
А. Подгорный

**СТОХАСТИЧЕСКИЕ СИМУЛЯЦИИ
МАКРОЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ ДЛЯ
ДОЛГОСРОЧНОГО ПЛАНИРОВАНИЯ В ПРОЕКТНОМ
ФИНАНСИРОВАНИИ**

В статье представлена комплексная методология прогнозирования макроэкономических показателей для долгосрочного планирования в проектном финансировании. Цель исследования состоит в разработке системы стохастических симуляций, способной генерировать правдоподобные сценарии экономического развития с учетом взаимосвязей между различными экономическими параметрами. Методология включает два ключевых компонента: алгоритм отбора значимых предикторов на основе разреженных графов и минимального дерева Штейнера, а также систему стохастических симуляций, интегрирующую модель CIR++ с методом Монте-Карло.

Автором разработан эффективный алгоритм построения регрессионных моделей, учитывающий структурные взаимосвязи между экономическими показателями.

Материалом исследования послужили исторические данные по широкому спектру макроэкономических показателей России: ВВП, инфляция, процентные ставки, индексы цен на недвижимость и показатели просроченной задолженности по кредитам. Результаты применения методологии демонстрируют высокую точность прогнозирования на исторических данных и интуитивно понятное поведение в долгосрочной перспективе. Система способна генерировать вероятностные сценарии с горизонтом до тридцати лет, что позволяет оценивать различные аспекты рисков, включая экстремальные сценарии. Модульная ар-

Андрей Сергеевич Подгорный – старший преподаватель, Дальневосточный федеральный университет (г. Владивосток).

хитектура системы обеспечивает гибкость и адаптируемость к различным экономическим условиям.

Результаты исследования имеют практическое значение для управления рисками в финансовых институтах и стратегического планирования в проектном финансировании.

Ключевые слова: разреженные графы, отбор признаков, дерево Штейнера, регрессионный анализ, метод Монте-Карло, модель Кокса – Ингерсолла – Росса, проектное финансирование, стохастическое моделирование, макроэкономическое прогнозирование.

УДК: 004.942:330.43

EDN: YG1K1Y

DOI: 10.51905/2073–038_2025_2S_58

Введение

Проектное финансирование требует принятия решений на основе долгосрочных макроэкономических прогнозов. Существующие модели часто не учитывают сложные взаимозависимости между разнородными экономическими индикаторами, что приводит к существенным ошибкам на протяженных горизонтах прогнозирования [1]. Представленное исследование призвано преодолеть эти ограничения путем разработки интегрированного подхода, синтезирующего методы машинного обучения и экономического моделирования. Классические регрессионные модели и ARIMA-подходы обладают присущими ограничениями при работе на длинных горизонтах: плохо адаптируются к внезапным структурным сдвигам и недостаточно хорошо отражают неопределенность и волатильность, характерные для макроэкономических факторов [2, 3].

Цель исследования

Основная цель – разработать гибкую и масштабируемую систему прогнозирования макроэкономических показателей, учитывающую структурные связи между переменными и генерирующую вероятностные сценарии для долгосрочного анализа рисков.

Для достижения цели поставлены следующие задачи:

- разработать методологию отбора значимых предикторов с использованием разреженных графов и алгоритма дерева Штейнера;
- интегрировать регрессионные модели в систему стохастических симуляций на основе модели $CIR++$ (Кокса – Ингерсолла – Росса) и методов Монте-Карло;
- оценить эффективность подхода на данных российской экономики, включая симуляции для ВВП, инфляции и процентных ставок.

Материалы и методы исследования

Представлен интегрированный подход к построению системы прогнозирования. Он сочетает методологию отбора макроэкономических предикторов посредством анализа разреженных графов и их использование в системе стохастических симуляций, генерирующей вероятностные экономические сценарии. Сценарии формируются на основе модели CIR++ [4, 5, 6] и метода Монте-Карло, применительно к таким индикаторам, как процентные ставки, индексы цен на жилье, ипотечные ставки и другие переменные, с явным учетом их корреляционных зависимостей. Методология включает калибровку к рыночным данным, стохастическое моделирование с месячной дискретизацией [7, 8] и регрессионные связи между переменными. Коэффициенты этих связей определяются с применением алгоритма минимального дерева Штейнера (*PCST*, *Prize-Collecting Steiner Tree*) на разреженных графах [9, 10].

