ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ
ФГАОУ ВО НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

Факультет компьютерных наук
Образовательная программа «Прикладная математика и информатика»

УДК 004.02

Отчет об исследовательском проекте на тему:
Прогнозирование многомерных хаотических временных рядов методами нелинейной динамики

Выполнил студент:

группы #БПМИ234, 2 курса Разухин Александр Сергеевич

Принял руководитель проекта:

Корней Кириллович Томащук
Преподаватель, магистр
Факультет компьютерных наук / Департамент анализа данных и искусственного интеллекта

Москва 2025

Содержание

Аннотация

Работа над данным проектом предполагает исследование методов прогнозирования временных рядов, основанных на использовании информации, полученной из других рядов, которые по некоторым критериям "подходят" для конкретной задачи.

Ключевые слова: Временные ряды, нелинейная динамика, прогнозируемые и непрогнозируемые точки

1 Введение

1.1 Релевантность

Современные научные и инженерные задачи требуют все более точных методов прогнозирования сложных систем, и особое внимание уделяется многомерным хаотическим временным рядам. Они встречаются в самых разных сферах — от метеорологии и финансов до биологии и механики. Проблема в том, что традиционные методы часто не справляются с прогнозированием таких систем из-за их высокой чувствительности к начальным условиям и нелинейного поведения. Поэтому активно развиваются подходы, основанные на методах нелинейной динамики, которые помогают находить скрытые закономерности и создавать более точные модели для предсказания будущих изменений.

1.2 Цели и задачи

Целью данной работы является исследование методов нелинейной динамики для прогнозирования многомерных хаотических временных рядов, а также оценка их эффективности. Для достижения поставленной цели необходимо решить следующие задачи:

1. Провести анализ существующих методов прогнозирования хаотических временных рядов.
2. Рассмотреть основные подходы нелинейной динамики, применяемые в данной области.
3. Разработать алгоритмы прогнозирования и протестировать их на примерах многомерных временных рядов.
4. Оценить точность и устойчивость предложенных методов на примере несинтетических рядов.

2 Обзор литературы

Хаотичные системы, как в природе, так и в социальных явлениях, активно исследуются в контексте прогнозирования временных рядов. На сегодняшний день существуют успешные алгоритмы для предсказания на один шаг вперед, но для того, чтобы предсказать сразу много шагов (MSA), задачи предсказания остаются сложными. Это связано с экспоненциальным ростом ошибки предсказания с увеличением горизонта прогноза, что отражает нестабильность Ляпунова, присущую хаотичным системам [6], которая не зависит от того, насколько мала начальная разница между соседними траекториями. Такая неустойчивость приводит к 'горизонту прогнозирования', который при заданной ошибке наблюдения $\epsilon(0)$, максимальной допустимой ошибке ϵ_{max} и экспоненциальном росте ошибки $e^{\lambda x}$ вычисляется как $T \approx \frac{1}{\lambda} \cdot \ln\left(\frac{\epsilon_{max}}{\epsilon(0)}\right)$ [7]. Таким образом, для прогнозирования существует теоретический предел, который ограничивает точность прогноза для более чем нескольких шагов вперед.

Предсказательная кластеризация [4] позволяет преодолеть некоторые из этих проблем. Этот метод использует повторяющиеся последовательности данных (мотивы) для прогнозирования будущих значений на основе схожих участков временного ряда. Если участок ряда "похож"на начало мотива с некоторой точностью, то можно предполагать, что дальше он будет вести себя подобно известному мотиву. Получаются мотивы кластеризацией векторов непоследовательных значений в z-векторы [9]. В отличие от глобальных моделей, предсказательные методы кластеризации строят локальные модели для каждого мотива, что улучшает точность предсказания [8]. Важно, что предсказания делаются не для всех точек, а для их части, значения в которых удовлетворяют некоторому критерию, остальные же точки помечаются как непрогнозируемые и не влияют на дальнейшие предсказания, что также может улучшить результаты. Данная стратегия довольно естественная, поскольку, к примеру, инвесторы не совершают действия каждый момент времени, а лишь тогда, когда удаётся точно спрогнозировать поведение рынка [1].

Кроме того, для более сложных случаев могут использоваться обобщенные z-векторы [9], которые представляют собой комбинацию непоследовательных наблюдений с заданными шаблонами. Шаблон определяется как последовательность расстояний между наблюдениями в ряде, что в сгенерированном векторе эти наблюдения становятся последовательными. Например, если вектор из $k+1$ последовательных значений соответствует шаблону $(\underbrace{1, \ldots, 1}_{k})$, то, заменив одну или несколько единиц на другие натуральные числа, вектор уже не будет являться подотрезком ряда, а лишь подпоследовательностью. Для каждого шаблона независимо строится список векторов, которые соответствуют данному шаблону, и все такие выборки кластеризуются по отдельности. Этот метод позволяет улучшить предсказания, поскольку при фиксированной длине шаблона k, если рассматривать только вектора из последовательных значений, есть всего один шаблон, а если взять $L > 1$ - максимальное значение элемента, то таких шаблонов будет уже L^k, что значительно увеличивает размер выборки для обучения. Также эта стратегия позволяет прогнозировать значения в позиции, даже если прямо перед ней были одна или несколько непрогнозируемых позиций, что достигается введением $L > 1$.

Таким образом, использование предсказательных кластеризационных методов, обобщенных z-векторов и различных подходов для выявления непрогнозируемых точек дает возможность получать предсказания для достаточно большого количества позиций временных рядов с выбранной точностью, даже когда классические методы не могут справиться с долгосрочным прогнозом.

