ТЕХНИКАЛЫҚ ҒЫЛЫМДАР
ТЕХНИЧЕСКИЕ НАУКИ
TECHNICAL SCIENCE

CENTRAL ASIAN SCIENTIFIC JOURNAL
ISSN 3135-3061
2026. №4 (32), Том 3. С. 59–65
DOI: https://doi.org/10.5281/zenodo.24622079
Сайт журнала: https://cajournal.kz

УДК 519.246.8

НАСКОЛЬКО НАДЁЖНЫ ИНТЕРВАЛЫ ПРОГНОЗА? КАЛИБРОВКА И СМЕЩЕНИЕ МЕСЯЧНЫХ ПРОГНОЗОВ ВЫРАБОТКИ ЭЛЕКТРОЭНЕРГИИ В КАЗАХСТАНЕ

И.С. Толеубаева
Евразийский национальный университет им. Л.Н. Гумилёва, Астана, Казахстан

Толеубаева Илялим Сериковна — магистрант 2 курса, образовательная программа 7М05402 – Математика, Евразийский национальный университет им. Л.Н. Гумилёва, Астана, Казахстан.

Научный руководитель: Бокаев Нуржан Адилханович — доктор физико-математических наук, профессор, Евразийский национальный университет им. Л.Н. Гумилёва.

© И.С. Толеубаева, 2026

Аннотация: Цель работы – проверить, насколько интервалы прогноза статистических моделей соответствуют заявленной вероятности и можно ли простыми средствами устранить систематическое занижение прогнозов месячной выработки электроэнергии в Казахстане. Для сезонной наивной модели, ETS (M, A, A), SARIMA и их комбинации по данным за январь 2018 – март 2026 г. построено по 402 вневыборочных прогноза с горизонтом 1-12 месяцев и оценены покрытие 80- и 95-процентных интервалов, их ширина и оценка Винклера. Интервалы ETS на горизонтах до шести месяцев оказались слишком узкими *(покрытие 85-87% вместо 95%)*, интервалы SARIMA и комбинации моделей были ближе к номиналу, а эмпирические интервалы по прошлым ошибкам работали хуже модельных. Все модели занижали прогноз, и смещение росло с горизонтом до 0,22-0,27 ТВт·ч в месяц; поправка на среднюю недавнюю ошибку улучшила оценку Винклера SARIMA на 9%, но долгосрочное смещение не устранила.

Ключевые слова: интервальный прогноз, покрытие интервала, оценка Винклера, смещение прогноза, SARIMA, экспоненциальное сглаживание, выработка электроэнергии, Казахстан.

Введение

Месячный прогноз выработки электроэнергии нужен системному оператору и генерирующим компаниям для планирования ремонтов, закупок топлива и перетоков. В Казахстане его цена в последние годы выросла: в 2024 г. внутреннее потребление превысило собственное производство примерно на 2 млрд кВт·ч, и разницу пришлось покрывать импортом [1]. В такой ситуации важна не только точечная оценка, но и диапазон, в котором с заданной вероятностью окажется фактическое значение: именно он показывает, какой резерв мощности или объём импорта нужно предусмотреть.

Интервалы прогноза статистических моделей рассчитываются в предположении, что модель верно описывает ряд, а её параметры известны. На практике оба условия нарушаются, поэтому фактическое покрытие интервалов часто оказывается ниже номинального [2]. Это подтвердило и соревнование M4, где интервалы большинства методов на месячных рядах были слишком узкими [3]. В литературе по прогнозированию электропотребления вероятностному прогнозу уделяется всё больше внимания [4], однако

для месячных рядов Казахстана проверка интервалов, насколько нам известно, не проводилась.

В предыдущей работе автора на том же ряду было показано, что ETS, SARIMA, XGBoost и LSTM дают практически одинаковую точность точечного прогноза (MAPE около 3,4%), но все модели занижали выработку в период ускоренного роста 2024–2025 гг. Настоящая статья продолжает это исследование и отвечает на два вопроса, которые там не рассматривались: (1) соответствует ли фактическое покрытие интервалов прогноза номиналу; (2) можно ли уменьшить систематическое смещение простыми приёмами – оценкой на скользящем окне, поправкой на недавнюю ошибку или комбинированием моделей.