Алгоритм *PCST* минимизирует целевую функцию [11]:

$$c(\mathcal{T}) = \sum_{v_i \in V \setminus V_T} p_i + \sum_{e_{ij} \in E_T} w_{ij}, \quad (1)$$

где V_T – множество выбранных вершин, E_T – множество выбранных ребер.

Преимущества этого метода для отбора признаков заключаются в балансе между индивидуальной значимостью признаков и их связностью, учете структурных взаимосвязей, контроле разреженности и полной интерпретируемости.

В качестве практической реализации предложенного метода были использованы регрессии для прогнозирования макроэкономических показателей.

Псевдокод для отбора признаков на основе *PCST* следующий.

Этап 1. Отбор потенциально значимых переменных.

1.1. Выбрать потенциально значимые признаки на основе экономического смысла и экспертных оценок.

1.2. Сформировать список признаков $X_{possible}$, потенциально влияющих на целевую переменную y .

Этап 2. Построение графа и отбор признаков.

2.1. Построить граф признаков G :

a) вычислить коэффициенты корреляции Пирсона между всеми парами признаков в $X_{possible}$;

b) для каждой пары признаков (i, j) :

– рассчитать вес ребра $W_{ij} = 1 - r_{ij}$, где r_{ij} – коэффициент корреляции между признаками i и j ;

– если $r_{ij} \geq V$ пороговое значение, добавить ребро между вершинами i и j с весом W_{ij} ;

c) создать список ребер: $edges = [(i, j, W_{ij})]$ для всех подходящих пар (i, j) ;

d) присвоить приз вершины (важность признака): $p_i = r_{iy} \forall$, где r_{iy} – корреляция признака i с целевой переменной y .

2.2. Настроить параметры алгоритма минимального дерева Штейнера (PCST):

– $edges_array$ = массив ребер;

– $prizes$ = массив призов вершин ;

– $root = -1$ (если корень не определен);

– $num_clusters$ = желаемое количество кластеров;

– $pruning$ = метод обрезки (strong или другой);

– $sparsity_low$ и $sparsity_high$ = границы разреженности (минимальное и максимальное количество признаков).

2.3. Применить алгоритм PCST для отбора признаков к графу G .

Найти подграф T , минимизирующий функцию стоимости:

$C(T) = \sum (p_i \text{ для вершин не в } T) + \sum (W_{ij} \text{ для ребер в } T)$.

a) получить выбранные признаки $V_T =$ вершины из подграфа T ;

b) сформировать новый набор признаков $X_selected$, содержащий только выбранные признаки V_T .

Этап 3. Обучение модели линейной регрессии с помощью нового метода.

3.1. Инициализация:

a) инициализировать вектор весов коэффициентов $W =$ нулевой вектор размерности, равной числу признаков в $X_selected$;

b) установить параметры обучения:

– количество итераций $num_iterations$;

– массив размеров шагов градиентного спуска $steps$ (например, $[0.1, 0.01, 0.001] \dots$ повторяется $num_iterations$ раз).

3.2. Итеративное обучение модели.

Для каждой итерации k от 1 до $num_iterations$:

a) вычислить градиент функции потерь MSE:

$grad = -2 * X_{selected}^T * (y - X_{selected} * W) / n$;

b) обновить веса модели для каждого размера шага в $steps_k$:

для каждого $step_size$ в $steps_k$:

i. Временно обновить веса:

$W_temp = W - step_size * grad$

ii. Проекция на структуру графа:

– возвести элементы W_temp в квадрат: $W_squared = W_temp^2$,

– применить функцию проекции разреженности с использованием PCST:

$W_proj_squared, _, _ = graph_proj_sparsity(W_squared, sparsity_low, sparsity_high, graph_opts, \dots)$

– Восстановить знаки коэффициентов:

$W_proj = sign(W_temp) * sqrt(W_proj_squared)$

iii. Вычислить функцию потерь:

$$loss = MSE(y, X_selected * W_proj)$$

iv. Если $loss < best_loss$:

– обновить $best_loss = loss$

– обновить $best_W = W_proj$

c) обновить веса модели:

$$W = best_W$$

d) мониторинг сходимости и корректировка параметров обучения.