3 Обзор существующих методов

В последнее время вышло множество работ, посвящённых прогнозированию временных рядов, однако бóльшая их часть описывает предсказание на один шаг вперёд, тогда как исследований, занимающихся прогнозированием на много шагов вперёд, намного меньше. Такая разница связана с ошибкой, растущей экспоненциально с увеличением горизонта прогнозирования.

Обычно, алгоритм прогнозирования на много шагов вперёд состоит из двух этапов: техника прогнозирования на один шаг и стратегии, которая используется для преобразования прогноза на один или несколько шагов в прогноз на много шагов вперёд. Для этих целей может быть использовано множество подходов, применяются концепции из практически всех областей машинного обучения и анализа данных: регрессия опорных векторов [2], расширенные свёрточные сети [10], кластерные центры в предиктивной кластеризации [5] и многие другие.

Ещё одним немаловажным фактором является стратегия предсказания MSA. Итерационная стратегия предполагает последовательное прогнозирование точек, основываясь на уже предсказанных значениях, и не вычисляет прогнозные значения для промежуточных точек. Прямая же стратегия [3] используется для немедленного получения результатов и не предполагает прогнозирования значений в промежуточных точках; она обеспечивает сразу множество прогнозов для прогнозируемой точки. Эти стратегии являются базовыми, и почти все исследования используют одну из них или применяют гибридные методы, основанные на обеих. Однако, разработанные в рамках этих стратегий методы также не защищены от уже упомянутой экспоненциально растущей ошибке прогнозирования, поэтому исследователи постоянно прикладывают усилия для создания новых стратегий.

В одном из обзоров [3] сравниваются базовые и новые стратегии (DirRec, MIMO, DIRMO). Стратегия DirRec является гибридом базовых, однако итеративно увеличивает число входов, чтобы учитывать значения только предсказанных позиций. MIMO же (Multiple Input Multiple Output) предполагает формирование массива значения для всех точек из горизонта предсказания, не ограничиваясь единственным значением, соответствующем горизонту предсказания, что позволяет выявлять закономерности и повышать качество прогнозирования. DIRMO, являющаяся гибридом DirRec и MIMO, делит ряд на блоки, к каждому из которых применяет стратегию MIMO.

Одной из неклассических стратегий, которая не так давно получила развитие, стала идея введения непрогнозируемых точек, предсказанные значения в которых не стоит учитывать, чтобы при прогнозировании на много шагов вперёд сохранять разумную точность [1].

4 Методы исследования

Методика обучения – использование z-векторов – частично была рассмотрена в обзоре литературы, поскольку эта идея не является новой. Прогноз делается, как уже было отмечено в обзоре существующих методов, на основании большого количества возможных значений. При прогнозировании точки T, для каждого шаблона (k_1, k_2, k_3, k_4) берётся кусок ряда таким образом, что последний элемент мотива имеет индекс T, хоть это значение ещё не спрогнозировано. Далее этот вектор длины L сравнивается со всеми префиксами собранных мотивов, каждый из которых имеет длину $L-1$, и в случае, если он отличается от префикса (по норме) не больше, чем на фиксированный ϵ, последний элемент мотива учитывается при прогнозировании интересующей точки. В качестве модели для прогноза использовалась мода *scipy.stats.mode*. В этом же блоке применялась концепция непрогнозируемых точек, то есть в случае отклонения прогнозного значения от реального более чем на δ точка помечалась как непрогнозируемая, и при анализе следующих точек не учитывалась (в случаях, когда сформированный мотив с окончанием в точке T содержал непрогнозируемую точку, данный шаблон пропускался). Данный подход позволяет прогнозировать на много шагов вперёд, не теряя точности.

Важным аспектом при анализе любых методов является оценка качества полученной модели. В данной работе было использовано ансамблевое усреднение прогнозов. Заключается этот подход в том, что для оценки качества строятся несколько независимых прогнозов с разными начальными условиями (в данном случае – смещённые во времени), а затем результаты усредняются по каждой временной точке, то есть если предсказания делаются M раз для N точек, то, к примеру, итоговая метрика $MAPE$ (Mean Absolute Percentage Error) для первой точки будет определяться как среднее $MAPE$ по точкам, которые были первыми на каком-либо из M шагов. Данный подход позволяет оценить модель более стабильно, без случайных колебаний. Для оценки методов прогнозирования использовались следующие значения: $M = 20$, $N = 50$, $\epsilon = 0.05$, $\delta = 0.05$, ошибка $MAPE$ и также доля непрогнозируемых точек за M шагов для каждого $i = 1, \ldots, N$.

5 Методы прогнозирования

Основная идея нелинейной динамики – использование других рядов для прогнозирования интересующего ряда. Данный подход не только увеличивает выборку, но и позволяет модели замечать поведение, характерное сразу для нескольких рядов, и использовать это при прогнозировании. Мой анализ состоял в том, чтобы проверить, следует ли из близости двух рядов по какой-либо метрике возможность использовать данные одного для прогнозирования другого. С данной целью мною были рассмотрены следующие подходы для анализа временных рядов.

5.1 Критерий Колмогорова-Смирнова

Один из основных подходов для сравнения двух выборок на предмет одинакового распределения – это непараметрический критерий Колмогорова-Смирнова, который анализирует расстояние между распределениями двух временных рядов. Метрикой в этом случае является максимальное расстояние между значениями распределений в каждой точке.

