

Применение кригинга при прогнозировании временных процессов

Н. В. Белецкая^{*,**}, С. А. Конанов^{*}, Д. А. Петрусевич^{*}, А. В. Стояновски^{*}

^{*}Российский Технологический Университет (МИРЭА), Москва, Россия

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

Поступила в редколлегию 02.09.2025 г. Принята 10.10.2025 г.

Аннотация—В представленной работе произведено сравнение применения кригинга для моделирования и прогнозирования временных рядов со стандартными методами (ARIMA/ETS). Исторически концепция кригинга была построена в интерполяции многомерной функции в приложении к моделированию пространственного распределения полезных ископаемых и (при включении в модели времени) геологических процессов. Если отвлечься от геологической природы данных, то этот подход можно применить и к моделированию произвольной многомерной величины (постоянной или меняющейся во времени). Интерес к этому методу связан, в частности, с тем, что настраиваемая модель относится к очень простому классу – к линейной или полиномиальной регрессии. В рамках представленного исследования решаются задачи прогнозирования на некоторый промежуток времени в будущем, а также обработка удалённых значений (сконцентрированных в одном месте или разбросанных по данным ряда). Построены комбинированные модели, в которых предсказание даётся на основе взвешенного голосования нескольких моделей. При этом, каждая следующая модель обучается на остатках первой модели (например, так работает беггинг). Исследуется комбинация моделей ARIMA, ETS и кригинга

КЛЮЧЕВЫЕ СЛОВА: Кригинг, временные ряды, ARIMA, ETS.

DOI: 10.53921/18195822_2025_25_3_261

1. ВВЕДЕНИЕ

Модель кригинга в применении к временным рядам представляет собой модель регрессии следующего вида [1]:

$$X(t) = \sum_{i=1}^k \beta_i f_i(x, t) + z(x, t),$$

где k – размерность задачи, β_i – параметры регрессии, задающей моделируемую величину, $f(x, t)$ – функции, по которым раскладывается величина $X(t)$. Чаще всего $z(x, t)$ – гауссовский процесс со средним значением 0 и вариацией σ^2 . В связи с применением регрессии модель настраивается на минимизацию отклонения прогнозируемой с помощью модели величины от реальных значений. При этом, чаще всего используют квадрат ошибки, хотя можно применить и другую норму в зависимости от природы данных в задаче. Зачастую в минимизируемую функцию добавляют слагаемые для регуляризации решения задачи [2].

Метод кригинга появился в геостатистике и впервые был описан Д. Кригом [3] в работе о статистической оценке запасов золота на месторождениях Витватерсранда, где предложил метод взвешенного усреднения для минимизации ошибок прогноза. В 1960-х гг. Ж. Матерон

разработал теоретическую базу кригинга и доказал, что эта модель обеспечивает наилучшую линейную несмещенную оценку [4, 5]. Данное исследование посвящено возможностям приложения модели Кригинга для прогнозирования временных рядов. Ранее исследователями было предложено две модели, решающие эту задачу. Основная идея первого подхода – использовать в качестве базисных функции вида $\cos(ax + \varphi)$, что роднит метод с преобразованием Фурье. Параметры функций определяются на основе спектрального анализа моделируемых данных [6, 7]. Второй подход предлагает использование в качестве ковариационной функции метрику особого вида, а в качестве базисных функций использовать $x^k, k \in N$, что по сути связано с разложением в ряд Тейлора [8]. Проблема этих подходов заключается в небольшом горизонте прогнозирования, который следует из небольшого радиуса сходимости разложения функции в ряд Тейлора. Кроме того, полученные модели относительно сложно интерпретировать [9–11].

При переходе к задаче, зависящей от времени, в кригинге пространственные параметры фактически заменяются на зависимость от времени. Схема построения моделей на основе кригинга для обработки временных процессов выглядит следующим образом. Временной процесс разбивается на три части: тренд, шум, сезонность [2, 12, 13]. Например, ряд $Z(t) = \mu(t) + s(t) + \varepsilon(t)$, разбивается на $\mu(t)$ – функцию, показывающую среднее поведение процесса Z , сезонные или циклические изменения $s(t)$, а также ε – случайную ошибку с нулевым математическим ожиданием. Предположим, мы знаем значение функции в N точках $t_1, t_2, \dots, t_N$. Кригинг использует линейную комбинацию известных значений для оценки значения функции в неизвестных точках: $\hat{Z}(t) = \sum_{i=1}^N \lambda_i Z(t_i)$, где коэффициенты подобраны таким образом, что $\hat{Z}(t)$ – несмещенная оценка для $Z(t)$ и дисперсия ошибки $\varepsilon(t)$ минимальна [14–16].

