

УДК 556.55:519.23(575.2)
AGRIS U10

<https://doi.org/10.33619/2414-2948/115/05>

**СРАВНИТЕЛЬНЫЙ АНАЛИЗ РЕГРЕССИОННЫХ МОДЕЛЕЙ
ДЛЯ АППРОКСИМАЦИИ И ПРОГНОЗИРОВАНИЯ ЕЖЕГОДНЫХ
МАКСИМАЛЬНЫХ ОБЪЕМОВ ВОДЫ
ТОКТОГУЛЬСКОГО ВОДОХРАНИЛИЩА КЫРГЫЗСКОЙ РЕСПУБЛИКИ**

©*Ташполотов Ы.* ORCID: 0000-0001-9293-7885, SPIN-код: 2425-6716, д-р физ.-мат. наук,
Ошский государственный университет, г. Ош, Кыргызстан, itashpolotov@mail.ru
©*Адылова Э. С.*, ORCID: 0000-0002-1886-6932, SPIN-код: 6609-9843, Кыргызско-узбекский
международный университет им. Б. Сыдыкова, г. Ош, Кыргызстан, A_elmira01@mail.ru

**COMPARATIVE ANALYSIS OF REGRESSION MODELS
FOR APPROXIMATING AND FORECASTING ANNUAL MAXIMUM WATER
VOLUMES OF THE TOKTOGUL RESERVOIR IN THE KYRGYZ REPUBLIC**

©*Tashpolotov Y.*, ORCID: 0000-0001-9293-7885, SPIN-code: 2425-6716, Dr. habil., Osh State
University, Osh, Kyrgyzstan, itashpolotov@mail.ru
©*Adylova E.*, ORCID: 0000-0002-1886-6932, SPIN-code: 6609-9843, Batyraly Sydykov Kyrgyz-
Uzbek International University, Osh, Kyrgyzstan, A_elmira01@mail.ru

Аннотация. Представлен сравнительный анализ различных моделей регрессии, включая линейную, полиномиальную, экспоненциальную, синусоидальную и сплайновую регрессию, для аппроксимации ежегодных максимальных объемов воды Токтогульского водохранилища Кыргызской Республики. Целью исследования является определение наиболее адекватной модели для описания динамики наполнения водохранилища, что имеет важное значение для планирования водохозяйственной деятельности и прогнозирования водных ресурсов в регионе. Проведен сравнительный анализ прогнозных возможностей моделей на основе имеющихся данных. Результаты демонстрируют различную степень соответствия моделей наблюдаемым данным, что подчеркивает необходимость тщательного выбора метода регрессионного анализа в гидрологических исследованиях. Анализ показал, что простые модели, такие как линейная, оказались недостаточными, в то время как сплайн-регрессия продемонстрировала значительно лучшие результаты.

Abstract. This work presents a comparative analysis of various regression models, including linear, polynomial, exponential, sinusoidal, and spline regression, for approximating the annual maximum water volumes of the Toktogul Reservoir in the Kyrgyz Republic. The aim of the study is to determine the most adequate model for describing the dynamics of reservoir filling, which is crucial for planning water management activities and forecasting water resources in the region. A comparative analysis of the predictive capabilities of the models was conducted based on the available data. The results demonstrate varying degrees of model fit to the observed data, highlighting the necessity of careful selection of regression analysis methods in hydrological studies. The analysis showed that simple models, such as linear regression, were insufficient, while spline regression demonstrated significantly better results.

Ключевые слова: регрессионный анализ, аппроксимация данных, Токтогульское водохранилище, линейная регрессия, полиномиальная регрессия, экспоненциальная регрессия, синусоидальная регрессия, сплайн-регрессия, прогнозирование.

Keywords: regression analysis, data approximation, Toktogul Reservoir, linear regression, polynomial regression, exponential regression, sinusoidal regression, spline regression, forecasting.

Токтогульское водохранилище, расположенное на реке Нарын в Кыргызской Республике, является крупнейшим водоемом страны и играет ключевую роль в обеспечении электроэнергией и водой для орошения обширных территорий Центральной Азии. Эффективное управление водными ресурсами водохранилища требует точного понимания динамики его наполнения, в частности, ежегодных максимальных объемов воды [1, 2].

Регрессионный анализ представляет собой мощный статистический инструмент для моделирования зависимостей между переменными и может быть использован для аппроксимации и прогнозирования гидрологических временных рядов. Выбор адекватной модели регрессии является критически важным для получения достоверных результатов [3, 4].

Целью данной работы является сравнительный анализ различных типов регрессионных моделей для аппроксимации ежегодных максимальных объемов воды Токтогульского водохранилища. В рамках исследования будут рассмотрены как простые (линейная), так и более сложные (полиномиальная, экспоненциальная, синусоидальная, сплайновая) модели регрессии, и проведена оценка их способности описывать наблюдаемые данные.

Материалы и методы

В качестве исходных данных использовались ежегодные максимальные объемы воды в Токтогульском водохранилище КР с 2012 по 2024 годы (Таблица 1).