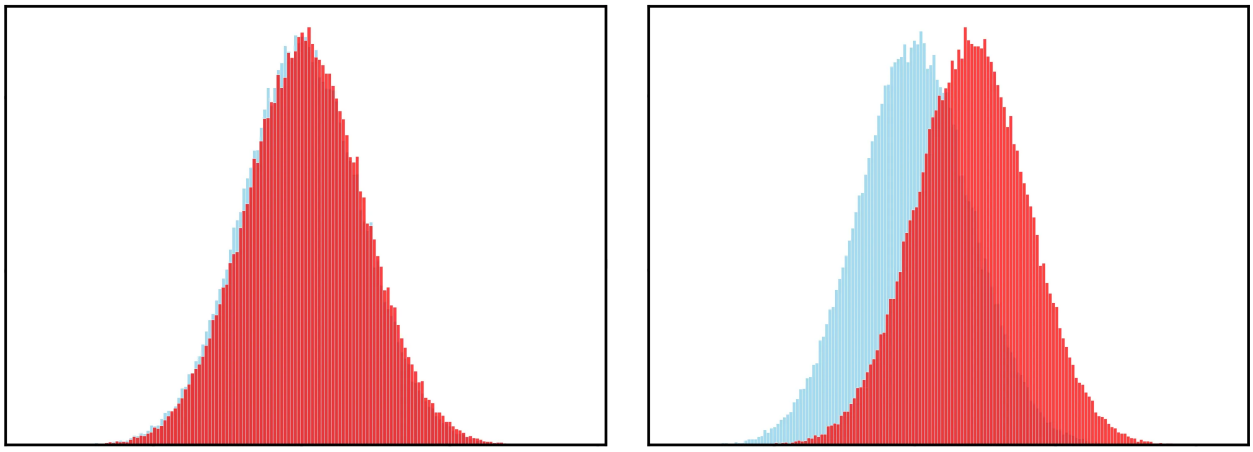

Рис. 1: Kolmogorov-Smirnov visualisation

5.2 Dynamic Time Warping

Метод DTW (Dynamic Time Warping) сравнивает временные ряды напрямую, но, в отличие от, к примеру, стандартного евклидова расстояния, он учитывает временные искажения при сравнении рядов (растяжение / сжатие, см. рис.) и теоретически может дать более точную оценку "похожести"рядов. Метрикой в этом случае также является максимальное расстояние между точками, между которыми алгоритм установил соответствие при анализе.

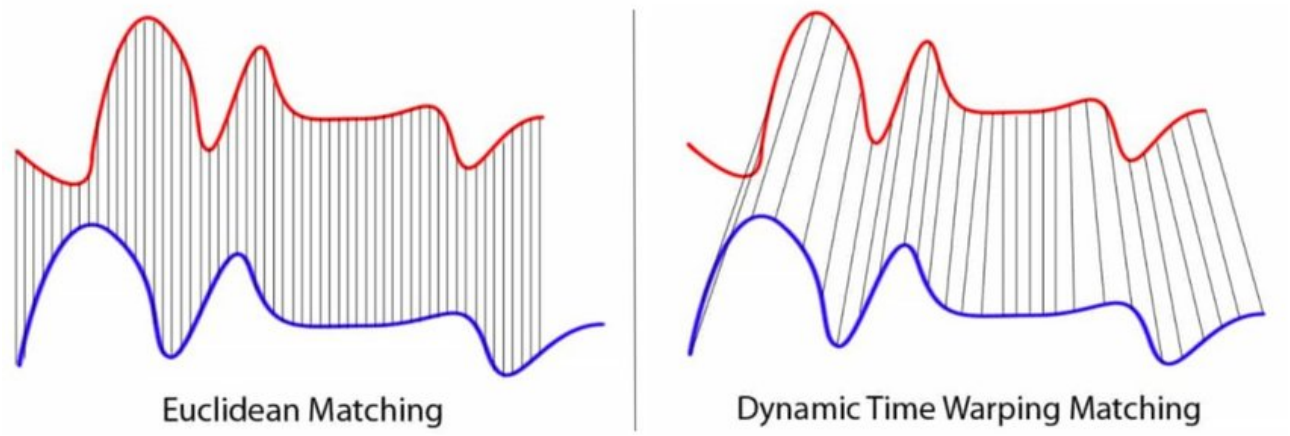

Рис. 2: DTW visualisation

5.3 Анализ ординальных структур

Некоторые методы сравнивают не временные ряды напрямую, а преобразовывают их в иные структуры и сводят анализ рядов к анализу этих структур. Одними из более распространённых являются ординальные структуры. Сам переход от ряда заключается в том, что каждый мотив заменяется перестановкой, которая, при её применении к данному мотиву, сортирует его элементы (к примеру, $(0.1, 0, 0.2) \to (1, 0, 2)$). Далее формируется вектор $p \in \mathbb{R}^{(L+1)!}$, где $i-$ая координата – доля $i-$ой перестановки среди всех.

Изначально у меня было 2 принципиально разных подхода для сравнения вероятностных векторов. Во-первых, обычная норма, то есть $\|p_1 - p_2\|_F$. Во-вторых, более продвинутый подход, использующий энтропию векторов как рядов: $H\left(\frac{p_1+p_2}{2}\right) - \frac{1}{2}H(p_1) - \frac{1}{2}H(p_2)$. Однако, после проведения сравнительного анализа этих двух подходов оказалось, что позиции каждого из 125 сгенерированных рядов Лоренца при сортировке по каждой из приведённых метрик сильно скоррелированы (см. рис.). Таким образом, я принял решение рассматривать эти два подхода вместе, учитывая каждый из них с весом 0.5 при анализе близости конкретного ряда.

