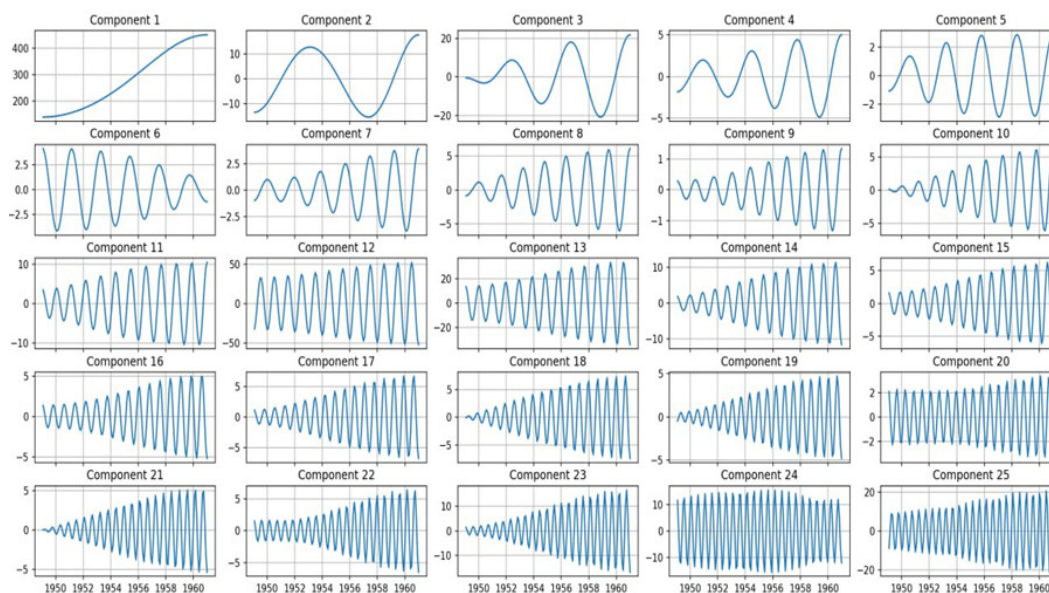


Рис. 2. Первые 25 компонент результата обработки ряда пассажиропотока с помощью преобразования EWT

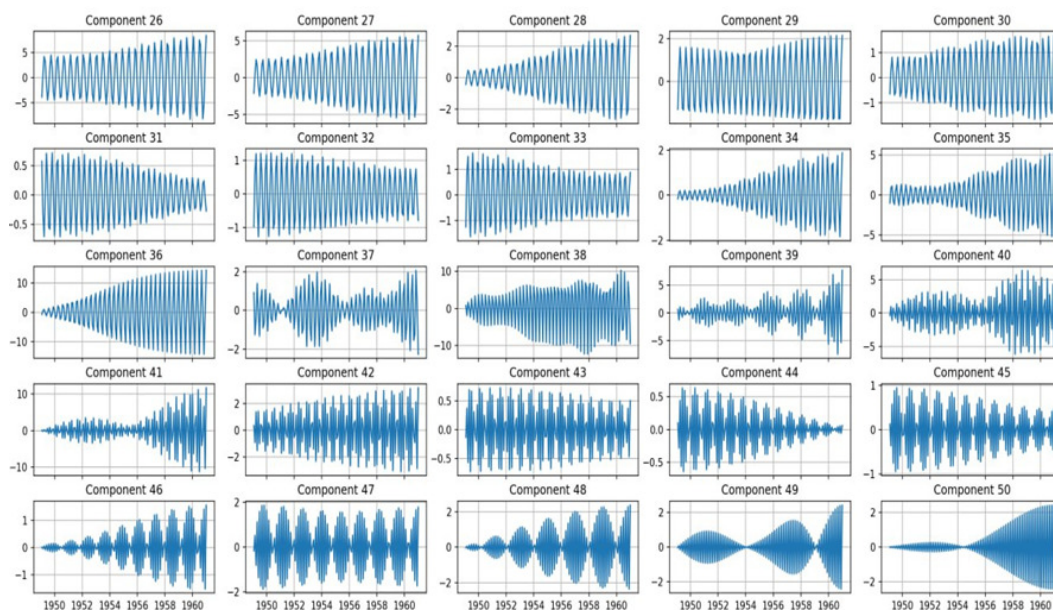


Рис. 3. Вторые 25 компонент результата обработки ряда пассажиропотока с помощью преобразования EWT