3.3. Выполнить статистические тесты на обученной модели:

a) тест значимости коэффициентов (*t*-тест);

b) тест нормальности распределения остатков (тест Шапиро – Уилка или Колмогорова – Смирнова);

c) тест гетероскедастичности остатков (тест Бройша – Пагана);

d) тест отсутствия автокорреляции остатков (тест Дарбина – Уотсона).

Этап 4. Визуализация и анализ результатов модели.

1. Построить графики фактических и прогнозных значений целевой переменной.

2. Проанализировать выбранные признаки и их веса для интерпретации модели.

Результаты исследования и их обсуждение

Разработанная методология была применена к реальным данным российской макроэкономики. Ниже представлены избранные результаты. Моделирование макроэкономических параметров осложняется сменами фаз развития российской экономики в сочетании с ограниченной доступностью данных и неоднородным качеством исторических данных. Для повышения точности и устойчивости моделей отбора признаков был применен алгоритм минимального дерева Штейнера, позволяющий интегрировать как индивидуальную значимость признаков, так и их взаимосвязи.

В рамках представленного исследования были разработаны различные модели для макроэкономических показателей: индекс потребительских цен, спреда между ключевой ставкой Банка России и ставками денежного рынка, индексы цен на недвижимость и другие показатели.

Во всех вышеупомянутых моделях алгоритм PCST был успешно применен для отбора признаков, обеспечивая учет как индивидуальной значимости переменных, так и их взаимосвязей. Это позволило создать стабильные и точные модели с высокой объяснительной способностью, подходящие для сценарного моделирования и прогнозирования макроэкономических показателей. Применение алгоритма PCST повысило качество моделей за счет интеграции вза-

имозависимостей признаков, тем самым улучшив надежность и интерпретируемость результатов.

Визуализация регрессии для спреда между ключевой ставкой и доходностью 1-месячной бескупонной облигации показана на рис. 1.

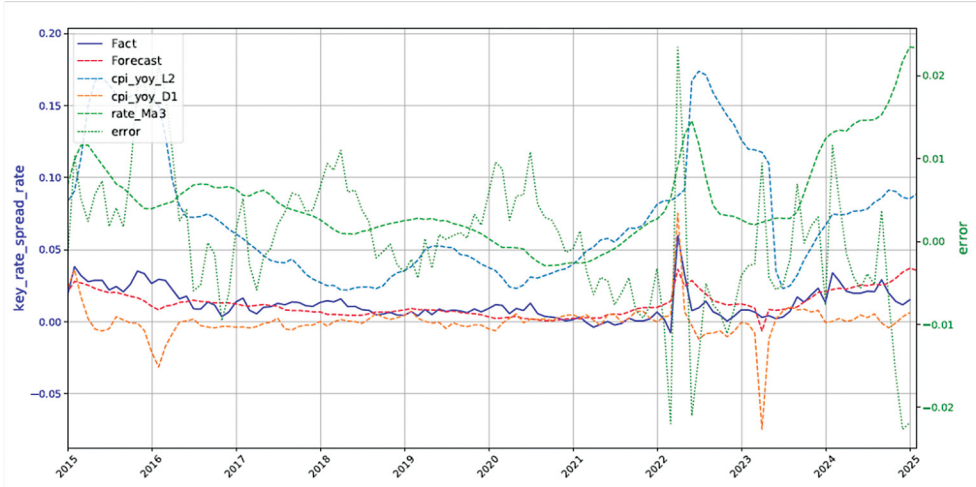


Рис. 1. Пример реализации *PCST* для прогнозирования спреда между ключевой ставкой и доходностью 1-месячной бескупонной облигации.

Источники: составлено автором.