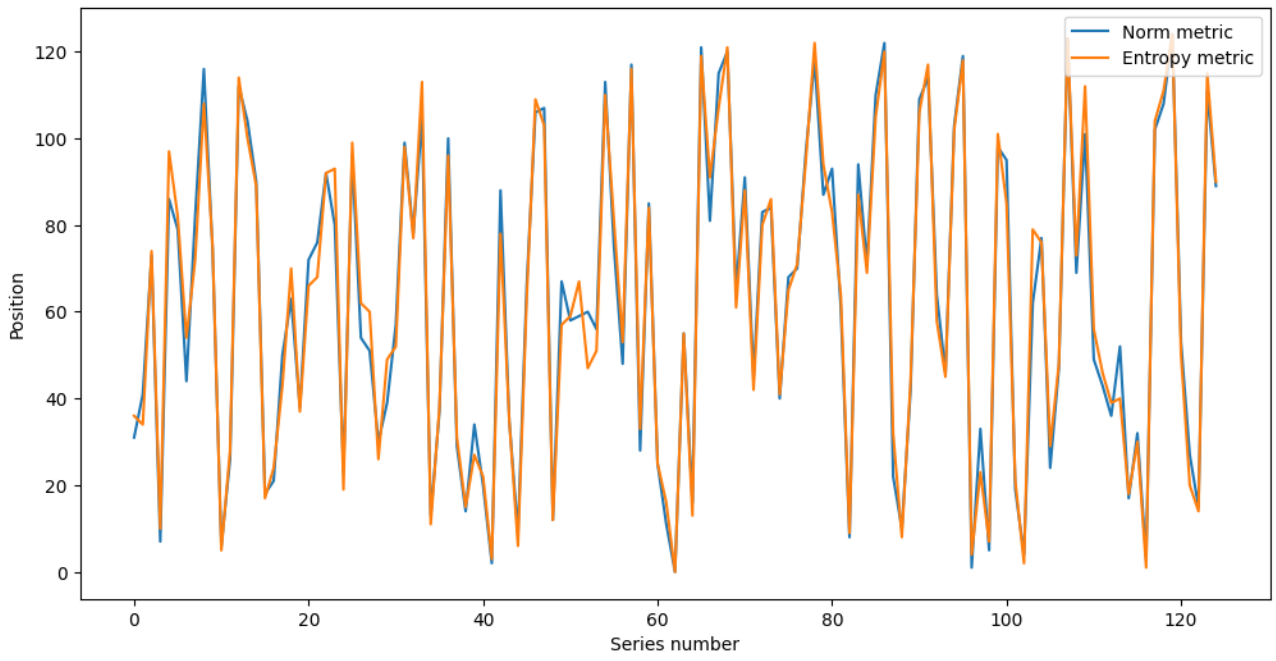

Рис. 3: Different metrics based on probability vectors comparison

6 Исследование синтетических временных рядов

В качестве данных для разработки алгоритмов прогнозирования было решено использовать аттрактор Лоренца, задаваемый системой дифференциальных уравнений:

$$\begin{cases} \dot{x} = \sigma(y - x) \\ \dot{y} = x(\rho - z) - y \\ \dot{z} = xy - \beta z \end{cases}$$

Ряд же Лоренца получается при проектировании аттрактора на плоскость Oxy:

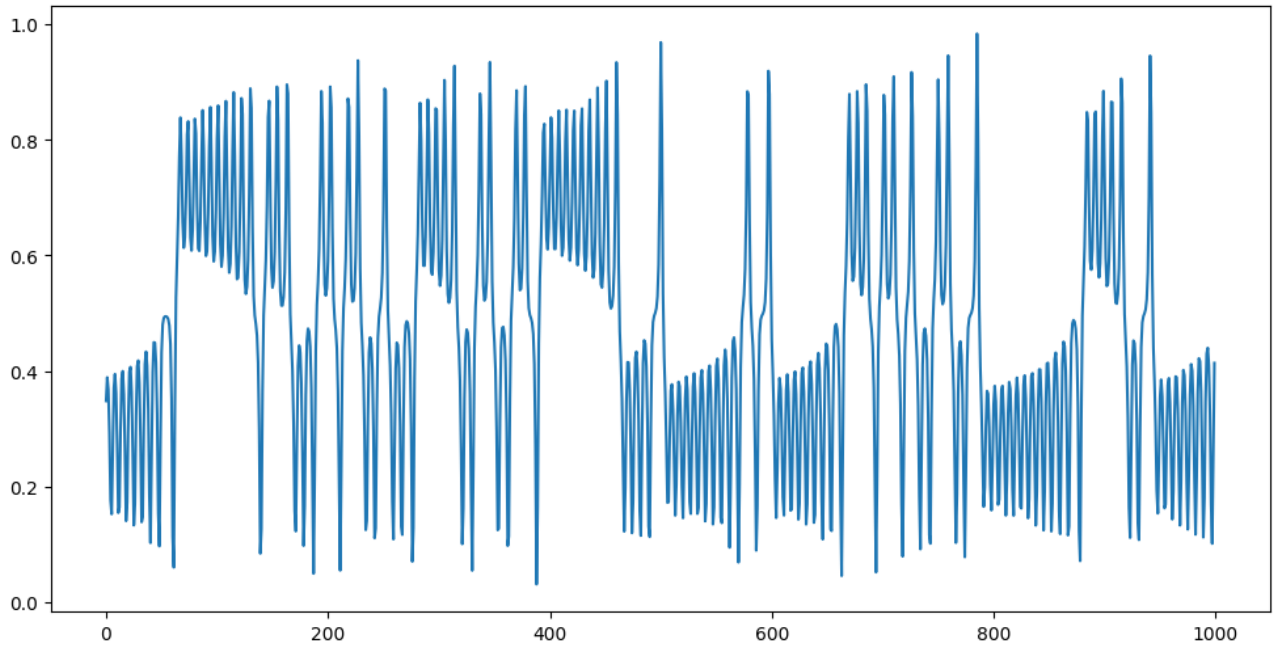

Рис. 4: Lorenz time series

Выбор именно этого ряда был сделан по нескольким причинам. Во-первых, этот синтетический ряд не содержит шумов и выбросов, но в то же время обладает сложной нелинейной структурой. Во-вторых, его динамика характерна для физических процессов, что позволяет экстраполировать результаты, полученные для ряда Лоренца, на реальные задачи. В-третьих, в большинстве научных исследований используется именно этот ряд, так что его использование упрощает сравнение результатов с существующими методами прогнозирования.