Для описания сезонной части в модели кригинга используется ковариационная функция. Она описывает взаимосвязь между значениями, находящимися на расстоянии d друг от друга. Из тех же соображений, как это проводится при настройке модели экспоненциального сглаживания, потребуем, чтобы $\lim_{d \rightarrow \infty} C(d) = 0$. Т.е. влияние данных, зафиксированных в определённый момент времени, уменьшается с ростом времени, прошедшего с этого момента. Помимо этого, есть измерения, которые должны влиять на прогноз сильнее, чем остальные: значения за прошлые сезоны. То есть функция должна иметь точки локального максимума с расстоянием L (соответствует длине сезона) между ними. Кроме того, из теоремы Бохнера-Хинчина следует, что для того, чтобы функция могла являться ковариационной, она должна быть положительно полуопределенной, т.е. для любого набора точек $t_1, \dots, t_n$ и любого набора действительных чисел $c_1, \dots, c_n$ выполнено неравенство $\sum_{i=0}^n \sum_{j=0}^n c_i c_j C(|t_i - t_j|) \geq 0$.

Далее следует подобрать вид функций, которые будут использоваться в качестве ковариационных. Одна из них – это функция $C(d) = \exp(-\alpha d) \cos^{2k}(d\pi/L)$, где $\alpha > 0$ – действительная константа и $k > 1$ – целая константа. Первые два требования, очевидно, выполняются. Докажем положительную полуопределенность.

Теорема Бохнера помимо ограничений на ковариационную функцию вводит критерий положительной полуопределенности функций: функция положительно полуопределена тогда и только тогда, когда преобразование Фурье этой функции – неотрицательная функция. Воспользуемся следствием теоремы о свертке: преобразование Фурье произведения функций с точностью до множителя $\frac{1}{2\pi}$ равно свертке образов этих функций (т.е. $2\pi \mathcal{F}\{a(t)b(t)\}(\omega) = \mathcal{F}\{a(t)\}(\omega) * \mathcal{F}\{b(t)\}(\omega)$). Из определения свертки нетрудно понять, что свертка двух неотрицательных функций есть неотрицательная функция. Тогда нужно доказать неотрицательность образов функций $\exp(-a|t|)$ и $\cos^{2k}(|t|\pi/L)$.

Функция $\exp(-a|t|)$ – это функция Орнштейна-Уленбека (стационарный процесс Гаусса-Маркова), известная ковариационная функция. В силу того, что она ковариационная, для нее

выполнено условие положительной полуопределенности, а следовательно и неотрицательности преобразования Фурье.

Рассмотрим функцию $\cos^{2k}(|x|\pi/L)$. Сделаем замену переменных и сведем эту функцию к функции $\cos^{2k}(t)$. Воспользуемся уже упомянутым свойством преобразования Фурье: преобразование Фурье произведения функций – это свертка преобразования функций с точностью до множителя. Преобразование Фурье функции $\cos x$ – это сумма двух дельта функций Дирака, которая в свою очередь, очевидно, неотрицательна. При возведении этой функции в степень k , ее образ будет являться сверткой k неотрицательных функций, что так же будет неотрицательной функцией. Следовательно, преобразование Фурье начальной функции – это неотрицательная функция.

Таким образом, образ Фурье функции $C(d) = \exp(-\alpha d) \cos^{2k}(d\pi/L)$ неотрицателен как свертка двух неотрицательных функций, следовательно, функция положительно полуопределена, следовательно, может являться ковариационной, что и требовалось доказать.

Сама по себе оптимизационная задача, которая решается при определении коэффициентов в модели кригинга, выглядит следующим образом [1, 3–5]. Кригинг использует линейную комбинацию известных значений для оценки значения функции в неизвестных точках: $\hat{Z}(t) = \sum_{i=1}^N \lambda_i Z(t_i)$, где коэффициенты подобраны таким образом, что $\hat{Z}(t)$ – несмещенная оценка для $Z(t)$, дисперсия ошибки $\varepsilon(t)$ минимальна.