Таблица 1

МАКСИМАЛЬНЫЕ ОБЪЕМЫ ВОДЫ В ТОКТОГУЛЬСКОМ ВОДОХРАНИЛИЩЕ
КЫРГЫЗСКОЙ РЕСПУБЛИКИ с 2012 по 2024 годы (млрд. м³)

2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
13,645	10,54	9,975	13,18	17,8	17,266	16,151	13,37	10,852	11,069	10,852	10,072	13,026

Для аппроксимации данных были применены следующие модели регрессии [3–9]:

1. *Линейная регрессия (эквивалентная простейшей нейронной сети).* В данном случае предполагается линейную зависимость между временем (годом) и объемом воды в водохранилище. Уравнение будет иметь вид:

$$y = ax + b, \quad (1)$$

где: a — вес (коэффициент наклона); b — смещение (свободный член).

Использование такой сети на основе данных Таблицы 1 будет заключаться в поиске оптимальных значений a и b направленные на минимизацию ошибки между предсказанными и фактическими значениями y . Для данных, представленные в Таблице 1 на основе метода наименьших квадратов получим следующее линейное регрессионное уравнение:

$$y = 0.157x + 13.18. \quad (2)$$

2. *Полиномиальная регрессия (степень 2).* Полиномиальная регрессия моделирует квадратичную зависимость, позволяя описывать криволинейное поведение:

$$y = ax^2 + bx + c. \quad (3)$$

Для уравнения (3) коэффициенты a , b , c определим исходя из данных таблицы 1, тогда получим следующее уравнение:

$$y = 0.05 x^2 - 1.2 x + 15.5. \quad (4)$$

3. *Экспоненциальная регрессия.* Экспоненциальная регрессия используется для моделирования экспоненциального роста или спада и выражается в виде формулы:

$$y = a * \exp(bx). \quad (5)$$

Для применения такой модели часто применяют логарифмическое преобразование y , чтобы привести уравнение к линейному виду:

$$\ln(y) = \ln(a) + bx. \quad (6)$$

Затем можно использовать линейную регрессию для определения $\ln(a)$ и b , а затем вернуться к исходным параметрам a и b . На основе такого подхода, используя данные таблицы 1 получили следующее уравнение:

$$y=14.028 \exp(-0.016x). \quad (7)$$

4. *Синусоидальная регрессия.* Если в данных наблюдается периодичность или волнообразное поведение, можно использовать тригонометрические функции, например синус:

$$y = a \sin(b x + c) + d. \quad (8)$$

где a — амплитуда, b — частота, c — фаза, d — вертикальное смещение.

Использование такой модели может быть более сложным и часто требует применения нелинейных методов оптимизации.

$$y=4,932\sin(0.51x - 1.045) + 12.031 \quad (9)$$

5. *Сплайн –регрессия* [7]. Регрессионные модели, в которых функции меняется в одной или нескольких точках на шкале значений предикторов, называется сплайнами, либо кусочными полиномами, а сами точки функции называются узловыми точками (узлами).

Сплайн-регрессия не параметрический метод, позволяющий гибко моделировать нелинейные зависимости путем построения кусочных полиномиальных функций. В данном исследовании использовали кубический сплайн с определенным количеством узлов: параметры сплайна определялись эмпирически [8] на основе анализа данных, представленные в Таблице 1.

Результаты и обсуждение

Результаты применения различных моделей регрессии к данным ежегодных максимальных объемов воды Токтогульского водохранилища представлены в табличной форме (Таблица 2).

Оценка адекватности каждой модели проводилась на основе сравнения прогнозных значений с фактическими данными, а также (опционально) с использованием статистических метрик, таких как средняя абсолютная ошибка, среднеквадратическая ошибка и коэффициент детерминации (R^2). Анализ результатов, представленных в Таблице 2 показал, что простая линейная регрессия не способна адекватно описать наблюдаемую динамику объемов воды, демонстрируя значительные отклонения от фактических значений. Полиномиальная регрессия второй степени показала некоторое улучшение, улавливая криволинейной зависимости, однако также имела существенные расхождения в отдельных

точках. Экспоненциальная и синусоидальная регрессии в данном случае оказались наименее подходящими, не отражая характерных особенностей временного ряда.

Таблица 2

СРАВНЕНИЕ ФАКТИЧЕСКИХ И ПРОГНОЗНЫХ ЗНАЧЕНИЙ ЕЖЕГОДНЫХ МАКСИМАЛЬНЫХ ОБЪЕМОВ ВОДЫ ТОКТОГУЛЬСКОГО ВОДОХРАНИЛИЩА КЫРГЫЗСТАНА (млрд. м³)

