

УДК 004.81

Гибридный подход на основе когнитивного картирования и регрессионного анализа для прогнозирования в сложных слабо формализованных системах

© 2025 г. Р. М. Романов, А. И. Гусева

Национальный исследовательский ядерный университет «МИФИ», Москва, 115409, Россия

В статье представлен оригинальный гибридный подход, объединяющий нечеткие когнитивные карты и регрессионный анализ для прогнозирования в слабо формализованных системах, характеризующихся высокой степенью неопределенности и сложными, неструктурированными взаимосвязями между переменными. Основная идея подхода заключается в использовании экспертных оценок, выраженных через лингвистические переменные и нечеткие числа, для адаптации весовых коэффициентов регрессионной модели. Полученные веса, рассчитанные на основе когнитивного анализа, интегрируются в процедуру многопараметрической линейной регрессии с использованием взвешенного метода наименьших квадратов. Это позволяет повысить точность прогнозов и улучшить интерпретируемость модели. Результаты эмпирического исследования, проведенного в статистической среде R , демонстрируют, что предложенный подход превосходит ($MAPE = 10.21\%$) не только классические одиночные методы прогнозирования, такие как линейная регрессия ($MAPE = 3.80\%$), а также нейросетевые подходы, например, многослойный перцептрон ($MAPE = 31.24\%$), но и успешно конкурирует с ансамблевыми методами, включая случайный лес ($MAPE = 12.35\%$) и градиентный бустинг ($MAPE = 7.75\%$). Данный подход может быть успешно применен для прогнозирования объемов поставок газа в Китай, а также для решения других задач, требующих интеграции качественных и количественных данных.

Ключевые слова: нечеткая когнитивная карта, нечеткие числа, регрессионный анализ, прогнозирование, слабо формализованные системы, экспертная оценка.

Введение

Прогнозирование в слабо формализованных системах, таких как топливно-энергетический комплекс, сопряжено с высокой степенью неопределенности, неполнотой данных и трудностями формализации экспертных знаний. Современные модели, включая регрессионные и нейросетевые, при всей точности демонстрируют низкую интерпретируемость и ограниченную способность учитывать качественные характеристики.

В данной работе обосновывается гибридный подход, основанный на интеграции нечетких когнитивных карт (далее – НКК) и регрессионного анализа. НКК позволяют формализовать причинно-следственные связи с учетом нечеткости экспертных оценок, в то время как регрессионная модель обеспечивает количественное уточнение прогнозов. Особое внимание уделено передаче весов из системы НКК в структуру регрессионной модели.

Цель исследования – разработка и обоснование прогностического подхода, адаптированного к условиям неопределенности и слабо формализованной системы.

✉ Р.М. Романов: Rodion.Romanov.M@yandex.com

А.И. Гусева: aiguseva@mephi.ru

Поступила в редакцию: 30.06.2025

После доработки: 15.07.2025

Принята к публикации: 22.07.2025

EDN SPPAWL

Эмпирическая проверка выполнена в среде статистического анализа *R* (версия 4.3.3), что обеспечивает воспроизводимость результатов и соответствует современным требованиям исследовательской практики. Сравнительный анализ демонстрирует превосходство предлагаемого подхода по точности и интерпретируемости прогнозов по сравнению с традиционными методами.

Методы исследования

Нечеткие когнитивные карты

Когнитивная карта представляет собой формализованную модель знаний экспертов о свойствах и закономерностях развития изучаемой системы, формально описываемую в виде ориентированного знакового графа [1]. В когнитивной карте факторы отображаются узлами, а причинно-следственные отношения между ними – направленными связями. При построении нечеткой когнитивной карты взаимодействие факторов в рассматриваемой системе формализуется в виде направленного взвешенного графа [1, 3, 5, 7]:

$$G = \langle C, E \rangle, \quad (1)$$

где C – множество узлов, соответствующих ключевым параметрам (факторам); E – совокупность дуг, описывающих причинно-следственные связи между ними.