Учитывая несмещенность оценки, следующее равенство верно:

$$E[\hat{Z}(t)] = \sum_{i=1}^N \lambda_i E[Z(t_i)]$$

Откуда, в силу линейности математического ожидания и $E[\varepsilon(t)] = 0$ получим

$$\mu(t) = \sum_{i=1}^N \lambda_i \mu(t_i)$$

Далее, предположим, что функция $\mu(t)$ может быть представлена в виде линейной комбинации конечного числа некоторых линейно независимых базисных функций f_l . В силу существования большого количества способов (например, разложением в ряд Фурье или аппроксимацией сплайнами) аппроксимации функции с заданной точностью конечной линейной комбинацией базисных функций, это предположение верно для любой функции. Тогда

$$\mu(t) = \sum_{l=1}^M a_l f_l(t)$$

Тогда из несмещенности оценки получаем ограничения:

$$f_l(t) = \sum_{i=1}^N \lambda_i f_l(t_i) \quad l = 1, \dots, M$$

Следовательно, минимизация дисперсии ошибки сводится к решению оптимизационной задачи с ограничениями типа "равенство":

$$\left\{ \begin{array}{l} D[\hat{Z}(t) - Z(t)] = \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \text{cov}[Z(t_i), Z(t_j)] - \\ \quad - 2 \sum_{i=1}^N \lambda_i \text{cov}[Z(t_i), Z(t)] + \\ \quad + D[Z(t)] \rightarrow \min \\ f_1 = \sum_{i=1}^N \lambda_i f_1(t_i) \\ \dots \\ f_M = \sum_{i=1}^N \lambda_i f_M(t_i) \end{array} \right.$$

При введении множителей Лагранжа u_l , $l = 1, \dots, M$, задача преобразуется в задачу оптимизации без ограничений:

$$\begin{aligned} & \sum_{i=1}^N \sum_{j=1}^N \lambda_i \lambda_j \text{cov}[Z(t_i), Z(t_j)] - \\ & - 2 \sum_{i=1}^N \lambda_i \text{cov}[Z(t_i), Z(t)] + \\ & + D[Z(t)] + 2 \sum_{l=1}^M u_l \left(-f_l(t) + \sum_{i=1}^N \lambda_i f_l(t_i) \right) \rightarrow \min \end{aligned} \quad (1)$$

Продифференцируем полученное выражение по переменным λ_i и u_l и получим систему из $N+M$ уравнений для $N+M$ переменных, которую в случае линейной независимости базисных функций f_i можно решить относительно параметров λ и u :

$$\left\{ \begin{array}{l} \sum_{i=1}^N \lambda_i f_1(t_i) = f_1(t) \\ \dots \\ \sum_{i=1}^N \lambda_i f_M(t_i) = f_M(t) \\ \sum_{l=1}^M u_l f_l(t_1) + \sum_{j=1}^N \lambda_j \text{cov}[Z(t_1), Z(t_j)] = \text{cov}[Z(t_1), Z(t)] \\ \dots \\ \sum_{l=1}^M u_l f_l(t_N) + \sum_{j=1}^N \lambda_j \text{cov}[Z(t_N), Z(t_j)] = \text{cov}[Z(t_N), Z(t)] \end{array} \right.$$

После решения этой системы уравнений, получим веса λ_i , благодаря которым можно выразить предсказание модели для точки с неизвестным значением.

Модель зависит от трех составляющих: базисных функций, параметра α и параметра k . Рассмотрим подбор каждой из них.

Подбор базисных функций – это нетривиальная задача, которую трудно автоматизировать. В общем случае, для каждого нового набора данных требуется подобрать свой набор базисных функций. Есть методы, позволяющие использовать некоторый набор функций для любых входных данных, в частности, использование степенных функций. Эта модель носит название Тейлор Кригинга [8], так как процесс аналогичен разложению некоторой функции в ряд Тейлора. Однако, несмотря на универсальность, это не всегда оптимальный выбор, так как разложение в ряд Тейлора имеет ограниченный интервал сходимости, что может оказаться существенным при прогнозировании на удалении от первоначальной выборки.

Остальные два параметра подобрать проще. Для подбора параметра k нетрудно заметить, что последовательность таких функций при неизменном параметре α сходятся в пространстве $C[a, b]$, следовательно, после некоторого значения k изменения, вызываемые изменением функции, будут иметь малое влияние на значения предсказаний. Для подбора параметра α нужно решить задачу минимизации квадрата ошибки. Таким образом, для нахождения оптимальных параметров можно перебрать значения k , для каждого из них выбрать оптимальный параметр α и выбрать среди этих пар ту, которая дает минимальное значение ошибки.