Мой анализ здесь и далее был основан на том, что, зафиксировав начальные параметры в системе, порождающей ряд Лоренца $(10, 28, 8/3)$, я генерировал ряды, в которых один или несколько значений отличались на 1-2% от оригинальных (из-за чувствительности ряда Лоренца к начальным параметрам порождаемые временные ряды сильно отличались от изначального), а затем анализировал «ближайшие» (по метрике) ряды и добавлял данные для обучения в оригинальный ряд для повышения точности предсказания.

Мой подход был основан на том, что я постепенно добавлял новые ряды к оригинальному, добиваясь всякий раз улучшения качества прогноза, выбирая очередной ряд из ближайших по выбранной метрике. Для каждого подхода было проведено не менее 15 операций по добавлению данных из других рядов с целью добиться наилучшего результата. Ниже приведены результаты исследования для каждого из подходов - показаны результаты применения нелинейной динамики в сравнении с прогнозированием лишь на оригинальном ряде.

Критерий Колмогорова-Смирнова

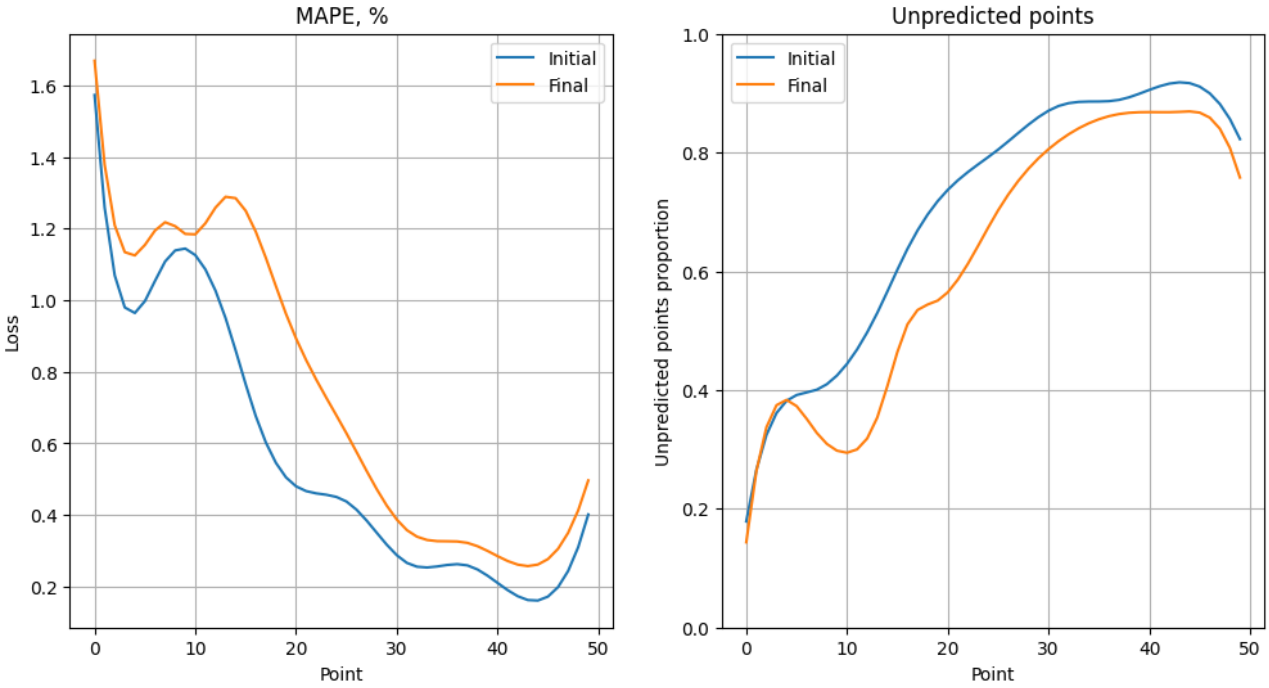

Рис. 5: Original series (train=1000) | Colmogorov mixing (train=5000)

Таким образом, удалось уменьшить число непрогнозируемых точек при небольшом увеличении ошибки, а значит, этот метод в целом применим в проверке на то, что ряд "подходит"для добавления данных из него.

