

Моделирование функциональных связей региональных экономических систем по малым выборкам на основе байесовских интеллектуальных измерений

Р. А. Жуков¹ , С. В. Прокопчина² , М. А. Плинская¹ , М. А. Желунцова¹ 

¹Тульский филиал Финансового университета
при Правительстве Российской Федерации,
г. Тула, Россия

²Финансовый университет при Правительстве Российской Федерации,
г. Москва, Россия

 pluszh@mail.ru

Аннотация. Экономика субъектов Российской Федерации, представляющих собой социо-эколого-экономические системы, характеризуется быстро меняющимися условиями, влияющими на их функционирование. Следовательно, при моделировании недостаточно традиционных подходов на основе долгосрочных трендов и построения соответствующих моделей и требуется их развитие, в том числе за счет сочетания с другими подходами, учитывающими изменчивость условий. Кроме того, информационные базы из открытых источников обладают неопределенностью (в разных источниках для одного показателя могут быть отличные значения). Некоторые показатели постоянно корректируются: вносятся изменения в значения для предыдущих периодов. Периодов оценки недостаточно для методологически обоснованного, с точки зрения математической статистики, построения моделей. Результаты функционирования зависят от влияющих факторов так же, как и результаты измерений. Обозначенные проблемы приводят к необходимости прибегать к вероятностным оценкам и учитывать неполноту и нечеткость данных, а также получать выборки достаточного объема из данных малой мощности. Целью исследования является моделирование связей между объемами валового внутреннего продукта и факторами, характеризующими функционирование экономической подсистемы регионов, в условиях неопределенности и ограниченности данных в заданный период времени. Рабочая гипотеза исследования: возможность методологически обоснованного построения эконометрических моделей по данным для отдельного субъекта Российской Федерации в определенный период времени. В основу исследования легли методы эконометрического моделирования и методология байесовских интеллектуальных измерений, а также методика формирования достаточного объема данных по малым выборкам. На примере Тульской области по данным за 2022 г. были построены степенные мультипликативные модели для объема ВРП (14 разделов по ОКВЭД), что позволило провести апробацию представленной методологии, базирующейся на сочетании эконометрического и байесовского подходов к моделированию функционирования сложных систем, с выводами теоретической и практической направленности на устойчивое развитие региона.

Ключевые слова: валовой региональный продукт; производственная функция; социо-эколого-экономическая система; модель, байесовские интеллектуальные измерения; малая выборка.

1. Введение

Рассматривая экономику субъектов Российской Федерации, представляющих собой социо-эколого-экономические системы (СЭЭС), исследователи оценивают их функционирование посредством макроэкономических показателей, а экономическую подсистему — объемом валового регионального продукта (ВРП) с детализацией по отдельным видам экономической деятельности (ОКВЭД). Зависимость между объемом ВРП и основными факторами производства традиционно моделируется посредством производственных функций (ПФ), чаще всего нелинейных функциональных форм. Каждый из субъектов развивается по собственной траектории, причем в каждый период времени условия функционирования (факторы) отличны от других временных интервалов, в которых проводились измерения, что подразумевает и изменение параметров моделей — коэффициентов чувствительности, учитывающих характер влияния таких факторов на результаты деятельности СЭЭС. Тогда при построении модели возникают дополнительно две проблемы: 1) ограниченность выборки; 2) неопределенность данных, связанная в том числе с их достоверностью.

В этом аспекте представляется актуальным для конструирования моделей использовать традиционные методы построения эконометрических моделей в сочетании с методологией байесовских интеллектуальных измерений (БИИ), хорошо работающей в условиях неопределенности, и методикой формирования совокупности данных достаточного объема, сгенерированной по малой выборке с известным типом распределения.

Исследовательский вопрос: возможно ли методологически обоснованно построить модели связи между ВРП по ОКВЭД и влияющими факторами по малым выборкам посредством инте-

грации эконометрического моделирования и методологии байесовских интеллектуальных измерений?

Цель исследования — моделирование связей между объемами валового внутреннего продукта и факторами, характеризующими функционирование экономической подсистемы регионов, в условиях неопределенности и ограниченности данных в заданный период времени.