Приведём решение задачи 1 согласно универсальному кригингу с использованием множителей Лагранжа m_i (вообще единица в качестве базисной функции не обязательна, но применяется очень часто):

$$\begin{pmatrix} \lambda_1 \\ \dots \\ \lambda_n \\ m_0 \\ m_1 \\ \dots \\ m_p \end{pmatrix} = \begin{pmatrix} \gamma(t_1 - t_1) & \dots & \gamma(t_1 - t_n) & 1 & f_1(t_1) & \dots & f_p(t_1) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \gamma(t_n - t_1) & \dots & \gamma(t_n - t_n) & 1 & f_1(t_n) & \dots & f_p(t_n) \\ 1 & \dots & 1 & 0 & 0 & \dots & 0 \\ f_1(t_1) & \dots & f_1(t_n) & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ f_p(t_1) & \dots & f_p(t_n) & 0 & 0 & \dots & 0 \end{pmatrix}^{-1} \begin{pmatrix} \gamma(t - t_1) \\ \dots \\ \gamma(t - t_n) \\ 1 \\ f_1(t_0) \\ \dots \\ f_p(t_0) \end{pmatrix} \quad (2)$$

Здесь функция $\gamma(\Delta t)$ – это дисперсия разницы значений функции в точках, расположенных на расстоянии Δt друг от друга. На практике удобно строить аналитическое приближение к вариограмме по известным значениям. Примеры применяемых функций эмпирических вариограмм [17]:

– Экспоненциальная модель:

$$\gamma(h) = (s - n)(1 - e^{-\frac{h}{ra}}) + n$$

– Гауссова модель:

$$\gamma(h) = (s - n)(1 - e^{-\frac{h^2}{r^2 a}}) + n$$

где s, n, r, a – параметры моделей.

Что касается вычислительной сложности задачи построения модели кригинга, обращаемая матрица 2 диагональной быть не может, поэтому трудно описать специальные случаи, когда с вычислительной точки зрения кригинг будет очень быстр [18]. С другой стороны, матрица системы зависит только от известных параметров (значений в точках и ковариационной функции), от неизвестной точки зависит только правая часть. Следовательно, можно один раз обратить матрицу (или, например, сделать LU-разложение) и после этого получать интерполяции за $O((N + M)^2)$ операций. Кроме того, если разброс известных точек довольно большой, а ковариационная функция быстро убывает и стремится к нулю (например, как экспонента), то матрицу системы можно считать разреженной, что может быть полезным при больших объёмах данных. Пока трудно определить, в каких случаях кригинг должен работать быстрее стандартных моделей или нейросетевых моделей прогнозирования, этому будут посвящены дальнейшие исследования.

Комбинирование моделей – сложная задача, требующая выбора подхода комбинирования. Он зависит от данных, их объема и зависимостей в них. В данной работе будет рассмотрено два основных подхода к комбинированию: взвешенные суммы предсказаний и гибридные подходы.

Взвешенные суммы предсказаний в общем случае работают так: каждая из комбинируемых моделей прогнозирует значение в определенной точке и результатом прогноза будет, очевидно,

взвешенная сумма полученных предсказаний. Основная задача таких моделей – определить вес предсказания каждой модели для определенной точки. Основные подходы следующие:

- Простое усреднение: у каждого предсказания одинаковый вес
- Постоянные веса: у каждого предсказания есть постоянный вес, определенный один раз при кросс-валидации модели
- Динамические веса: поддерживается скользящее окно для оценки ошибок каждой из моделей, на основании чего для каждого предсказания выбирается свой вес предсказаний моделей
- Мета-моделирование: есть некоторая модель машинного обучения, предназначенная для выбора оптимальных значений весов

Каждый из этих подходов реализуем с любыми моделями. К примеру, для комбинированной модели, основанной на трех статистических моделях, получим предсказания вида

$$\text{prediction}[x] = \alpha \cdot \text{ARIMA}[x] + \beta \cdot \text{Exp.Smoothering}[x] + \gamma \cdot \text{Kriging}[x]$$

Гибридный подход к комбинированию моделей заключается в том, что прогноз одной модели улучшается за счет другой модели, обученной на остатках первой. Этот подход используется в случаях, когда каждая из моделей сильнее в той или иной области прогнозов. К примеру, как мы убедимся далее, модель на основе кригинга хорошо умеет прогнозировать тренды, но несколько хуже – сезонность. Возьмем ее за основу. Обучим на одной части тренировочных данных и спрогнозируем значения для второй части. Возьмем вторую модель, к примеру, Хольта-Винтерса, и обучим ее на полученных остатках. За счет чего мы надеемся компенсировать недостатки работы модели Кригинга на сезонных данных, вторая модель может найти зависимости в остатках, которые первая модель упустила. Комбинировать таким способом можно любые модели в любом порядке и, в принципе, в любом количестве (третью модель можно обучить на ошибках второй и первой и так далее).