Dynamic Time Warping

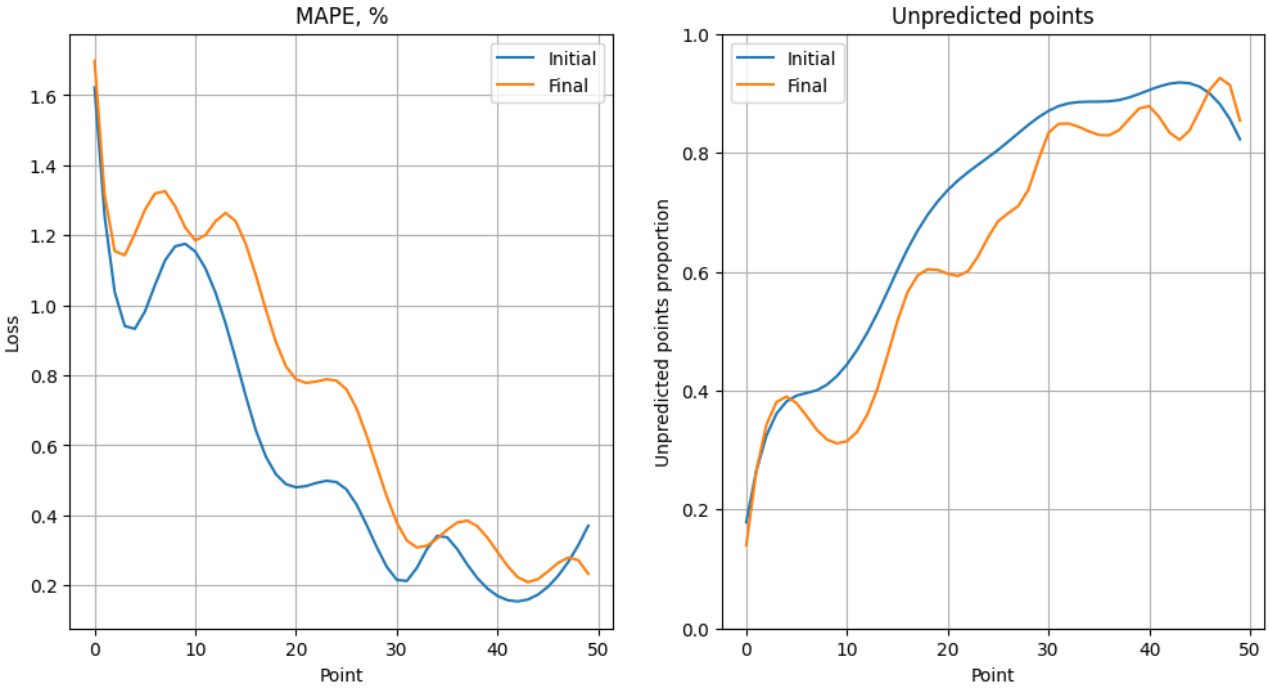

Рис. 6: Original series (train=1000) | DTW mixing (train=5000)

Можно заметить, что для этого метода результаты даже немного хуже, чем для критерия Колмогорова-Смирнова, хотя теоретически этот метод более точный, посольку учитывает поведение рядов. Я могу объяснить этот факт тем, что у ряда Лоренца сложная

структура, которая, по-видимому, не может быть точно проанализирована методом DTW, и даже обычное евклидово расстояние оказывается более репрезентативной метрикой.

Анализ ординальных структур

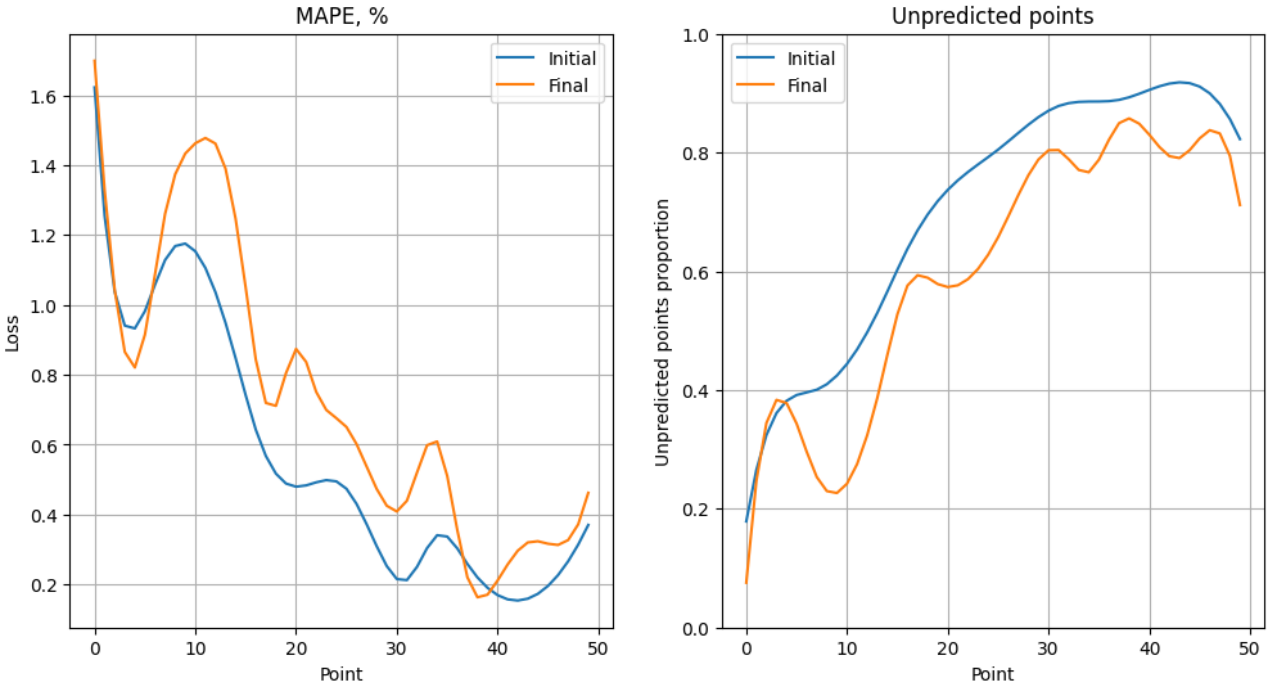

Рис. 7: Original series (train=1000) | Ordinal Structures mixing (train=5000)

Можно заметить, что ошибка в данном случае выросла сильнее, но число непрогнозируемых точек заметно уменьшилось в сравнении с прошлым методом, что особенно видно в конце, когда по экспоненциальному закону ошибки график должен стремиться к 1 (большинство точек непрогнозируемые), однако тот лишь слегка превышает 0.8.