На основе предварительно построенных регрессионных моделей была разработана комплексная система стохастических макроэкономических симуляций¹. Архитектура системы состоит из четырех взаимосвязанных модулей.

1. Модуль моделирования процентных ставок (*CIR++*).

Процентные ставки моделируются с использованием стохастической модели *CIR++*, модифицированной для соответствия наблюдаемой рыночной кривой доходности:

- латентная краткосрочная ставка $x(t)$ следует стохастическому дифференциальному уравнению:

$$dx_t = a(\theta - x_t)dt + \sigma\sqrt{x_t}dW_t ;$$

- наблюдаемая краткосрочная ставка вычисляется как:

$$r_t = x_t + \varphi(t) ,$$

где $\varphi(t)$ – детерминированная функция сдвига, откалиброванная для соответствия рыночной форвардной кривой;

¹ Экспериментальный код доступен на https://github.com/andreypodgorny10/macro_sims

- модель дискретизируется с использованием схемы Эйлера – Маруямы с месячным шагом $\Delta t = \frac{1}{12}$:

$$x_{t+1} = x_t + [\alpha \cdot \theta_0 - (\alpha + B_{T-t} \cdot \sigma^2 \cdot \mathbf{1}_{T-measure}) \cdot x_t] \cdot \Delta t + \sigma \cdot \sqrt{\Delta t} \cdot \varepsilon_t ;$$

- функции B_{T-t} и $\varphi(t)$ вычисляются аналитически на этапе калибровки на основе форвардной кривой ставок.

1. Модуль генерации коррелированных стохастических шоков [12]:

- из исторических данных строится матрица корреляций R между макроэкономическими переменными;
- для симуляции коррелированных шоков используется разложение Холецкого:

$$L = \text{cholesky}(R);$$

- независимые стандартные нормальные векторы $Z \sim N(0, I)$ преобразуются в коррелированные векторы:

$$\varepsilon = L \cdot Z;$$

- полученные шоки одновременно применяются ко всем переменным в регрессионных моделях, обеспечивая внутренне согласованные траектории сценариев.

2. Модуль регрессионного моделирования макроэкономических переменных:

- каждая макроэкономическая переменная моделируется с использованием стохастической регрессии, включающей
 - лаговые члены (например, X_{t-1}, X_{t-2}),
 - разности (например, $X_t - X_{t-1}$),
 - скользящие средние,
 - индикаторные переменные и эффекты взаимодействия (например, условные зависимости от режимов инфляции).

Общая структура регрессионной модели для переменной Y_t задается:

$$Y_t = \beta_0 + \sum_{i=1}^k \beta_i X_{i,t-i} + \sum_{j=1}^m \gamma_j (Z_{j,t} - Z_{j,t-d_j}) + \sum_{p=1}^n \delta_p Y_{t-p} + \sigma \cdot \varepsilon_t,$$

где $X_{i,t-i}$: лаговые объясняющие переменные, $(Z_{j,t} - Z_{j,t-d_j})$: разности первого или более высокого порядка, Y_{t-p} : авторегрессионные члены целевой переменной, $\varepsilon_t \sim N(0, \mathbf{1})$: стохастический шок, коррелированный между переменными согласно модулю 2.

3. Оркестрация симуляций.

Сводка содержит такие статистики, включая следующие:

- *Value-at-Risk (VaR)* и *Expected Shortfall (ES)*;
- агрегаты сценариев на основе медианы;
- доверительные интервалы (например, 5-й и 95-й процентиля).

Модель *CIR++* служит центральным элементом системы симуляций, генерируя прогнозы процентных ставок с учетом их стохастической природы и динамических характеристик [13].

Метод Монте-Карло [14, с. 594] используется для генерации распределения возможных будущих состояний для каждой макроэкономиче-

ской переменной. Ключевым аспектом реалистичного моделирования макроэкономических переменных является учет корреляций между ними. Система использует метод разложения Холецкого для генерации коррелированных случайных величин. Основное преимущество модели $CIR++$ заключается в ее способности одновременно моделировать возврат ставок к долгосрочному среднему значению и поддерживать калибровку к наблюдаемой рыночной кривой доходности [15].