Выбор подхода во многом зависит от того, какие данные обрабатываются. Первый шаг выбора – оценка размера данных: гибридный подход требует больше исходных данных, чем взвешенные данные ($\sum n_i < \max(n_i)$, где n_i – минимальное необходимое количество данных для нормальной работы модели). Кроме того, гибридные подходы обычно работают медленнее за счет необходимости последовательного построения моделей (подход взвешенных сумм можно сделать параллельным: пусть каждая модель обучается и вычисляет прогнозы в своем потоке). Но в случае правильного подбора моделей, гибридный подход обычно выигрывает у взвешенных сумм по точности.

2. ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ

Рассмотрен набор пассажиропотока авиакомпаний [20] с монотонным трендом, сильной сезонностью и небольшим шумом. Длина тренировочной выборки – 100, длина тестовой выборки – 40, период сезонности: 12. В эксперименте под комбинированной моделью будет подразумеваться взвешенная сумма предсказаний трёх моделей с постоянными весами. Выбор именно этого подхода к комбинированию моделей был сделан на основании того, что представленные наборы входных данных, в основном, небольшие, в связи с чем гибриднему подходу может не хватить данных для полного обучения всех моделей. Наборы данных не имеют сложных нелинейных зависимостей, поэтому, выбор модели с динамическими весами или с мета-моделью был бы нерациональным. Построим прогнозы рассматриваемых моделей на одном графике. Обозначим тренировочные данные черной сплошной кривой, тестовые данные – черной пунктирной кривой, предсказания модели Кригинга – красной сплошной кривой, модели ARIMA

– зеленой сплошной кривой, модели Хольта-Винтерса – синей сплошной кривой, комбинированной модели – пурпурной сплошной кривой 1. В таблице 1 приведено сравнение прогнозов моделей по метрике RMSE [2, 13] с точностью до 0.01.

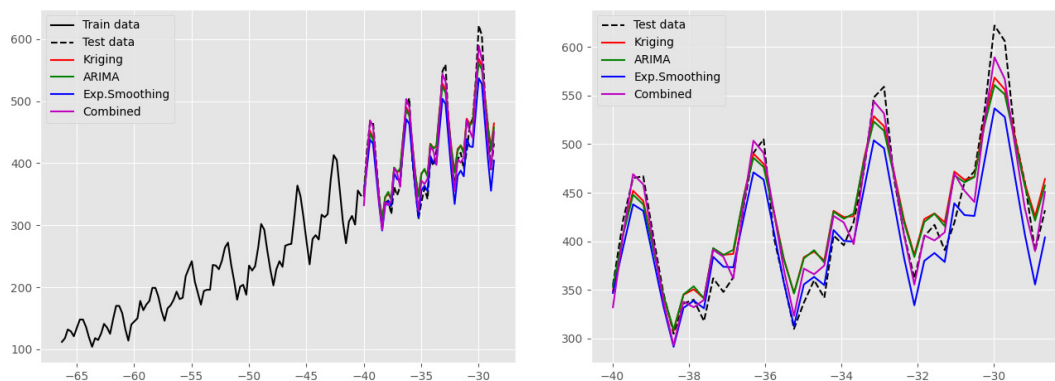


Рис. 1. Графики обучающей выборки и прогнозов моделей на тестовый период для ряда пассажиропотока (слева) и сравнение прогнозов (справа)

Таблица 1. Сравнение точности предсказаний на первых k прогнозах для ряда пассажиропотока

Модель, k	10	20	30	40
ARIMA	15.39	23.51	24.28	26.72
Хольт-Винтерс	19.79	19.78	23.08	30.85
Кригинг	13.99	22.91	23.65	25.92
Комбинация	12.78	17.90	18.10	20.50

По графикам и таблице можно заметить, что кригинг и модель ARIMA дают примерно одинаковые предсказания на всем интервале предсказания. Экспоненциальное сглаживание показывает примерно такие же результаты в течение первого сезона предсказания. После этого модель начинает давать значительно отклоняющиеся от действительных значения, поэтому на всем интервале прогнозирования ее метрики будут несколько хуже, чем у двух других моделей. Комбинированная модель, в свою очередь, ожидаемо дает увеличение точности прогноза: у нее получается лучше остальных моделей следовать за локальными экстремумами, ошибки в которых обычно сильно влияют на итоговую точность моделей.