7 Исследование реальных временных рядов

В качестве реального ряда для оценки полученных методов было решено использовать временные ряды потребления электроэнергии в разных странах. Для того, чтобы нелинейная динамика давала прирост качества, необходимо, чтобы ряды были не только были примерно похожи, но и чтобы они имели структуру, похожую на ряды Лоренца. Данное свойство позволило бы успешнее перенести результаты исследования для ряда Лоренца на реальный ряд. Таким образом, идеальным решением стал ряд потребления электроэнергии, в котором промежуток между соседними точками составляет 1 час. Такие ряды, как нетрудно понять, имеют амплитуду, обусловленную пониженным потреблением ночью и повышенным потреблением днём. Ниже приведены части временных рядов для 20 европейских стран. Использование нормализации обусловлено тем, что в разных странах в силу различных факторов разные значения потребления, так что для анализа необходимо было ограничить все ряды отрезком $[0, 1]$.

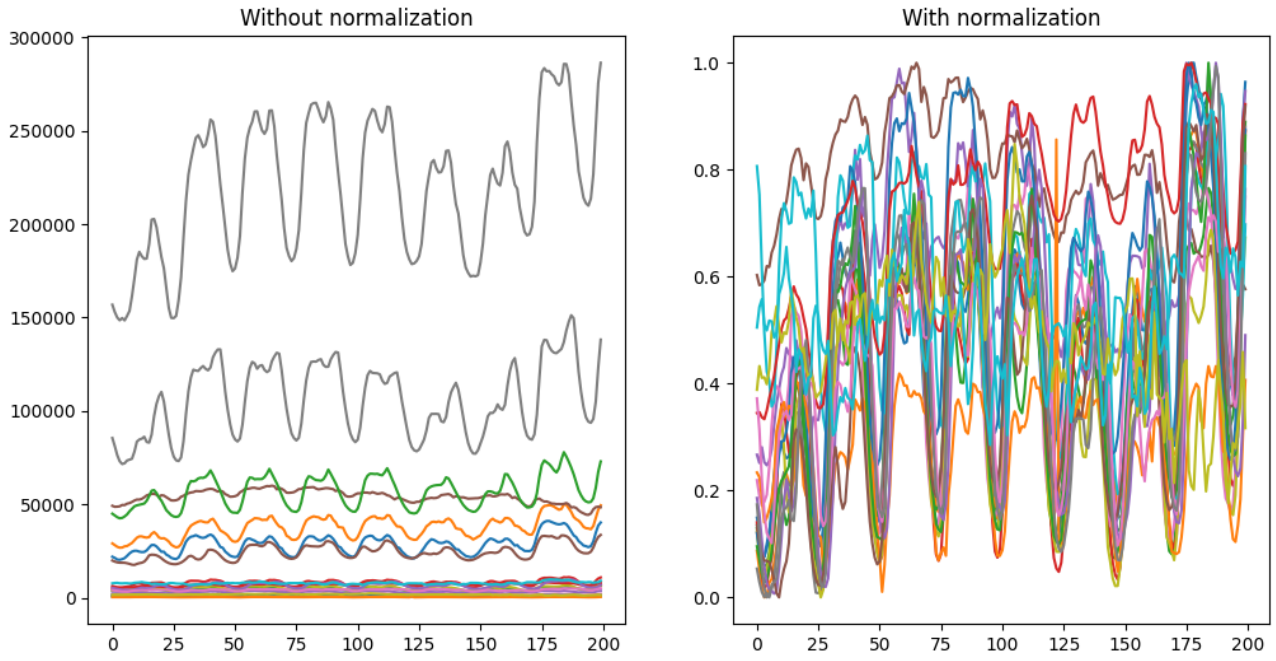

Рис. 8: Electricity consumption in Europe in 2024

Далее я применил метод, который лучше всех показал себя при прогнозировании ряда Лоренца - анализ ординальных структур. В результате анализа 10 добавлений данных из близких по метрике рядов были получены следующие результаты.