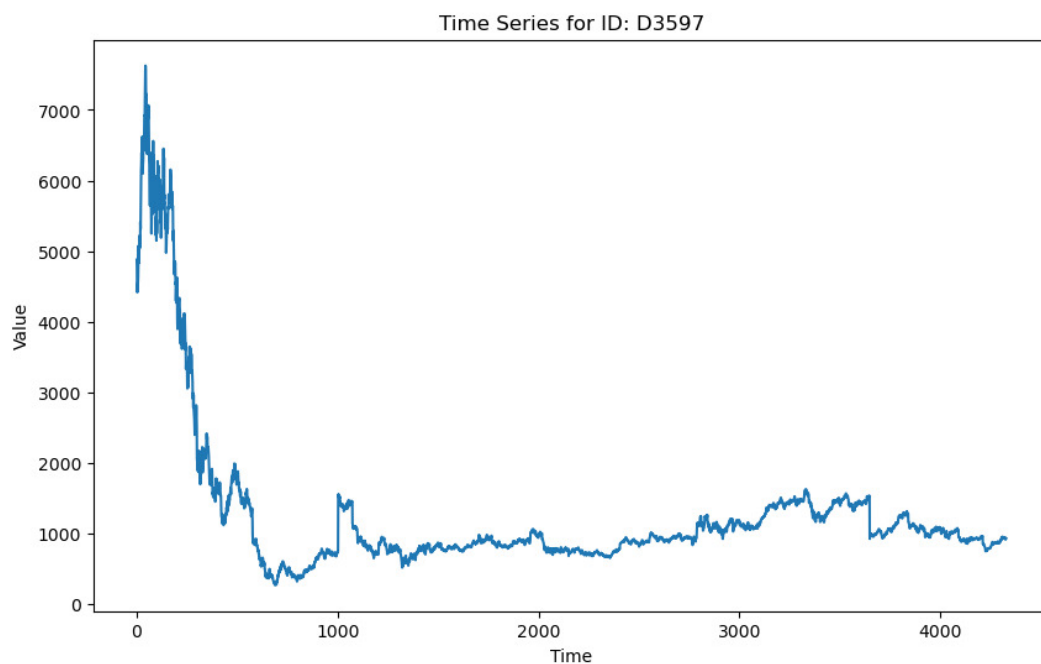


Рис. 4. График ряда D3597 из набора данных M4

СПИСОК ЛИТЕРАТУРЫ

1. Пашпоев Б., Петрусевиц Д.А., Анализ нейросетевых моделей для прогнозирования временных рядов. Российский Технологический Журнал, 2024, 12, №4, с. 106–116. doi: 10.32362/2500-316X-2024-12-4-106-116.
2. Gilles J., Empirical Wavelet Transform. IEEE Transactions on Signal Processing, 2013, vol. 61, pp.3999–4010. doi: 10.1109/TSP.2013.2265222.