Финальным элементом реализации является модуль визуализации. Этот модуль обрабатывает и представляет результаты симуляций, включая расчет статистических метрик и идентификацию различных сценариев. Примеры вывода модуля визуализации представлены на рис. 2.

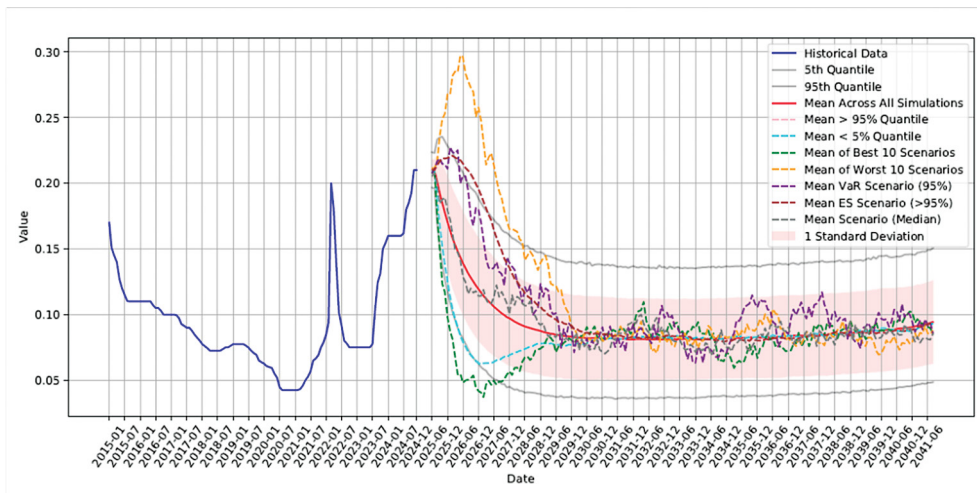


Рис. 2. Пример вывода модуля визуализации для ключевой ставки.
Источники: составлено автором.

* * *

Результаты показывают эффективность предложенного подхода к прогнозированию макроиндикаторов с учетом их взаимозависимостей и стохастической природы. Ключевые достоинства системы – гибкость, масштабируемость и адаптивность к разным экономическим условиям. Подход повышает точность оценки рисков для долгосрочных проектов жилищной недвижимости и улучшает качество финансового планирования.

Список литературы / References

1. Bekaert G., Engstrom E., Xu N. R. The time variation in risk appetite and uncertainty. *Management Science*. 2022. Vol. 68. Is. 6. pp. 3975–4004. <https://doi.org/10.1287/mnsc.2021.4068>.

2. *Antolin-Diaz J., Petrella I., Rubio-Ramirez J. F.* Structural scenario analysis with SVARs. *Journal of Monetary Economics*. 2021. Vol. 117. pp. 798–815. <https://doi.org/10.1016/j.jmoneco.2020.06.001>.

3. *Kaushik M., Giri A.* Forecasting Foreign Exchange Rate: A Multivariate Comparative Analysis between Traditional Econometric, Contemporary Machine Learning & Deep Learning Techniques. 2020. <https://doi.org/10.48550/arXiv.2002.10247>.

4. *Akhtari B., Hanwu Li.* The Cox-Ingersoll-Ross process under volatility uncertainty. *Journal of Mathematical Analysis and Applications*. 2024. Vol. 531. Is. 1. Part 1. pp. 2–31. DOI: 10.1016/j.jmaa.2023.127867.

5. *Bayraktar E., Clément E.* Estimation of a pure-jump stable Cox-Ingersoll-Ross process. *Bernoulli*. 2025. Vol. 31. Is. 1. pp. 484–508. DOI: 10.3150/24-BEJ1736.

6. *Chen R-R, Scott L.* Multi-Factor Cox-Ingersoll-Ross models of the term structure: estimates and tests from a Kalman filter model. *The Journal of Real Estate Finance and Economics*. 2003. Vol. 27. pp. 143–172. DOI: 10.1023/A:1024736903090.