Следующий набор данных – ежемесячный ряд средней температуры воздуха [21] – с отсутствующим трендом, сильной сезонностью и небольшим шумом. Длина тренировочной выборки – 800, длина тестовой выборки – 100, период сезонности: 12. По графикам 2 и таблице 2 можно отметить, что все модели с достаточно большой точностью описывают тестовые данные.

Таблица 2. Сравнение точности предсказаний на первых k прогнозах для ряда средней температуры

k	20	40	60	80	100
ARIMA	1.68	1.51	1.33	1.38	1.39
Хольт-Винтерс	1.42	1.30	1.26	1.33	1.32
Кригинг	1.49	1.48	1.40	1.39	1.38
Комбинация	1.22	1.23	1.16	1.16	1.19

Также по набору данных пассажиропотока решалась и другая задача. В наборе 144 наблюдений с января 1949 года по декабрь 1960 года. Небольшой непрерывный участок данных был удалён. Различные модели должны были сделать прогноз на тестовый период, который расположен не в конце, а между заданными точками [19]. На рисунке 3 представлен исходный ряд

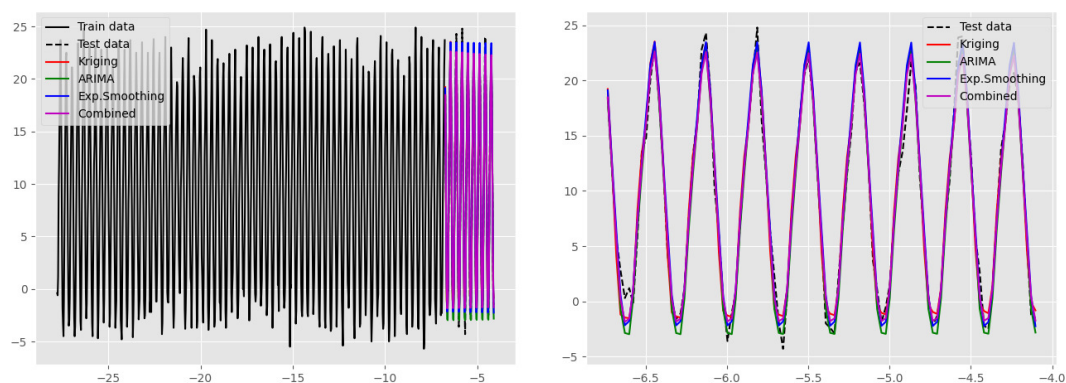


Рис. 2. Графики обучающей выборки и прогнозов моделей на тестовый период для ряда средней температуры (слева) и сравнение прогнозов (справа)

после удаления промежутка и результаты прогнозирования с помощью сплайнов, кригинга, метода STL [2]. Сравнение прогнозов с реальными данными приведено в таблице 3. Заметим, что модель на основе кригинга не просто сделала прогноз, близкий к реальности, но и уловила сезонность в данных.

В следующем эксперименте удалённые промежутки разбросаны по данным ряда. На рисунке 4 представлен исходный ряд после удаления промежутка и результаты прогнозирования с помощью сплайнов, кригинга, метода STL [2]. Сравнение прогнозов с реальными данными приведено в таблице 3.

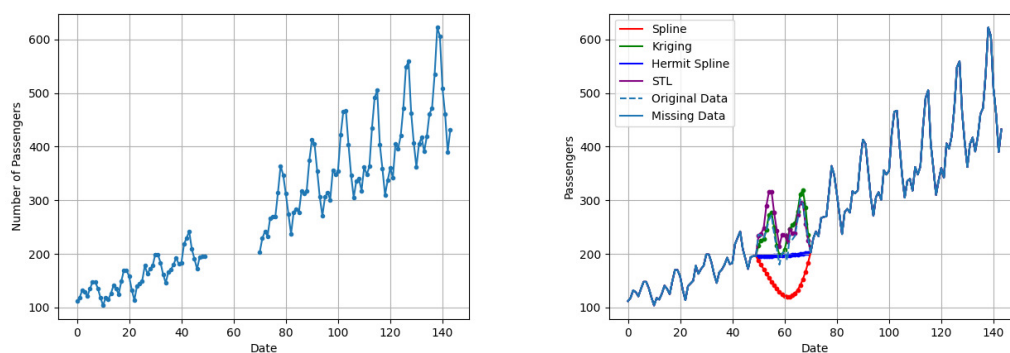


Рис. 3. Графики ряда с удалённым промежутком значений (слева) и прогнозов моделей на удалённый участок (справа)

Таблица 3. Оценка кригинга и других методов при заполнении удалённого промежутка.