Годы (X)	Фактический объем (Y)	Прогноз (Линейная)	Прогноз (Полиномиальная)	Прогноз (Экспоненциальная)	Прогноз (Синусоидальная)	Прогноз (Сплайн)
2012	13.00	13.337	14.350	14.028	13.899	13.579
2013	11.25	13.494	13.600	13.809	11.978	11.688
2014	10.10	13.651	12.950	13.595	10.416	10.284
2015	13.00	13.808	12.400	13.386	9.351	13.080
2016	17.50	13.965	11.950	13.182	8.821	17.479
2017	17.20	14.122	11.600	12.983	8.867	17.056
2018	16.10	14.279	11.350	12.789	9.533	15.275
2019	13.40	14.436	11.200	12.600	10.860	13.370
2020	11.00	14.593	11.150	12.415	12.881	11.786
2021	11.10	14.750	11.200	12.235	15.628	11.069
2022	10.90	14.907	11.350	12.059	19.133	11.664
2023	10.00	15.064	11.600	11.888	23.429	11.915
2024	13.00	15.221	11.950	11.721	28.548	11.267
2025	----	15.378	12.400	11.559	34.523	13.63
2026	----	15.535	12.950	11.401	41.387	12.80

Сплайн-регрессия продемонстрировала наибольшую гибкость и наилучшее соответствие фактическим данным, плавно следуя их изменениям. Это подчеркивает эффективность непараметрических методов для моделирования сложных нелинейных зависимостей в гидрологических процессах.

Выводы

1. Сравнительный анализ различных моделей регрессии показал, что выбор метода имеет решающее значение для адекватной аппроксимации ежегодных максимальных объемов воды в водохранилище. Простые модели, такие как линейная, оказались недостаточными, в то время как более гибкие подходы, например, сплайн-регрессия, продемонстрировали значительно лучшие результаты.

2. Полученные результаты подчеркивают важность учета нелинейности и потенциальной сложности гидрологических процессов при их моделировании.

Дальнейшие исследования будут направлены на оптимизацию параметров сплайн-регрессии (выбор оптимального количества и расположения узлов), а также на исследование других нелинейных методов и моделей временных рядов для повышения точности прогнозирования водных ресурсов КР.

Список литературы:

1. Михайлов В. Н., Добролюбов С. А. Гидрология. М.; Берлин: Директ-Медиа, 2017. 752 с.

2. Ткачев А. А., Ольгаренко И. В. Современные проблемы в управлении водораспределением в магистральных каналах оросительных систем // Научный журнал Российского НИИ проблем мелиорации. 2021. Т. 11, № 2. С. 1–23. <https://doi.org/10.31774/2222-1816-2021-11-2-1-23>

3. Дрейпер Н. Р., Смит Г. Прикладной регрессионный анализ. М., 2007. 912 с.

4. Chatterjee S., Hadi A. S. Regression analysis by example. John Wiley & Sons, 2015.

5. Rencher A. C., Schaalje G. B. Linear models in statistics. John Wiley & Sons, 2008.

6. Ryan T. P. Modern regression methods. John Wiley & Sons, 2008.

7. Saber G. A. F., Wild C. J. Nonlinear regression. 1989.

8. Вапник В. Н. Восстановление зависимостей по эмпирическим данным. М.: Наука, 1979. 448с.

9. De Boor C., De Boor C. A practical guide to splines. New York : springer, 1978. V. 27. P. 325.

References:

1. Mikhailov, V. N., & Dobrolyubov, S. A. (2017). *Gidrologiya*. Moscow. (in Russian).

2. Tkachev, A. A., & Ol'garenko, I. V. (2021). *Sovremennye problemy v upravlenii vodoraspredeleniem v magistral'nykh kanalah orositel'nykh sistem. Melioratsiya i gidrotekhnika*, 11(2), 1-23. (in Russian). <https://doi.org/10.31774/2222-1816-2021-11-2-1-23>

3. Dreiper, N. R., & Smit, G. (2007). *Prikladnoi regressionnyi analiz*. Moscow. (in Russian).

4. Chatterjee, S., & Hadi, A. S. (2015). *Regression analysis by example*. John Wiley & Sons.

5. Rencher, A. C., & Schaalje, G. B. (2008). *Linear models in statistics*. John Wiley & Sons.

6. Ryan, T. P. (2008). *Modern regression methods*. John Wiley & Sons.

7. Saber, G. A. F., & Wild, C. J. (1989). *Nonlinear regression*.

8. Vapnik, V. N. (1979). *Vosstanovlenie zavisimostei po empiricheskim dannym*. Moscow. (in Russian).

9. De Boor, C., & De Boor, C. (1978). *A practical guide to splines* (Vol. 27, p. 325). New York: springer.

Работа поступила
в редакцию 12.04.2025 г.

Принята к публикации
25.04.2025 г.

Ссылка для цитирования:

Ташполотов Ы. Адылова Э. С. Сравнительный анализ регрессионных моделей для аппроксимации и прогнозирования ежегодных максимальных объемов воды Токтогульского водохранилища Кыргызской Республики // Бюллетень науки и практики. 2025. Т. 11. №6. С. 36-40. <https://doi.org/10.33619/2414-2948/115/05>

Cite as (APA):

Tashpolotov, Y., & Adylova, E. (2025). Comparative Analysis of Regression Models for Approximating and Forecasting Annual Maximum Water Volumes of the Toktogul Reservoir in the Kyrgyz Republic. *Bulletin of Science and Practice*, 11(6), 36-40. (in Russian). <https://doi.org/10.33619/2414-2948/115/05>