Объект исследования — экономическая подсистема социо-эколого-экономической системы с однотипными результативными признаками (Тульская область).

Предмет исследования — особенности построения моделей в условиях неполноты и нечеткости данных.

Рабочая гипотеза исследования — возможность методологически обоснованного построения эконометрических моделей по данным для отдельного субъекта Российской Федерации в определенный период времени.

Структура статьи включает введение, степень проработанности проблемы, раздел методологии и информационной базы исследования, результаты исследования, обсуждение результатов и заключение.

2. Степень проработанности проблемы

Моделируя экономику регионов, представляющую собой экономическую подсистему СЭЭС, обычно строят эконометрические модели, отображающие связь между результатами функционирования элементов СЭЭС и условиями (факторами) их деятельности. Элементы СЭЭС — совокупность хозяйствующих субъектов на территории региона (резиденты), функционирующих в соответствии с СНС.

Модели позволяют изучать поведение элементов, подсистем и системы

в целом в зависимости от изменяющихся условий, проводить оценку и анализ с целью разработки управленческих решений, направленных на обеспечение сбалансированного и устойчивого функционирования и развития СЭЭС.

Значения результативных признаков СЭЭС дают возможность проводить их сравнительную оценку с ожидаемыми (плановыми, нормативными) значениями, установленными органами управления либо вычисленными по специальным методикам, в том числе при построении соответствующих моделей.

Функционирование сложных систем, какими являются СЭЭС, сопряжено с неопределенностью внешних условий, а также воздействием факторов, которые неидентифицируемы и не могут быть явно включены модель. Поэтому при изучении СЭЭС, в том числе в экономических исследованиях, помимо использования традиционных эконометрических методов, прибегают к специальным методам, работающим в условиях нечеткости и неполноты данных.

Akperov & Kurianov [1], используя логику Zadeh [2], предложили оценивать региональный экономический потенциал (низкий, средний, высокий) посредством достаточности ресурсов (недостаточный, достаточный объем и избыточный), возможности привлечения дополнительных ресурсов (низкая, средняя, высокая) и эффективности управления ресурсами (низкая, средняя, высокая). Такая оценка осуществляется при использовании набора параметров, характеризующих экономику региона и сводящихся к агрегированным трехкомпонентным лингвистическим факторам с применением элементов нечеткой логики и нейронных сетей.

Zolfani et al. [3] применили нечеткие множества, включающие социально-экономические показатели,

для оценки потенциала стран Южной и Восточной Европы и их взаимодействия. Результатом исследования стало обоснование эффективности предложенной модели и факт наличия соперничества между соответствующими регионами.

Huang et al. [4] предложили модель прогнозирования 3 продуктового промышленного сектора китайской экономики посредством формирования нечетких временных рядов первого порядка, что позволило осуществить оценку будущих значений переменных в условиях нечеткости данных. Результаты показали, что многомерная нечеткая модель обладает большей точностью по сравнению с традиционными ARIMA моделями и моделью, построенной на основе нейронной сети обратного распространения (BPNN). Такой подход в какой-то мере коррелирует с эконометрическим моделированием, однако переход от нечетких к классическим моделям параметрической формы не был исследован.

Ouifak & Idri [5] на основании глубокого анализа 48 исследований, посвященных методам моделирования на нейросетевых системах различного типа, доказали, что модели, построенные на базе нейро-нечетких ансамблей, поддаются лучшей интерпретации по сравнению с остальными системами. Это еще раз доказывает необходимость сочетания эконометрических и нечетких подходов, в том числе байесовских интеллектуальных технологий при изучении сложных систем, какими являются социо-эколого-экономические системы.

Huning & Wahl [6] применили нечеткую пространственную регрессию степенной мультипликативной формы для оценки экономики Западной Германии, что позволило им сделать обоснованный вывод о влиянии исторического развития территорий, их экономической специализации и традиций на по-

вышение доходов, плотность населения и уровень индустриализации в соответствующих регионах. Данные были взяты, начиная с XIX по XXI век. Однако если применить такой подход для России, то очевидно, что обоснованные выводы сделать будет трудно, в силу ограниченности данных. Это приводит к необходимости искусственно формировать выборку достаточной для обоснованности применения эконометрических моделей мощности.