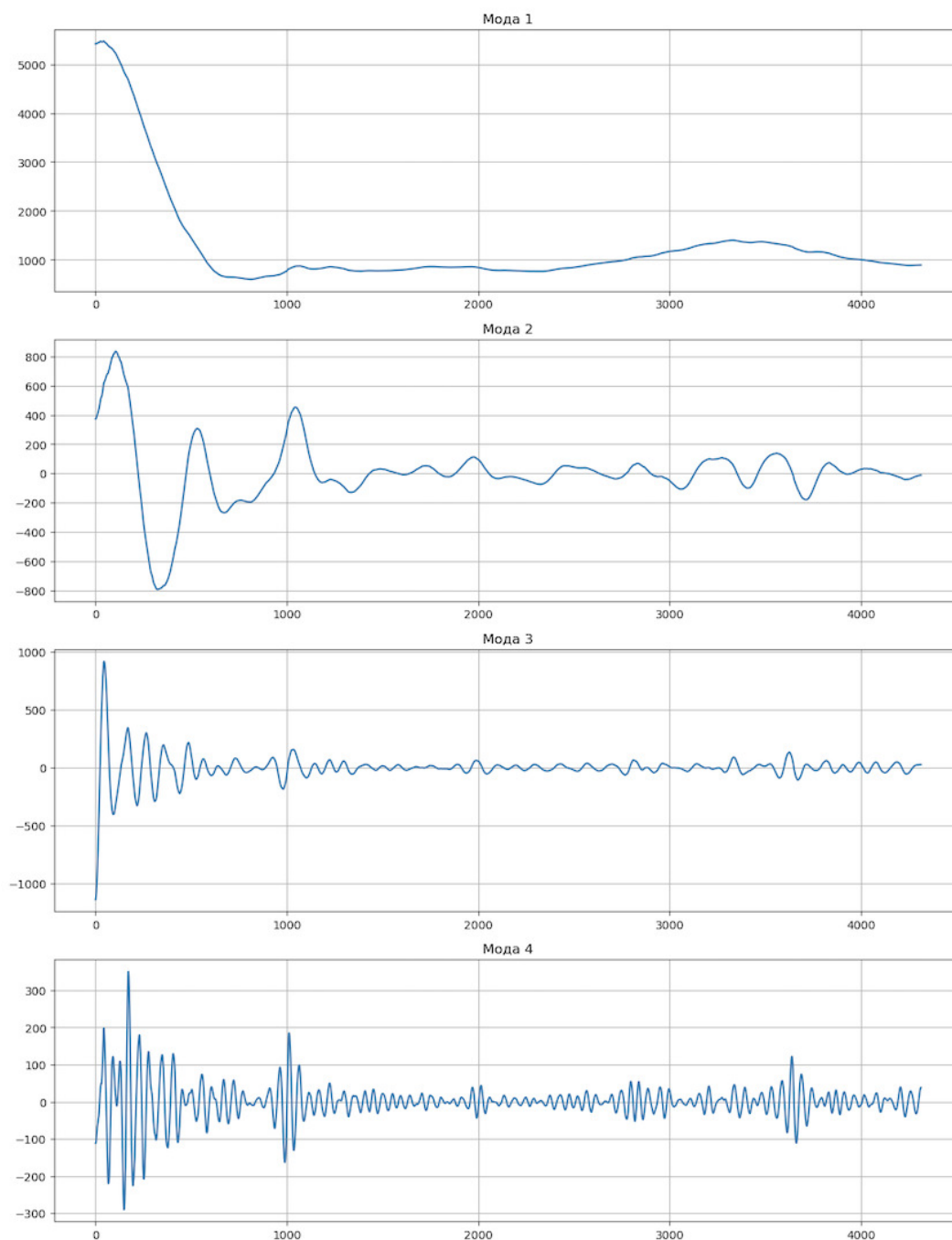


Рис. 5. Первые 4 компоненты разложения ряда с помощью преобразования VMD

3. Li J., Cheng B., Chen Y., Gao G., Shi J., Zeng T., EWT: Efficient Wavelet-Transformer for single image denoising. *Neural Networks*, 2024, vol. 177, p. 106378. doi: 10.1016/j.neunet.2024.106378.
4. Gu J., Tian F., Shi J., Tan F., Noise reduction and analysis of leaf electrical signals of strap-leaved plants based on VMD-EWT. *Computers and Electronics in Agriculture*, 2024, vol. 226, p. 109441. doi: 10.1016/j.compag.2024.109441.
5. Rawat A., Singh M.P., Sharma R.R., Data-driven 2D-EWT based diabetic retinopathy identification using hybrid neural network. *Image and Vision Computing*, 2024, vol. 150, p. 105194. doi: 10.1016/j.imavis.2024.105194.