Каждой дуге $(C_i, C_j) \in E$ приписывается вес $w_{ij} \in [-1, 1]$, численно выражающий характер и силу влияния одного фактора на другой. Значения весов формируются на основе экспертных суждений, заданных в виде лингвистических переменных (например, «умеренное положительное», «значительное отрицательное»), которые затем дефазифицируются методом центра тяжести [1]:

$$z = \frac{\sum \mu(x_i) \cdot x_i}{\sum \mu(x_i)}, \quad (2)$$

где $\mu(x_i)$ – степень принадлежности функции принадлежности, x_i – значение.

Интерпретация весов, следующая:

$w_{ij} > 0$ – положительное влияние (усиление C_j при росте C_i);

$w_{ij} < 0$ – отрицательное влияние (ослабление C_j при росте C_i);

$w_{ij} = 0$ – отсутствие прямого влияния (C_i на фактор C_j).

Для исключения противоречий при передаче влияний между факторами используется механизм согласованности. Коэффициент совместимости $c_{ij} \in [0, 1]$, рассчитываемый по формуле

$$c_{ij} = \frac{v_{ij}^+}{v_{ij}^+ + v_{ij}^-}, \quad (3)$$

отражает уровень согласия между влияющим и принимающим факторами. Здесь v_{ij}^+ и v_{ij}^- – положительные и отрицательные компоненты консонанса, отражающие степень согласованности C_i на C_j .

Собственно балансировка выполняется с помощью медианного оператора [2]:

$$C_j = \text{med}(C_i, w_{ij}; c_{ij}), \quad (4)$$

где влияние передается от C_i на C_j , w_{ij} – вес, а c_{ij} – коэффициент совместимости, определяемый по формуле (3).

Значение же концепта C_j , на которое передаются непосредственные влияния с нескольких концептов $C_i, i = 1, \dots, N$, зависит от порядка просмотра степеней совместимости c_{ij} , обуславливающего очередность их аккумуляции, например, от меньших к большим.

Такой механизм обеспечивает фильтрацию несогласованных воздействий и гарантирует согласованность структуры факторов при распространении влияний.

Регрессия

Построение многопараметрической регрессионной модели включает серию этапов, направленных на выявление зависимости между прогнозируемой переменной Y и множеством факторов $X_1, X_2, \dots, X_N$. Общая форма линейной модели имеет вид [4]:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_N X_N + \varepsilon, \quad (5)$$

где $\beta_0, \beta_1, \beta_2, \dots, \beta_N$ – коэффициенты регрессии, подлежащие определению на основе данных, ε – случайная ошибка, предполагаемая нормально распределенной.

Первоначально формируется репрезентативная обучающая выборка, проводится обработка пропусков (средними, медианой, kNN, GMM и др.), проверяется применимость регрессии: отсутствуют мультиколлинеарность и вырожденность матрицы X . Оценка параметров производится методом наименьших квадратов (МНК):

$$\hat{\Theta}_{\text{МНК}} = (X^T X)^{-1} X^T Y, \quad (6)$$

что позволяет минимизировать сумму квадратов отклонений между предсказанными и фактическими значениями.

Для повышения устойчивости моделей применяются взвешенные (WLS), обобщенные (GLS), робастные (LAD) модификации, а также регуляризация (Ridge, Lasso) и жадные методы отбора признаков (Forward/Backward Selection) [6]. Эти подходы уменьшают дисперсию оценок при наличии мультиколлинеарности и улучшают обобщающую способность моделей. Качество модели оценивается по остаткам, которые должны соответствовать условиям нормальности, гомоскедастичности и отсутствия автокорреляции (например, по критерию Дарбина – Уотсона). Дополнительно рассчитываются метрики: R^2 , $Adjusted R^2$, MSE , MAE , $RMSE$ [8].

Результаты

Проведем сравнительный анализ классических ансамблевых и нейросетевых методов и предложенной гибридной подходе, основанной на НКК в сочетании с многофакторной регрессией для прогнозирования объема поставок газа в Китай ежемесячно на 2023 г. (рис. 1 и табл. 1).