Материалы и методы

Данные. Используется ряд месячной суммарной выработки электроэнергии в Казахстане (ТВт·ч) по данным Ember [5] за январь 2018 – март 2026 г., всего 99 наблюдений. Ряд имеет выраженную годовую сезонность с максимумом в декабре–январе и минимумом летом, а также восходящий тренд: годовая выработка выросла со 106,8 ТВт·ч в 2018 г. до 123,1 ТВт·ч в 2025 г. Следует отметить, что в 18 случаях значение месяца в точности повторяет предыдущее; по-видимому, это оценки, которыми Ember заполняет пропуски в первичных месячных данных. Мы сохраняем ряд без изменений, чтобы результаты были сопоставимы с предыдущей работой автора.

Модели. Рассматриваются три модели, чьи интервалы прогноза выводятся из самой модели. Первая – сезонная наивная, прогнозирующая значение того же месяца год назад; её интервал строится по стандартному отклонению сезонных разностей [6]. Вторая – экспоненциальное сглаживание ETS(M,A,A) с мультипликативной ошибкой, аддитивными трендом и сезонностью [7]; так как ошибка мультипликативная, интервалы получены по 2000 смоделированным траекториям. Третья – SARIMA$(1,0,0)(0,1,1)_{12}$ с константой [8]. Структуры ETS и SARIMA выбраны по критерию AICc на данных 2018-2022 гг. и далее не менялись, параметры переоценивались в каждой точке начала прогноза. Четвёртый вариант – комбинация ETS и SARIMA: точечный прогноз равен среднему двух прогнозов [9], а границы интервала – средним соответствующих границ, то есть усреднению квантилей [10].

Эмпирические интервалы. В качестве альтернативы модельным интервалам использованы интервалы, построенные по фактическим вневыборочным ошибкам [11]. Для прогноза из точки t на горизонт h берутся все прошлые ошибки того же горизонта, целевой месяц которых уже наблюдался к моменту t, и интервал задаётся их выборочными квантилями:

$$\left[\hat{y}_{t+h\vee t}+q^{(h)}_{\alpha/2};\ \hat{y}_{t+h\vee t}+q^{(h)}_{1-\alpha/2}\right] \qquad (1)$$

где $q^{(h)}$ - квантили прошлых ошибок горизонта h, α = 0,05 или 0,20. Чтобы к началу тестового периода накопились ошибки, прогнозы дополнительно строились из 24 «разогревочных» точек начала с декабря 2020 по ноябрь 2022 г.; в оценку они не входят.

Коррекция смещения. Проверены два приёма. Первый – оценка моделей на скользящем окне последних 60 месяцев вместо всей доступной истории; он должен помочь, если параметры ряда меняются во времени [12]. Второй – поправка свободного члена (intercept correction) [13]: к прогнозу добавляется средняя ошибка внутривыборочной подгонки за последние 12 месяцев,

$$\hat{y}^{ic}_{t+h\vee t}=\hat{y}_{t+h\vee t}+\frac{1}{12}\sum_{j=0}^{11} e_{t-j} \qquad (2)$$

где e_t – остатки модели. Если модель в последний год систематически недооценивала ряд, поправка сдвигает прогноз и интервал вверх.

Схема оценки. Как и в предыдущей работе автора, прогноз строился методом скользящего начала: из каждой из 39 точек с декабря 2022 по февраль 2026 г. все модели заново оценивались на доступных данных и давали прогноз на 1-12 месяцев вперёд. Для каждой модели получено 402 вневыборочных прогноза (от 39 на горизонте 1 до 28 на

горизонте 12). Точность точечного прогноза измерялась средней абсолютной процентной ошибкой (MAPE) и средней ошибкой ME = факт − прогноз; положительная ME означает занижение. Качество интервалов оценивалось по фактическому покрытию

$$C_h = \frac{1}{N_h} \sum_{t}^{\square} 1\left(l_{t+h \vee t} \le y_{t+h} \le u_{t+h \vee t}\right) \quad (3)$$

где $\mathbb{1}(\cdot)$ – индикатор, N_h – число прогнозов на горизонте h, по средней ширине интервала и по оценке Винклера [14]:

$$W_t = \left(u_t - l_t\right) + \frac{2}{\alpha}\left(l_t - y_t\right) 1\left(y_t < l_t\right) + \frac{2}{\alpha}\left(y_t - u_t\right) 1\left(y_t > u_t\right) \quad (4)$$

Оценка Винклера (4) объединяет ширину интервала и штраф за выход факта за его границы: чем она меньше, тем лучше интервал [15]. Отклонение покрытия от номинала проверялось тестом Купика [16], значимость смещения на каждом горизонте – t-тестом средней ошибки со стандартными ошибками Ньюи–Уэста [17] с лагом h − 1. Поскольку прогнозы с соседних точек начала перекрываются, результаты обоих тестов следует понимать, как ориентировочные. Расчёты выполнены в Python (statsmodels).