Модель	RMSE (случ. пробелы)	RMSE (большой пробел)
Кригинг (эксп.)	16.39	16.90
Кригинг (гаус.)	15.86	17.21
Кригинг (нейр.)	15.86	16.78
Пол. сплайн 3-го пор.	17.68	96.25
Эрмитов сплайн	17.75	49.85
STL	16.89	27.39

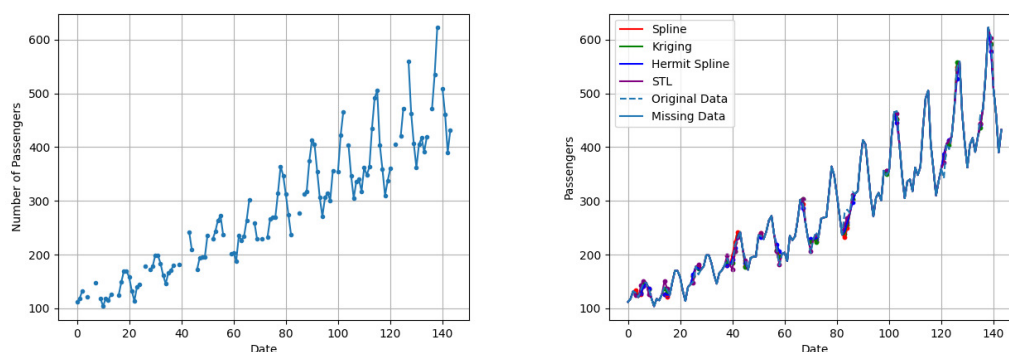


Рис. 4. Графики ряда с удалёнными значениями (слева) и прогнозов моделей на удалённые участки (справа)

3. ЗАКЛЮЧЕНИЕ

Модель кригинга, построенная в работе, дает более точные прогнозы, чем популярные модели ARIMA и Хольта-Винтерса, особенно когда данные имеют сложный тренд. Если же тренд в данных простой или его нет, кригинг показывает результаты, близкие к стандартным моделям временных рядов. При этом, [6, 7] показано, что кригинг очень хорошо работает в случаях сильных шумов. Кригинг может быть введен в комбинированные модели для прогнозирования временных рядов. В гибридной модели, построенной в работе, он дополняет традиционные статистические методы, такие как ARIMA и экспоненциальное сглаживание, за счет учета сложных зависимостей, особенно в задачах со сложными компонентами тренда. Использование кригинга в гибридных моделях позволяет повысить точность прогнозов и адаптироваться к нелинейным трендам. Модель универсального кригинга хорошо справляется как с прогнозированием рядов на будущий промежуток, так и с обработкой удалённых значений. К недостаткам метода следует отнести то, что работа происходит только для данных с сезонностью, а также требует ручного подбора базисных функций, так как алгоритма для их автоматического выбора пока нет. Дальнейшие исследования могут быть направлены на преодоление этих ограничений и расширение применимости кригинга в задачах прогнозирования временных процессов.

СПИСОК ЛИТЕРАТУРЫ

1. Xie X., Lu X., Theta-regularized Kriging: Modeling and algorithms. *Applied Mathematical Modelling*, 2024, vol. 136, p. 115627. doi: 10.1016/j.apm.2024.07.034.
2. Hyndman R.J., Athanasopoulos G., *Forecasting: principles and practice*. Melbourne: OTexts, 2021 (3rd edition)
3. Krige D.G., A statistical approach to some basic mine valuation problems on the Witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, 1951, vol. 52, no. 6, p. 119–139.
4. Matheron G., *Principles of Geostatistics*. *Economic Geology*, 1963, vol. 58, no. 8, p. 1246–1266.
5. Matheron G., *Traite de geostatistique appliquee*. Paris: Editions Technip, 1962, 334 p.
6. Gaidos A., Hancova M., Hanc J., *Kriging Methodology and Its Development in Forecasting Econometric Time Series*. *Statistika*, 2017, vol. 97, no. 1, p. 59–73.
7. Hancova M., Gaidos A., Hanc J., *Estimating variances in time series kriging using convex optimization and empirical BLUPs*. *Stat Papers*, 2021, vol. 62, p. 1899–1938.