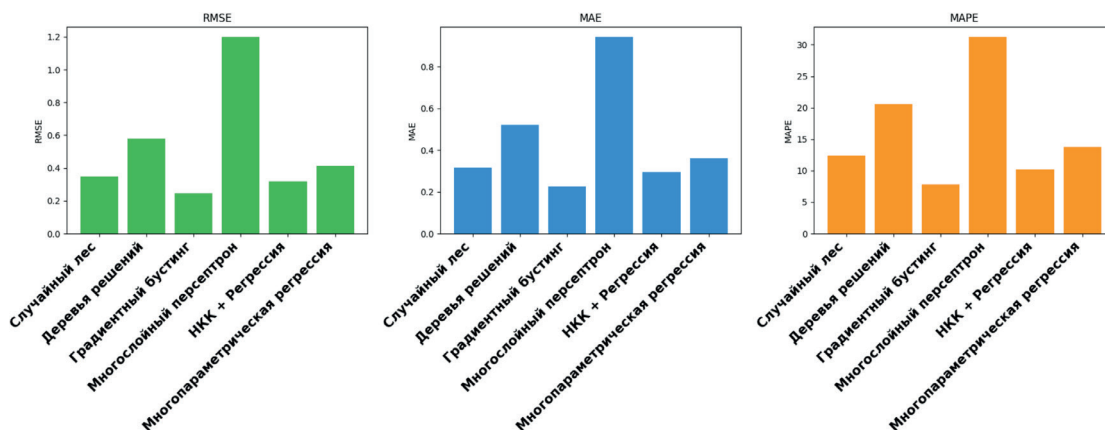


Рис. 1. Сравнительный анализ значений метрик классических ансамблевых подходов, нейросетевых методов нечеткой когнитивной карты и регрессионного анализа

Таблица 1. Сравнение классических ансамблевых подходов, нечеткой когнитивной карты и регрессионного анализа

Модель	RMSE	MAE	MAPE (%)
Случайный лес	0.351	0.317	12.35
Деревья решений	0.577	0.522	20.52
Градиентный бустинг	0.246	0.226	7.75
Многослойный персептрон (MLP)	1.200	0.943	31.24
НKK и регрессия	0.318	0.297	10.21
Многофакторная регрессия	0.4123	0.3625	13.80

Для проведения прогноза набор данных содержит 48 наблюдений – ежемесячные значения с января 2019 по декабрь 2022 г. В качестве прогнозируемой целевой переменной выступал объем совокупных поставок газа из РФ в Китай.

В качестве факторов, потенциально влияющих на динамику объема, были выбраны как внутренние, так и внешние индикаторы, способные отражать экономические, энергетические и политические условия.

Для всех переменных была предварительно проведена нормализация. Для моделирования использовались как классические методы машинного обучения, так и предложенная гибридная модель на основе нечеткой когнитивной карты и регрессии.

Исследование проводилось в специальной среде статистического анализа данных R версии 4.3.3, с использованием следующих пакетов:

randomForest – для случайного леса (500 деревьев, mtry = 3);
 rpart – для дерева решений;
 gbm – для градиентного бустинга (n.trees = 100, depth = 3, shrinkage = 0.05);
 nnet – для MLP с одним скрытым слоем (5 нейронов);
 lm (встроенная функция в R) – для многофакторной регрессии.

Погрешность вычислений в табл. 1 составляет в среднем $\pm 5\%$ при уровне доверия 95 %, что обусловлено вариативностью оценок качества моделей в рамках процедур перекрестной валидации.

Согласно экспериментальным результатам (см. рис. 1 и табл. 1), модель НKK продемонстрировала значение средней абсолютной процентной ошибки (MAPE), сопоставимое с градиентным бустингом (10.2 % против 7.75 %) на уровне погрешности вычисления, а также существенно более низкие значения по сравнению с деревьями решений (20.5 %) и многослойным персептроном (31.2 %). Аналогичные выводы можно сделать на основании сравнимых значений средней абсолютной ошибки (MAE) и корня средней квадратичной ошибки (RMSE), что подтверждает высокую точность гибридной модели при общем улучшении интерпретируемости.