Результаты и обсуждение

Сводные результаты по всем горизонтам приведены в таблице 1.

Таблица 1 – Точность и качество интервалов прогноза, январь 2023 – март 2026 г. (402 прогноза на модель; ME и ширина – в ТВт·ч, покрытие – в %)

Модель	MAPE, %	ME	Покры-тие 80%	Покры-тие 95%	Ширина 95%	Winkler 95%
Сезонная наивная	3,93	0,30	79,3	97,0	1,81	2,06
ETS(M,A,A)	3,56	0,12	73,6	90,3	1,55	2,14
SARIMA	3,38	0,13	73,9	92,5	1,43	1,98
Комбинация ETS+SARIMA	3,35	0,12	74,6	93,3	1,49	1,89
ETS, окно 60 мес.	3,53	0,10	69,7	87,3	1,42	2,25
SARIMA, окно 60 мес.	3,42	0,11	71,1	93,0	1,49	2,02
ETS + поправка (2)	3,60	0,10	72,4	89,5	1,55	2,20
SARIMA + поправка (2)	3,32	0,13	74,9	92,5	1,43	1,81
ETS, эмпирич. интервалы	3,56	0,12	65,4	82,3	1,55	2,67
SARIMA, эмпирич. интервалы	3,38	0,13	70,4	86,3	1,48	2,33
Комбинация, эмпирич. интервалы	3,35	0,12	66,9	84,1	1,45	2,47

Интервалы ETS слишком узкие на коротких горизонтах. При номинале 95% фактическое покрытие интервалов ETS составило 86,8% на горизонтах 1-3 месяца и 84,8% на горизонтах 4–6 месяцев (таблица 2, рисунок 1); тест Купика отвергает соответствие номиналу ($p < 0{,}001$). Только на горизонтах 7-12 месяцев покрытие ETS приблизилось к номиналу. Интервалы SARIMA ведут себя ровнее: 91-93% на всех группах горизонтов. Восьмидесятипроцентные интервалы ETS, SARIMA и их комбинации покрыли на горизонтах до шести месяцев лишь 68-74% фактических значений. Это согласуется с выводом [2, 3] о том, что модельные интервалы не учитывают неопределённость выбора модели и оценки параметров.

Таблица 2 – Фактическое покрытие 80- и 95-процентных интервалов по группам горизонтов h, %

Модель	h = 1-3		h = 4-6		h = 7-12	
	80%	95%	80%	95%	80%	95%
Сезонная наивная	80,7	97,4	81,9	97,1	77,0	96,7

Модель	h = 1-3		h = 4-6		h = 7-12	
	80%	95%	80%	95%	80%	95%
ETS(M,A,A)	68,4	86,8	67,6	84,8	80,3	95,6
SARIMA	73,7	92,1	71,4	91,4	75,4	93,4
Комбинация ETS+SARIMA	69,3	89,5	68,6	91,4	81,4	96,7
SARIMA + поправка (2)	74,6	93,9	75,2	91,4	74,9	92,3
SARIMA, эмпирич. интервалы	74,6	92,1	68,6	85,7	68,9	83,1