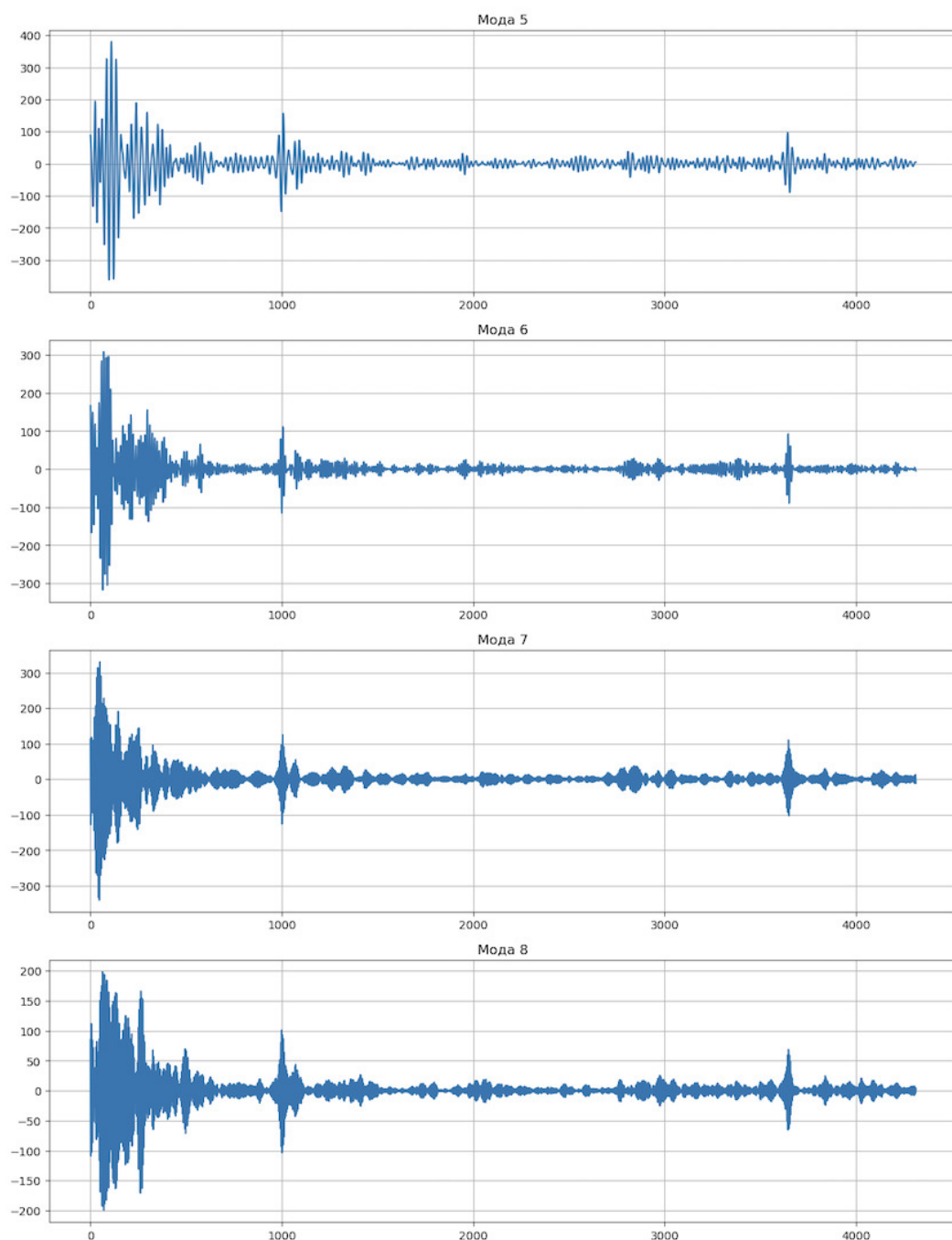


Рис. 6. Вторые 4 компоненты разложения ряда с помощью преобразования VMD

6. Celebi M., Ozturk S., Kaplan K., An emotion recognition method based on EWT-3D-CNN-BiLSTM-GRU-AT model. *Computers in Biology and Medicine*, 2024, vol. 169, p. 107954. doi: 10.1016/j.combiomed.2024.107954.
7. Zhu Y., Xia H., Wang Z., Zhang J., Time Series Prediction of Key Parameters of Pump-Type Machinery in Nuclear Power Plants Based on Empirical Wavelet Transform-Gate Recurrent Unit (EWT-GRU). *Nuclear Engineering and Technology*, 2025, p. 103431. doi: 10.1016/j.net.2024.103431.
8. Karijadi I., Chou S.-Y., Dewabharata A., Wind power forecasting based on hybrid CEEMDAN-EWT deep learning method. *Renewable Energy*, 2023, vol. 218, p. 119357 doi: 10.1016/j.renene.2023.119357.