Обсуждение результатов

Полученные экспериментальные результаты показывают, что предложенная гибридная модель (НKK + регрессионный анализ) обеспечивает более высокую точность прогнозирования по сравнению с традиционными методами машинного обучения. В частности, гибридный подход показывает конкурентоспособные результаты при сравнении с ансамблевыми моделями (случайный лес, градиентный бустинг) и значительно превосходит отдельные методы, такие как деревья решений, многослойный персептрон и многофакторная регрессия. Например, другой современный подход сочетает НKK первого уровня с нейросетевыми моделями (MLP/LSTM) на втором уровне для анализа сложных многомерных временных рядов для задачи прогнозирования энергопотребления [9]. В этом подходе данные группируются в кластеры, и на каждом кластере строятся нейросети. Результаты данного исследования показывают, что лучшие LSTM-модели второго уровня дают очень низкую ошибку: например, для одного кластера получены $MSE = 0.0013$, $MAE = 0.0277$. Авторы подчеркивают, что предложенная вложенная структура обеспечила более низкие MAE и MSE по сравнению с классической НKK и обычными ANN [9]. Иными словами, гибрид НKK + MLP/LSTM продемонстрировал высокую точность прогнозирования энергопотребления, превосходя отдельные FCM-модели и классические сети по MSE и MAE. Таким

образом, экспериментальные выводы подтверждают, что гибридный подход обеспечивает значительное преимущество по метрикам качества прогноза.

Практическая значимость предлагаемого подхода особенно очевидна при прогнозировании объема поставок газа в Китай. Как отмечено в литературе, прогнозирование индикаторов спроса и предложения природного газа является важной опорой для принятия стратегических решений органами управления и энергетическими компаниями. В условиях устойчивого роста газового потребления в Китае и реализации масштабных энергетических программ точное прогнозирование объема потребления природного газа приобретает особую важность [10]. Гибридный метод объединяет количественные методы регрессионного анализа и экспертные знания, позволяет учесть широкий спектр влияющих факторов (экономический рост, демографию, сезонность потребления, политические и климатические риски и т.д.) и, следовательно, повышает обоснованность прогнозных оценок. Это делает подход особенно полезным для планирования инфраструктурных проектов и мониторинга динамики импорта сжиженного и трубопроводного газа. Метод имеет и ограничения. В НКК требуются ручное формализация концептов и задание весов влияния экспертами, что вносит субъективность и возможные погрешности в модель [11]. Неточность или неполнота экспертных оценок может существенно исказить результат прогноза. Для уменьшения влияния такой субъективности применяются статистические методы обработки экспертных данных (например, метод наименьших квадратов), однако полностью устранить эти погрешности сложно [12, 13]. Кроме того, настройка численных параметров модели (весов связей, пороговых функций активации и др.) часто требует итеративной калибровки или эвристических алгоритмов (например, генетической оптимизации), что усложняет применение метода при большом числе факторов и связей. Эти факторы ограничивают масштабируемость подхода и требуют щепетильности при выборе экспертов и валидации весов НКК.

Перспективными направлениями дальнейших исследований могут стать динамическая калибровка весовых связей и интеграция с другими ИИ-методами. Стоит исследовать методы адаптации весов НКК на основе данных. В литературе предложены различные алгоритмы обучения весов (от правил Хебба до эволюционных методов), позволяющие автоматически корректировать параметры НКК по результатам наблюдений [14]. Наконец, интересен путь комбинирования НКК с современными архитектурами глубокого обучения. Недавно были предложены гибридные структуры, объединяющие когнитивные сети (расширенные НКК) с нейросетевыми модулями (CNN, ESN, автоэнкодерами и др.), которые демонстрируют качество прогнозирования не хуже или лучше по сравнению с классическими глубокими нейросетями [14]. Анализ подобных гибридных архитектур может открыть новые возможности повышения точности и адаптивности прогнозов в слабо формализованных системах.

Финансирование

Авторы заявляют об отсутствии источников финансирования.

Конфликт интересов

Конфликт интересов отсутствует.

Вклад авторов

Р.М. Романов – выполнение экспериментальных работ, обработка результатов; проведение численных расчетов; выбор методов исследования; подготовка текста статьи.

А.И. Гусева – постановка задачи и целей исследования.