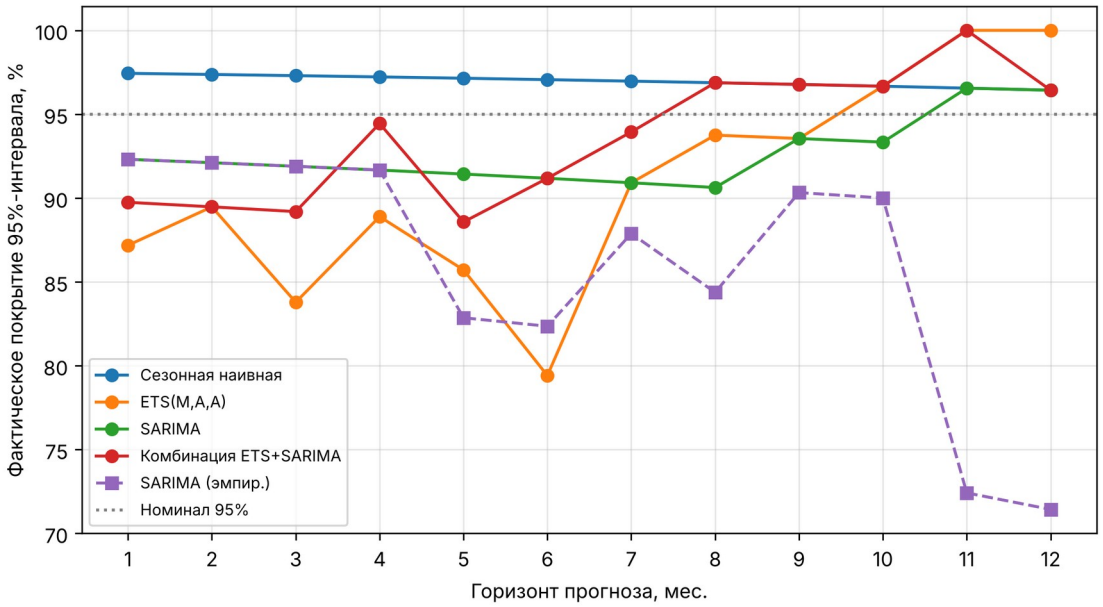


Рисунок 1 – Фактическое покрытие 95-процентного интервала по горизонтам прогноза

Сезонная наивная модель даёт «честные», но широкие интервалы. Её покрытие составило 97% при номинале 95%, то есть интервалы даже шире необходимого: их средняя ширина 1,81 ТВт·ч против 1,43 ТВт·ч у SARIMA. В результате по оценке Винклера наивная модель (2,06) уступает SARIMA (1,98), хотя и опережает ETS (2,14).

Комбинация моделей улучшает интервалы. Усреднение ETS и SARIMA дало покрытие 93,3% и наилучшую среди модельных интервалов оценку Винклера без поправок (1,89); тест Купика не отверг для неё номинал 95% ($p = 0,13$), тогда как для ETS и SARIMA отверг ($p < 0,001$ и $p = 0,03$). Выигрыш объясняется тем, что ошибки двух моделей частично компенсируют друг друга, а ширина усреднённого интервала остаётся промежуточной.

Эмпирические интервалы оказались хуже модельных. Вопреки ожиданиям, интервалы по прошлым ошибкам покрыли лишь 82–86% фактических значений. Причина в короткой истории: в начале тестового периода квантили оценивались по нескольким десяткам ошибок спокойных 2021–2022 гг., которые не отражали последующего ускорения роста. Для месячных рядов длиной около 100 наблюдений этот подход, следовательно, требует осторожности.