Kutty et al. [7] в виду множественности факторов, влияющих на устойчивое развитие «умных» городов Европы, успешно использовали методы нечеткой оценки для их рейтингования по набору индексов.

Ruben et al. [8] на основе построения модели структурных взаимосвязей с нечеткими входными данными выявили барьеры внедрения технологий индустрии 4.0.

Особое место в экономических исследованиях занимает байесовский подход, позволяющий нечеткость интерпретировать вероятностными оценками.

Ziesmer et al. [9] удалось скомбинировать байесовский подход и CGE-модель для оценки устойчивого развития экономики. В качестве основных параметрических моделей авторы рассмотрели полиномиальные и линейные формы, встречавшиеся в работах Forrester et al. [10] и Myers [11].

Iacopini et al. [12] на основе одномерной байесовской квантильной регрессии осуществили прогноз условных значений ВВП США, которые строго связаны с количественной оценкой стоимости, подверженной риску ожидаемого дефицита и разницы между процентиями прогнозов реального ВВП, в том числе на недостающих низкочастотных данных, восстановленных по данным более высокой частоты в предположении о нормальности их распределений.

Иными словами, в этой ситуации байесовский подход был применен в условиях неполноты датасетов, что характерно и для российских данных, имеющихся в открытых источниках.

D'Andrea [13] для сглаживания неопределенности во влияющих на экономический рост факторах применила усреднение по байесовской модели, показав, что для данной выборки (по 19 европейским странам) модель является лучшим статистическим инструментом для оценки переменных, включаемых в регрессию роста. В исследовании применялось как эконометрическое моделирование (линейная регрессия), так и байесовский подход с оценкой вероятностей (апостериорных и априорных вероятностей) моделей из выбранного набора на базе метода Монте-Карло. Авторами выборка генерировалась на основании гипотез с заданными параметрами распределения факторов (в данном случае нормальными). Однако исходные, фактические данные не «размывались» и не были представлены в виде нечеткого числа, идентифицирующего неопределенность имеющейся информации.

Duan et al. [14] использовали байесовский стохастический поиск и метод Монте-Карло, примененный D'Andrea, для исследования влияния набора социально-экономических факторов при анализе цен на жилье в Китае. В качестве типа распределения было выбрано распределение Бернулли.

Canoz & Kalkavan [15] провели аналогичное исследование для рынка жилья Стамбула.

Salles et al. [16] применили байесовский подход для определения плотности вероятностей исходных данных с целью анализа экономического развития лесной промышленности в Бразилии.

Debnath et al. [17] ориентировались на аналогичную [3] трехкомпонентную структуру оценки развития мебельной

промышленности в Бангладеш на основе BWM (байесовский метод анализа наилучшего и наихудшего) и концепции устойчивого бережливого производства.

Klein et al. [18] использовали байесовский подход при расчете неэффективности, что позволило авторам скорректировать оценки неэффективности функционирования фермерских хозяйств в сочетании с моделью стохастической границы.

Szafranek et al. [19] на основе байесовской структурной модели VAR (BSVAR) смогли описать эволюцию восьми ключевых макроэкономических переменных с набором нулевых и знаковых ограничений. Это дало возможность корректно проанализировать внутренние экономические шоки, в том числе связанные со всплеском инфляции в Польше. Применение BSVAR позволило избежать излишней параметризации уравнений за счет использования априорного распределения коэффициентов. В исследовании, как и в большинстве аналогичных работ, распределение предполагалось нормальным.

Mozdzen et al. [20] комбинацией байесовского моделирования и кластеризации попытались устранить проблемы обоснованности применения теоретической базы и возможности использования стандартных вычислительных методов при работе с пространственно-временными данными. Как утверждают авторы, их модель дает более точные оценки, а дополнительная информация, полученная в результате кластеризации, позволяет расширить экономическую интерпретацию уровня безработицы в итальянских провинциях.

Elhorst [21] — яркий представитель школы пространственной эконометрики — анализирует байесовскую оценку при работе с динамической пространственной эконометрической моделью об-

щей вложенности для пространственных панелей с общими факторами. То есть байесовский подход он использует как средство построения эконометрической модели и, в отличие от Mozdzen, подступает с другой стороны к решению проблемы неопределенности данных.