9. Heddam S., Merabet K., Difi S., Kim S., Ptak M., Sojka M., Zounemat-Kermani M., Kisi O., River water temperature prediction using hybrid machine learning coupled signal decomposition: EWT versus MODWT. *Ecological Informatics*, 2023, vol. 78, p. 102376. doi: 10.1016/j.ecoinf.2023.102376.
10. Li Y., Wu H., Liu H., Multi-step wind speed forecasting using EWT decomposition, LSTM principal computing, RELM subordinate computing and IEWT reconstruction. *Energy Conversion and Management*, 2018, vol. 167, pp. 203–219. doi: 10.1016/j.enconman.2018.04.082.
11. Air Passenger Data for Time Series Analysis. URL: <https://www.kaggle.com/datasets/ashfakyeafi/air-passenger-data-for-time-series-analysis/data>, 23-06-2025
12. Dragomiretskiy K., Zosso D. Variational Mode Decomposition. *IEEE Transactions on Signal Processing*, 2014, vol. 62, pp. 531–534. doi:10.1109/TSP.2013.2288675.
13. Zhu J., Wang C., Hu Z., Kong F., Liu X., Adaptive variational mode decomposition based on artificial fish swarm algorithm for fault diagnosis of rolling bearings. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 2015, vol. 231, pp. 635–654. doi:10.1177/0954406215623311.
14. Huang N.E., Shen Z., Long S.R., Wu M.C., Shih H.H., Zheng Q., Yen N.-C., Tung C.C., Liu H.H., The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. A.*, 1998, vol. 454, pp. 903–995. doi:10.1098/rspa.1998.0193.
15. Zhang Y., Zhong K., Xie X., Huang Y., Han S., Liu G., Chen Z., VMD-ConvTSMixer: Spatiotemporal channel mixing model for non-stationary time series forecasting. *Expert Systems with Applications*, 2025, p. 126535. doi: 10.1016/j.eswa.2025.126535.
16. Zhu Q., Zhang F., Liu S., Wu Y., Wang L., A hybrid VMD–BiGRU model for rubber futures time series forecasting. *Applied Soft Computing*, 2019, vol. 84, p. 105739. doi: 10.1016/j.asoc.2019.105739.
17. Zhao Y., Peng X., Tu T., Li Z., Yan P., Li C., WOA-VMD-SCINet: Hybrid model for accurate prediction of ultra-short-term Photovoltaic generation power considering seasonal variations. *Energy Reports*, 2024, vol. 12, pp. 3470–3487. doi: 10.1016/j.egyr.2024.09.025.
18. Su T., Liu D., Cui X., Dou X., Lei B., Cheng X., Yuan M., Chen R., Prediction of DEDI index for meteorological drought with the VMD-CBiLSTM hybrid model. *Journal of Hydrology*, 2024, vol. 641, p. 131805. doi: 10.1016/j.jhydrol.2024.131805.
19. Fang S., Chen Y., Wu X., Zhao N., Liu Y., Radon exhalation rate prediction and early warning model based on VMD-GRU and similar day analysis. *Journal of Environmental Radioactivity*, 2025, vol. 281, p. 107593. doi: 10.1016/j.jenvrad.2024.107593.
20. Zhang S., Zhao Z., Wu J., Jin Y., Jeng D.-S., Li S., Li G., Ding D., Solving the temporal lags in local significant wave height prediction with a new VMD-LSTM model. *Ocean Engineering*, 2024; vol. 313, no. 1, p. 119385. doi: 10.1016/j.oceaneng.2024.119385
21. Geng D., Zhang Y., Zhang Y., Qu X., Li L., A hybrid model based on CapSA-VMD-ResNet-GRU-attention mechanism for ultra-short-term and short-term wind speed prediction. *Renewable Energy*, 2025, vol. 240, p. 122191. doi: 10.1016/j.renene.2024.122191
22. Makridakis S., Spiliotis E., Assimakopoulos V., The M4 Competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, 2018, vol. 34, no. 4, pp. 802–808. doi:10.1016/j.ijforecast.2018.06.001.
23. Vaswani A., Noam S., Niki P., Attention Is All You Need. *Conference on Neural Information Processing Systems (NIPS 2017)*, 2017.
24. Zhou H., Zhang S., Peng J., Informer: beyond efficient transformer for long sequence time-series forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, vol. 35, no. 12, pp. 11106–11115.
25. Liu D., Wang H., Time series analysis model for forecasting unsteady electric load in buildings. *Energy and Built Environment*, 2024, vol. 5, no. 6, pp. 900–910. doi: 10.1016/j.enbenv.2023.07.003