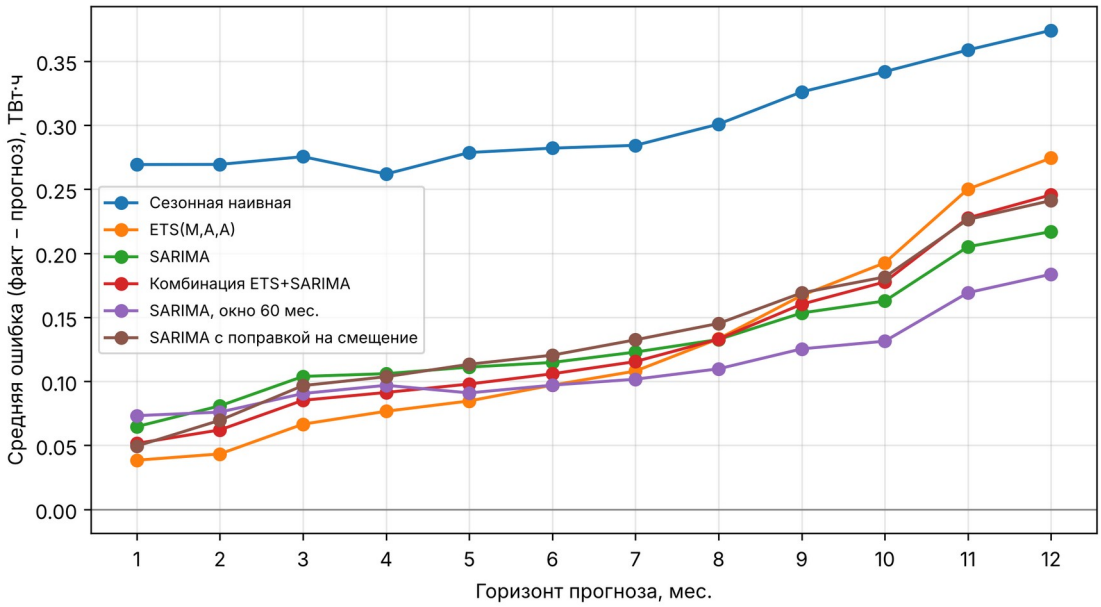


Рисунок 2 - Средняя ошибка прогноза (факт – прогноз) по горизонтам

Все модели систематически занижают прогноз. Средняя ошибка положительна на всех горизонтах и растёт с горизонтом – от 0,04-0,07 ТВт·ч на горизонте 1 до 0,22-0,27 ТВт·ч на горизонте 12, то есть примерно до 2,5% месячной выработки (рисунок 2). На горизонтах от 9-10 месяцев и далее смещение статистически значимо ($p < 0,05$). Разбивка по годам показывает, что оно связано со сменой режима: в 2023 г. модели, напротив, завышали прогноз (ME SARIMA -0,31 ТВт·ч), а в 2024 и 2025 гг. занижали его (+0,15 и +0,32 ТВт·ч). Показательно, что все 20 случаев выхода факта выше верхней границы 95-процентного интервала SARIMA пришлись на 2025 г., а все 10 случаев выхода ниже нижней границы – на 2023 г. Модели, оценённые на периоде умеренного роста, не успевали за ускорением спроса (рисунок 3).

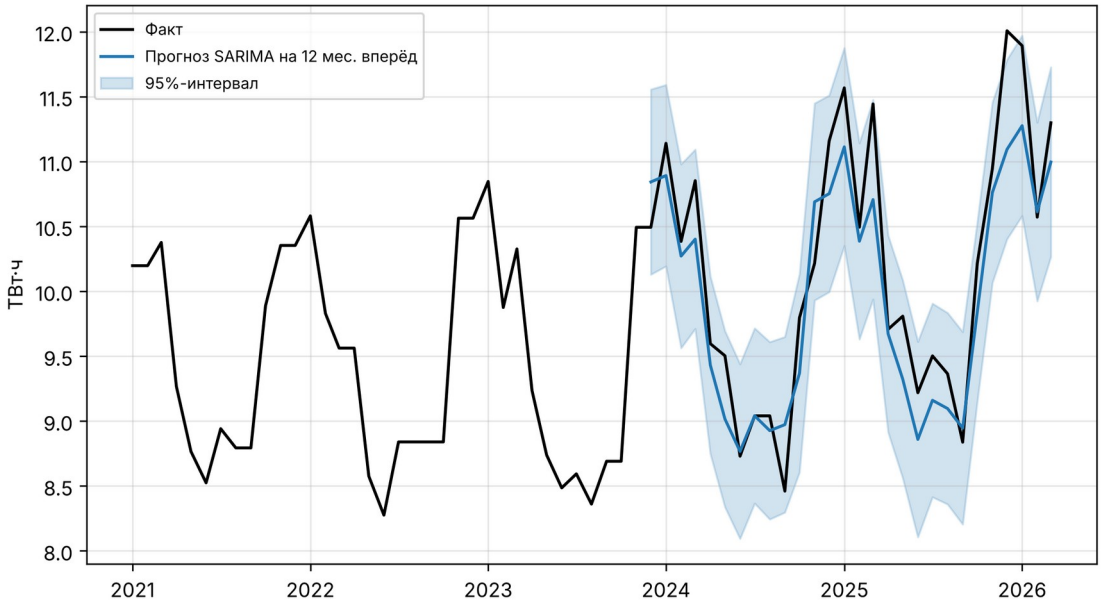


Рисунок 3 – Прогноз SARIMA на 12 месяцев вперёд и его 95-процентный интервал

Простые приёмы коррекции помогают лишь частично. Оценка на скользящем 60-месячном окне не дала выигрыша: MAPE SARIMA выросла с 3,38 до 3,42%, а покрытие интервалов ETS снизилось до 87,3% из-за меньшего объёма данных. Поправка (2) оказалась

полезнее: для SARIMA она уменьшила MAPE до 3,32%, подняла покрытие 95-процентного интервала на горизонтах 1-3 месяца до 93,9% и улучшила оценку Винклера с 1,98 до 1,81, то есть на 9% - это лучший результат среди всех вариантов. Однако общее смещение поправка не устранила (ME 0,13 ТВт·ч): она переносит в прогноз недавний уровень ошибок, но не изменение наклона тренда, которое и порождает растущее с горизонтом занижение.