Список литературы

1. Кулинич А.А. Систематизация когнитивных карт и методов их анализ // Труды VII Международной конференции «Когнитивный анализ и управление развитием ситуаций (CASC'2007)». М.: ИПУ РАН, 2007. С. 50–56.
2. Борисов В.В., Федулов А.С. «Совместимые» нечеткие когнитивные карты // Системы компьютерной математики и их приложения, 2016. № 17. С. 41–43.

3. Axelrod R. Structure of decision: The cognitive maps of political elites. Princeton, NJ: Princeton University Press, 1976. 422 p.
4. Арутюнов А.Л., Иванюк В.А., Цвиркун А.Д. Разработка инструментальных средств прогнозирования в социально-экономических системах // Управление развитием крупномасштабных систем, 2015. С. 241–293.
5. Kosko B. Fuzzy cognitive maps // International journal of man-machine studies, 1986. V. 24. № 1. P. 65–75.
6. Eng K., Chen Y. Y., Kiang J. E. User's guide to the weighted-multiple-linear-regression program (WREG version 1.0) // US Geological Survey Techniques and Methods, 2009. book 4, A8, 21 p.
7. Романов Р.М., Гусева А.И. Прогнозирование с использованием нечетких когнитивных карт и экспертной информации // GSOM Economy and Management conference 2024. Conference book. St. Petersburg, SPSU, 2024. P. 399–406.
8. Бабешко Л.О., Орлова И.В. Эконометрика и эконометрическое моделирование в Excel и R: учебник. М.: ИНФРА-М, 2025. 300 с.
9. Poczetka K., Papageorgiou E.I. Energy use forecasting with the use of a nested structure based on fuzzy cognitive maps and artificial neural networks // Energies, 2022. V. 15. № 20. P. 7542.
10. De G., Gao W. Forecasting China's natural gas consumption based on AdaBoost-particle swarm optimization-extreme learning machine integrated learning method // Energies, 2018. V. 11. № 11. P. 2938.
11. Christodoulou P., Christoforou A., Andreou A.S. A hybrid prediction model integrating fuzzy cognitive Maps with Support vector machines // International Conference on Enterprise Information Systems. SCITEPRESS, 2017. V. 2. P. 554–564.
12. Ротштейн А.П. Нечеткие когнитивные карты в анализе надежности систем // Надежность, 2019. Т. 19. № 4. С. 24–31.
13. Подгорская С.В. и др. Построение нечетких когнитивных моделей социально-экономических систем на примере модели управления комплексным развитием сельских территорий // Бизнес-информатика, 2019. Т. 13. № 3. С. 7–19.
14. Karatzinis G.D. et al. Fuzzy cognitive networks in diverse applications using hybrid representative structures // International Journal of Fuzzy Systems, 2023. V. 25. № 7. P. 2534–2554.

Vestnik Natsional'nogo Issledovatel'skogo Yadernogo Universiteta «MIFI», 2025, vol. 14, no. 4, pp. 332–338

Hybrid approach based on cognitive mapping and regression analysis for forecasting in complex weakly formalized systems

R. M. Romanov[✉], A. I. Guseva[✉]

National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Moscow, 115409, Russia

[✉] Rodion.Romanov.M@yandex.com

[✉] aiguseva@mephi.ru

Received June 30, 2025; revised July 15, 2025; accepted July 22, 2025

The article presents an original hybrid approach that integrates fuzzy cognitive maps and regression analysis for forecasting in weakly formalized systems characterized by high levels of uncertainty and complex, unstructured interrelationships among variables. The core idea of the approach lies in utilizing expert judgments, expressed through linguistic variables and fuzzy numbers, to adapt the weighting coefficients of the regression model. The weights obtained from cognitive analysis are incorporated into the procedure of multiparametric linear regression using the weighted least squares method. This integration enhances both the accuracy of forecasts and the interpretability of the model. The results of an empirical study conducted in the R statistical environment demonstrate that the proposed approach outperforms (MAPE = 10.21%) not only classical single forecasting methods, such as linear regression (MAPE = 13.80%) and neural network models, e.g., the multilayer perceptron (MAPE = 31.24%), but also successfully competes with ensemble methods, including random forest (MAPE = 12.35%) and gradient boosting (MAPE = 7.75%). The proposed approach can be effectively applied to forecasting natural gas supply volumes to China, as well as to solving other tasks that require the integration of qualitative and quantitative data.