8. Liu H., Shi J., Erdem E., Prediction of wind speed time series using modified Taylor Kriging method. *Energy*, 2010, vol. 35, no. 12, p. 4870–4879. doi: 10.1016/j.energy.2010.09.001.
9. Yuan Y., Au F.T.K., Yang D., Zhang J., An active learning Kriging-based Bayesian framework for probabilistic structural model exploration. *Journal of Sound and Vibration*, 2025, vol. 596, p. 118730. doi: 10.1016/j.jsv.2024.118730.
10. Li Z.-A., Dong X.-W., Zhu C.-Y., Chen C.-H., Zhang H., Vectorial surrogate modeling method based on moving Kriging model for system reliability analysis. *Computer Methods in Applied Mechanics and Engineering*, 2024, vol. 432B, p. 117409. doi: 10.1016/j.cma.2024.117409.
11. Hu M., Li H., Lu J., Zou H., Ma Q., An online modeling virtual sensing technique based on kriging interpolation for active noise control. *Mechanical Systems and Signal Processing*, 2025, vol. 224, p. 112186. doi: 10.1016/j.ymssp.2024.112186.
12. Мохнаткина У.С., Парфенов Д.В., Петрусевич Д.А., Анализ подходов к определению тренда в структуре временного ряда. *Российский Технологический Журнал*, 2024, 12, №3, 2024, с. 93–103. doi: 10.32362/2500-316X-2024-12-3-93-103.
13. Stock J.H., Watson M.W., *Introduction to Econometrics*. Pearson, 2019
14. Chen C., Gao X., Yu G., Xu W., Naing W., Time history response prediction of stochastic SDOF structures through Kriging-NARX modeling and adaptive real-time hybrid simulation. *Structures*, 2024, vol. 64, p. 106674. doi: 10.1016/j.istruc.2024.106674.
15. Zhan H., Liu H., Xiao N.-C., Time-dependent reliability analysis of structural systems based on parallel active learning Kriging model. *Expert Systems with Application*, 2024, vol. 247, p. 123252. doi: 10.1016/j.eswa.2024.123252.
16. Vinuales A., Montes F., Guijarro M., Gomez C., de la Calle I., Madrigal J., Real-time assessment of live forest fuel moisture content and flammability by using space-time universal kriging. *Ecological Modelling*, 2024, vol. 498, p. 110867. doi: 10.1016/j.ecolmodel.2024.110867.
17. Li Y., Baorong Z., Xiaohong X., Zijun L., Application of a semivariogram based on a deep neural network to Ordinary Kriging interpolation of elevation data. *PLOS ONE*, 2022, vol. 17, no. 4, p. 1–12. doi: 10.1371/journal.pone.0266942.
18. den Hertog D., Kleijnen J.P.C., Siem A.Y.D., The Correct Kriging Variance Estimated by Bootstrapping. *CentER Discussion Paper No. 2004-46*, 2004. doi: 10.2139/ssrn.557862.
19. Drozdov I.Yu., Aleksakhin A.V., Aleksakhina Yu.V., Petrusevich D.A., Mathematical models of water pollution evaluation. *IOP Conference Series: Earth and Environmental Science*, 2021, vol. 684, p. 012026. doi: 10.1088/1755-1315/684/1/012026.
20. Air Passenger Data for Time Series Analysis. URL: <https://www.kaggle.com/datasets/ashfakyeafi/air-passenger-data-for-time-series-analysis/data>, 17-06-2025
21. Temperature trends around Aomori City. URL: <https://www.kaggle.com/code/akioonodera/temperature-trends-around-aomori-city>, 17-06-2025

Implementation of kriging in time series forecasting

N. V. Beletskaya, S. A. Konanov, D. A. Petrushevich, A. V. Stoyanovsky

This paper contains comparison of kriging technique implementation for time series modeling and forecasting. In the computational experiment kriging is compared to the standard methods (ARIMA/ETS). Basically, the concept of kriging was built in the task of multivariate function interpolation in the study of spatial distribution modeling of minerals and (when time is included in the models) geological processes. If one abstracts from the geological nature of the data, this approach can also be applied to modeling of an arbitrary multivariate quantity (constant or time-varying). The interest in this method is also related to the fact that the adjustable model belongs to a very simple class – linear or polynomial regression. In the presented research forecasting for some time interval in the future is made and also processing of missing values is investigated (concentrated in one place or scattered over the data series). Combined models are built (the prediction is given as weighted combination of forecasts of several models). Each successive model is trained on the residuals of the previous models (it's close to bagging technique). The combination of ARIMA, ETS and kriging models is studied.

KEYWORDS: kriging, time series, ARIMA, ETS.