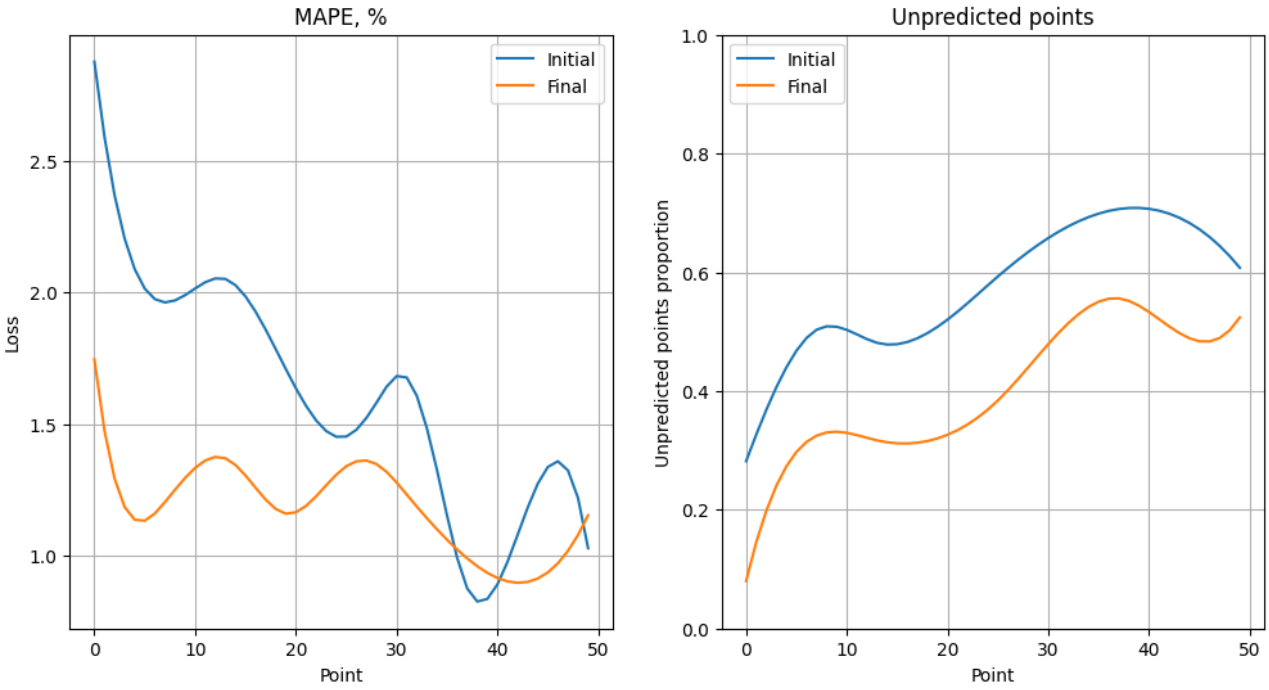

Рис. 9: Original series (train=1000) | Ordinal Structures mixing (train=5000)

Таким образом, добавление других рядов положительно сказалось как на количестве непрогнозируемых точек, так и на ошибке. А значит данный метод вполне применим не только на синтетических рядах таких, как ряд Лоренца, но и на реальных данных, которые меньше поддвержены математическим законам и являются, по большей части, хаотическими.

8 Заключение

Итак, в результате работы над данным проектом я сделал следующие выводы

1. Применение методов нелинейной динамики значительно улучшило качество прогнозирования многомерных хаотических временных рядов. Использование данных, полученных из других временных рядов, привело к увеличению обучающей выборки и, как следствие, повышению точности прогнозов при фиксированной максимальной допустимой ошибке.

2. Рассмотренные методы - критерий Колмогорова-Смирнова, Dynamic Time Warping, анализ ординальных структур - продемонстрировали эффективность при прогнозировании. В то время как метод DTW теоретически должен был показать лучшее качество в сравнении с критерием Смирнова, поскольку учитывает растяжение и сжатие рядов, последний оказался более успешным при прогнозировании. В то же время анализ ординальных структур существенно повысил качество прогнозов, если сравнивать с упомянутыми выше методами, что особенно важно при предсказаниях на много шагов вперёд.

3. Исследование продемонстрировало, что результаты, полученные на синтетическом ряде Лоренца, могут быть успешно экстраполированы на реальные ряды такие, как данные о потреблении электроэнергии. Наличие сложных структурных особенностей, характерных как для реальных, так и для синтетических рядов, подтверждает потенциал методов нелинейной динамики для их применения на практике.

4. Теоретические выводы говорят о том, что увеличение объёма обучающей выборки может привести к более точным и стабильным прогнозам. Так что дальнейшее исследование может быть направлено на оптимизацию различных алгоритмов прогнозирования и разработку новых методов, которые бы объединяли преимущества разных подходов для решения задач прогнозирования хаотических систем с целью возможности расширения обучающей выборки

Таким образом, проделанная работа доказывает, что методы нелинейной динамики являются мощным инструментом для прогнозирования многомерных хаотических временных рядов, а также имеют большой потенциал для практического применения в различных областях, в которых требуются стабильно точные прогнозы.

Список литературы

[1] Gromov V. A. and Borisenko E. A. Predictive clustering on nonsuccessive observations for multi-step ahead chaotic time series prediction. *Neural Computing and Applications*, 26(8):1827–1838, 2015. URL: https://doi.org/10.1007/s00521-015-1845-8.

[2] Xiong T. Bao Y. and Hu Z. Multi-step-ahead time series prediction using multiple-output support vector regression. *Neurocomputing*, (129):482–493, 2014. URL: https://doi.org/10.1016/j.neucom.2013.09.010.

[3] Atiya A. F. Ben Taieb S., Bontempi G. and Sorjamaa A. A review and comparison of strategies for multi-step ahead time series forecasting based on the nn5 forecasting competition. *Computational and Applied Mathematics*, 39(8):7067–7083, 2012. URL: https://doi.org/10.1016/j.eswa.2012.01.039.

[4] De Raedt L. Blockeel H. Top-down induction of first-order logical decision trees. *Artificial Intelligence*, 101(1-2):285–297, 1998. URL: https://doi.org/10.1016/S0004-3702(98)00034-4.

[5] Konev A. S. Gromov V. A. Precocious identification of popular topics on twitter with the employment of predictive clustering. *Neural Computing and Applications*, 28(11):3317–3322, 2017. URL: https://doi.org/10.1007/s00521-016-2256-1.

[6] Schreiber T. Kantz H. Nonlinear time series analysis (2nd ed.). *Cambridge University Press*, 2003. URL: https://doi.org/10.1017/CBO9780511755798.

[7] Potapov A. Malinetskii G. *Modern Problems of Nonlinear Dynamics*. Editorial URSS, 2003.

[8] Ben Taieb S., Sorjamaa A., and G Bontempi. Multiple-output modeling for multi-step-ahead time series forecasting. *Neurocomputing*, 73(10-12):1950–1957, 2010. URL: https://doi.org/10.1016/j.neucom.2009.11.030.

[9] M. Small. Applied nonlinear time series analysis: applications in physics, physiology and finance. *In World Scientific series in nonlinear science*, Series A(Issue 52), 2005.

[10] Z. Wang R. Peng C. Gao J. Gao and Jiang H. A dilated convolution network-based lstm model for multi-step prediction of chaotic time series. *Computational and Applied Mathematics*, 39(1):3317–3322, 2020. URL: https://doi.org/10.1007/s40314-019-1006-2.