Ограничения. Тестовый период охватывает чуть более трёх лет, прогнозы с соседних точек начала перекрываются, а часть месячных значений в исходных данных является оценками. Поэтому численные значения покрытия следует рассматривать как ориентировочные; качественные выводы при этом устойчивы для всех рассмотренных моделей.

Заключение

Проверка интервальных прогнозов месячной выработки электроэнергии в Казахстане показала, что заявленная вероятность интервалов статистических моделей выполняется не всегда. Интервалы ETS на горизонтах до полугода слишком узки, восьмидесятипроцентные интервалы ETS и SARIMA покрывают около трёх четвертей фактических значений, а эмпирические интервалы при короткой истории работают ещё хуже. Наиболее надёжный результат дали SARIMA с поправкой на недавнюю ошибку и комбинация ETS и SARIMA. Для практического планирования это означает, что номинальные интервалы стоит расширять или калибровать, особенно на горизонте до шести месяцев.

Систематическое занижение прогноза в 2024-2025 гг. не устраняется ни скользящим окном, ни поправкой свободного члена, поскольку вызвано ускорением роста спроса, которого нет в истории ряда. Дальнейшее повышение качества прогноза связано с включением внешней информации – температуры воздуха и данных о вводе крупных потребителей, – что будет рассмотрено в последующих работах в рамках магистерской диссертации.

Список литературы:

1. АО «KEGOC». Годовой отчёт 2024: баланс электроэнергии. – URL: https://ar2024.kegoc.kz/ru/electricity-balance.html (дата обращения: 17.11.2026).

2. Chatfield C. Calculating interval forecasts // Journal of Business & Economic Statistics. – 1993. – Vol. 11, No. 2. – P. 121–135.

3. Makridakis S., Spiliotis E., Assimakopoulos V. The M4 Competition: 100,000-time series and 61 forecasting methods // International Journal of Forecasting. – 2020. – Vol. 36, No. 1. – P. 54–74.

4. Hong T., Fan S. Probabilistic electric load forecasting: a tutorial review // International Journal of Forecasting. – 2016. – Vol. 32, No. 3. – P. 914–938.

5. Ember. Monthly electricity data. – URL: https://ember-energy.org/data/monthly-electricity-data/ (дата обращения: 17.11.2026).

6. Hyndman R.J., Athanasopoulos G. Forecasting: principles and practice. – 3rd ed. – Melbourne: OTexts, 2021. – URL: https://otexts.com/fpp3/ (дата обращения: 17.11.2026).

7. Hyndman R.J., Koehler A.B., Snyder R.D., Grose S. A state space framework for automatic forecasting using exponential smoothing methods // International Journal of Forecasting. – 2002. – Vol. 18, No. 3. – P. 439–454.

8. Box G.E.P., Jenkins G.M., Reinsel G.C., Ljung G.M. Time series analysis: forecasting and control. – 5th ed. – Hoboken: Wiley, 2015. – 712 p.

9. Bates J.M., Granger C.W.J. The combination of forecasts // Operational Research Quarterly. – 1969. – Vol. 20, No. 4. – P. 451–468.

10. Lichtendahl K.C., Grushka-Cockayne Y., Winkler R.L. Is it better to average probabilities or quantiles? // Management Science. – 2013. – Vol. 59, No. 7. – P. 1594–1611.

11. Williams W.H., Goodman M.L. A simple method for the construction of empirical confidence limits for economic forecasts // Journal of the American Statistical Association. – 1971. – Vol. 66, No. 336. – P. 752–754.

12. Pesaran M.H., Timmermann A. Selection of estimation window in the presence of breaks // Journal of Econometrics. – 2007. – Vol. 137, No. 1. – P. 134–161.

13. Clements M.P., Hendry D.F. Forecasting economic time series. – Cambridge: Cambridge University Press, 1998. – 368 p.

14. Winkler R.L. A decision-theoretic approach to interval estimation // Journal of the American Statistical Association. – 1972. – Vol. 67, No. 337. – P. 187–191.

15. Gneiting T., Raftery A.E. Strictly proper scoring rules, prediction, and estimation // Journal of the American Statistical Association. – 2007. – Vol. 102, No. 477. – P. 359–378.

16. Kupiec P.H. Techniques for verifying the accuracy of risk measurement models // The Journal of Derivatives. – 1995. – Vol. 3, No. 2. – P. 73–84.

17. Newey W.K., West K.D. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix // Econometrica. – 1987. – Vol. 55, No. 3. – P. 703–708.

© Central Asian Scientific Journal, 2026