26. Zhang S., Zhu C., Guo X., Wind-Speed Multi-Step Forecasting Based on Variational Mode Decomposition, Temporal Convolutional Network, and Transformer Model. *Energies*, 2024, vol. 17. doi: 10.3390/en17091996
27. Montero-Manso P., Athanasopoulos G., Hyndman R.J., Talagala T.S., FFORMA: Feature-based forecast model averaging. *International Journal of Forecasting*, 2020, vol. 36, no. 1, pp. 86-92. doi: 10.1016/j.ijforecast.2019.02.011
28. Hyndman R.J., Athanasopoulos G., *Forecasting: principles and practice*. Melbourne: OTexts, 2021 (3rd edition).
29. Stock J.H., Watson M.W., *Introduction to Econometrics*. Pearson, 2019.
30. Wu Z., Huang N. E., Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method. *Advances in Adaptive Data Analysis*, 2009, vol. 1, no. 1.
31. Torres M.E., Colominas M.A., Schlotthauer G., Flandrin P., A complete ensemble empirical mode decomposition with adaptive noise. *Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 4144-4147.
32. Colominas M.A., Schlotthauer G., Torres M.E., Improved complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomed. Sig. Proc. and Control*, 2014, vol. 14, pp. 19-29.
33. Грамович Я.В., Мусатов Д.Ю., Петрусевич Д.А., Применение беггинга в прогнозировании временных рядов. *Российский Технологический Журнал*. 2024, 12, №1, с. 101-110. doi: 10.32362/2500-316X-2024-12-1-101-110.

Hybrid time series models using wavelet transform

**D. V. Bakulin, N. V. Beletskaya, M. I. Djioeva, N. N. Gorshkov,
U. S. Mokhnatkina, D. A. Petrusevich, A. A. Polyakov**

The paper contains comparison of accuracy of the system using wavelets (EWT and VMD transformations), the transformer network that uses results of these transformations as input and standard time series models. The basic idea of these EWT and VMD transformations is to use the Fourier transform and extract information on behavior of time series from their results. At the same time, the transformations can be applied in different ways: entirely to the series, separately to the trend and seasonality and so on. Moreover, the components obtained after decomposition can be either assembled back or used in other models in order to describe the interdependencies in the data of the series. To describe the trend and seasonality, there is a wide set of approaches including STL transformation, ARIMA/ETS statistical models, trend processing using polynomials, decomposition of seasonality with use of Fourier series, etc. These approaches can be combined using various models to describe parts of the series (trend, seasonality and remainder) due to the fact that usually each model can describe only a part of time series behaviour. For example, ARIMA/ETS statistical models perfectly describe seasonality and remainder part having means to describe the trend. Neural network models work better to describe the trend but they've got problems in seasonality description. One can assume that model combination should improve their forecasts. In this paper EWT components are used to describe trend and seasonality. Their results together with the original signal are used as input to neural network models (transformers). The predictions of hybrid models are compared with forecasts of standard models.

KEYWORDS: EWT, VMD, Fourier series, wavelet, transformer, time series.