Keywords: fuzzy cognitive map; fuzzy numbers; regression analysis; forecasting; weakly formalized systems; expert judgment.

References

1. *Kulinich A.A.* Sistematizaciya kognitivnyh kart i metodov ih analiz [Systematization of cognitive karting and methods of analysis]. Trudy VII Mezhdunarodnoj konferencii «Kognitivnyj analiz i upravlenie razvitiem situacij (CASC'2007)» [Proceedings of the VII International Conference «Cognitive Analysis and Business Development Management (CASC'2007)»]. Moscow, IPU RAN Publ., 2007. Pp. 50–56 (in Russian).
2. *Borisov V.V., Fedulov A.S.* «Sovmestimye» nechetkie kognitivnye karty [«Compatible» fuzzy cognitive maps]. Sistemy komp'yuternoj matematiki i ih prilozheniya, 2016. No. 17. Pp. 41–43 (in Russian).
3. *Axelrod R.* Structure of decision: The cognitive maps of political elites. Princeton, NJ: Princeton University Press, 1976. 422 p.
4. *Arutyunov A.L., Ivanyuk V.A., Tsvirkun A.D.* Razrabotka instrumental'nyh sredstv prognozirovaniya v social'no-ekonomicheskikh sistemah [Development of forecasting tools in socio-economic systems]. Upravlenie razvitiem krupnomasshtabnyh sistem, 2015. Pp. 241–293 (in Russian).
5. *Kosko B.* Fuzzy cognitive maps. International Journal of Human-Machine Research, 1986. Vol. 24. No. 1. Pp. 65–75.
6. *Eng K., Chen Y.Y., Kiang J.E.* User's guide to the weighted-multiple-linear-regression program (WREG version 1.0). US Geological Survey Techniques and Methods, 2009. book 4, A8, 21 p.
7. *Romanov R.M., Guseva A.I.* Prognozirovanie s ispol'zovaniem nechotkih kognitivnyh kart i ekspertnoj informacii [Forecasting using fuzzy cognitive maps and expert information]. GSOM Economy and Management conference 2024. Conference book. St. Petersburg, SPSU Publ., 2024 Pp. 399–406 (in Russian).
8. *Babenko L.O., Orlova I.V.* Ekonometrika i ekonometricheskoe modelirovanie v Excel i R: uchebnik [Computer graphics and computer modeling in Excel and R: textbook]. Moscow, INFRA-M Publ., 2025. 300 p.
9. *Pocheta K., Papageorgiu E.I.* Forecasting energy consumption using a nested structure based on fuzzy cognitive maps and artificial neural networks. Energy, 2022. Vol. 15. No. 20. P. 7542.
10. *De G., Gao U.* Forecasting natural gas consumption in China based on the integrated learning method AdaBoost-particle swarm optimization-extreme learning machine. Energy, 2018. Vol. 11. No. 11. P. 2938.
11. *Christodoulou P., Christoforou A., Andreu A.S.* Hybrid forecasting model integrating fuzzy cognitive maps with support vector methods. International Conference on Corporate Information Systems. SCITEPRESS, 2017. Vol. 2. Pp. 554–564.
12. *Rotstein A.P.* Nechetkie kognitivnye karty v analize nadezhnosti sistem [Fuzzy cognitive maps in the analysis of system reliability]. Nadezhnost', 2019. Vol. 19. No. 4. Pp. 24–31 (in Russian).
13. *Podgorskaya S.V., et al.* Postroenie nechetkih kognitivnyh modelej social'no-ekonomicheskikh sistem na primere modeli upravleniya kompleksnym razvitiem sel'skih territorij [Building fuzzy cognitive models of socio-economic systems using the example of a management model for integrated rural development]. Biznes-informatika, 2019. Vol. 13. No. 3. Pp. 7–19 (in Russian).
14. *Karatsinis G.D. et al.* Fuzzy cognitive networks in various applications using hybrid representative structures. International Journal of Fuzzy Systems, 2023. Vol. 25. No. 7. Pp. 2534–2554.