7. *Poignard B., Asai M.* High-dimensional sparse multivariate stochastic volatility models // *Journal of Time Series Analysis*. 2023. Vol. 44. Is. 1. pp. 4–22. DOI: 10.1111/jtsa.12647.

8. *Rhijn J., Oosterlee C., Grzelak L., Shuaiqiang Liu.* Monte Carlo simulation of SDEs using GANs // *Japan Journal of Industrial and Applied Mathematics*. 2022. Vol. 40. pp. 1359–1390. DOI: 10.1007/s13160-022-00534-x.

9. *Hegde C., Indyk P., Schmidt L.* A nearly-linear time framework for graph-structured sparsity // *Proceedings of the 32th International Conference on Machine Learning*. Vol. 37. pp. 928–937 (2015).

10. *Giuan Yu, Yufeng Liu.* Sparse regression incorporating graphical structure among predictors // *Journal of the American Statistical Association*. 2016. Vol. 111. Is. 514. pp. 707–720. DOI: 10.1080/01621459.2015.1034319.

11. *Khodamoradi K., Krishnamurti R.* Prize collecting traveling salesman problem – Fast Heuristic Separations // *Proceedings of 5th the International Conference on Operations Research and Enterprise Systems*. 2016. Vol. 1. pp. 380–387. DOI: 10.5220/0005758103800387.

12. *Fernandez-Villaverde J., Guerron-Quintana P. A.* Uncertainty shocks and business cycle research. *Review of Economic Dynamics*. 2020. Vol. 37. pp. 118–146. DOI: 10.1016/j.red.2020.06.005.

13. *Dehtiar O., Mishura Y., Ralchenko K.* Two methods of estimation of the drift parameters of the Cox-Ingersoll-Ross process: continuous observations. 2020. <https://doi.org/10.48550/arXiv.2005.05262>.

14. *Халл Джон К.* Опционы, фьючерсы и другие производные финансовые инструменты / Джон К. Халл; [пер. с англ. и ред. Д. А. Ключина]. 8-е изд. Москва [и др.]: Вильямс, 2014. 1070 с.

15. Балабушкин А., Гамбаров Г., Никитин А., Шевчук И. Кривая бескупонной доходности на рынке ГКО-ОФЗ // Рынок ценных бумаг. 2006. №3. С. 68–77.

16. Di Francesco M., Kamn K. How to handle negative interest rates in a CIR framework // SeMa Journal. 2022. Vol. 79. pp. 593–618. DOI: 10.1007/s40324–021–00267-w.

Дата предоставления рукописи: 14 октября 2025 г.

About the Author

Andrei S. Podgorny – *a Senior Lecturer at the Far Eastern Federal University (Vladivostok).*

podgorny_as@dvfu.ru

Framework Simulations of Macroeconomic Indicators for Long-Term Planning in Project Financing

Abstract. The article presents a comprehensive methodology for forecasting macroeconomic indicators for long-term planning in project finance. The purpose of the research is to develop a stochastic simulation system capable of generating plausible scenarios of economic development taking into consideration the relationships between various economic parameters. The methodology includes two key components: an algorithm for selecting significant predictors based on sparse graphs and the minimum Steiner tree, and a system of stochastic simulations integrating the CIR++ model with the Monte Carlo method.

The author has developed an efficient algorithm for constructing regression models that takes into consideration structural relationships between economic indicators.

The study utilized historical data on a wide range of Russia's macroeconomic indicators: GDP, inflation, interest rates, real estate price indices, and loan delinquency rates. The results of applying the methodology demonstrate high accuracy of forecasting based on historical data and intuitively understandable behavior in the long term. The system is capable of generating probabilistic scenarios with horizon of up to 30 years, enabling the assessment of various risk aspects, including extreme scenarios. The system's modular architecture ensures flexibility and adaptability to various economic conditions.

The results of the study have practical significance for risk management in financial institutions and strategic planning in project finance.

Keywords: spars graphs, Feature selection, Prize-collecting Steiner Forest, regression, Monte Carlo method, Cox-Ingersoll-Ross Model, Project finance.