Noia et al. [22] расширили класс байесовских моделей скрытых переменных за счет учета данных о размере и форме откликов гауссовского распределения.

Aivazian [23] достаточно полно описал применение байесовского подхода в эконометрических исследованиях. К сожалению, большая часть исследований, связанная с развитием байесовского подхода и нечеткого моделирования, относится к техническим и информационным системам, в том числе включая использование искусственного интеллекта, что отражено, например, в работах Potapenko et al. [24] и Averkin & Yarushev [25].

Бормотов и др. [26] использовали методологию байесовских интеллектуальных измерений (далее — БИИ) для оценки ценностного мира российской молодежи, продемонстрировав возможность применения БИИ для систем различного типа на основе развиваемой Прокопчиной методологии байесовских интеллектуальных технологий (далее — БИТ) в рамках регуляризирующего байесовского подхода (далее — РБП) [27] и методики определения типа распределения по малым выборкам [28].

При описании сложных систем исследователи часто сталкиваются с проблемой выбора функциональных форм моделей.

Zhong et al. [29] применили линейные модели для формирования интегральных индикаторов, сформированных на основе влияющих факторов посредством унифицированного преобразования, использованного как один из этапов формирования системы универсальных

индикаторов оценки функционирования социо-эколого-экономических систем в работе Zhukov et al. [30].

Andrews et al. [31] использовали инструментальные переменные для построения регрессионных моделей.

Одним из базовых показателей оценки экономики субъектов Российской Федерации является объем валового регионального продукта, в том числе оцениваемого по видам экономической деятельности. Его можно описать с помощью модели, имеющей вид степенной мультипликативной функции, аналогичной форме Кобба — Дугласа, как это представлено в работе Макарова и др. [32].

Жуков и др. [33] показали, что в случае статистической адекватности моделей они являются инвариантными относительно изучаемого процесса и выбор можно осуществлять по критерию приоритетности того или иного направления развития субъектов экономики.

Как видно, все больше исследователей применяют нечеткие модели, модели, построенные в рамках байесовского подхода, а также комбинируют их с классическими моделями, задавая мейнстрим в экономических и других исследованиях на мировом уровне. Это означает обоснованность выбора авторов данной статьи в сторону сочетания соответствующих моделей, в том числе конструирование эконометрических моделей на основе байесовского подхода в условиях неполноты и нечеткости данных. Кроме того, байесовский подход используют как для построения моделей, которые должны адекватно описывать функционирование систем в условиях неопределенности (в том числе для отдельных отраслей промышленности), так и при решении частных задач, например для генерации выборки, обеспечивающей полноту данных.

Отметим, что наряду с выявленными проблемами выбора функциональных форм моделей, методологии оценки

параметров в условиях неполноты и нечеткости данных существует проблема, связанная с тем, что в каждый период времени субъекты экономики функционируют в уникальных для себя условиях даже при неизменных пространственных характеристиках, а измерения являются единичными. Следовательно, модели должны строиться на основании таких измерений, что приводит к существенному ограничению объема имеющейся в распоряжении исследователя выборки и к необходимости ее расширения.

Это обуславливает необходимость развития методов моделирования, которые могли бы сочетать в себе как традиционные и часто используемые в практике эконометрического моделирования, так и вероятностные подходы, в том числе на основе БИТ с возможностью получения датасетов достаточной мощности.

3. Методология и информационная база исследования

3.1. Методология исследования

Обозначим за k номер субъекта Российской Федерации ($k = 1, \dots, K$), а i — номер элемента экономической подсистемы региона ($i = 1, \dots, J$). В данном случае это совокупность институциональных единиц, осуществляющих деятельность i вида по ОКВЭД. Каждый элемент характеризуется результирующим признаком y_{ik} и набором факторов (факторных признаков) x_{ikj} ($j = 1, \dots, J$), последние из которых представляют собой условия функционирования ik элемента.

Эконометрическое соотношение между результатами и факторами определим как:

$$y_i = \hat{y}_i + \varepsilon_i. \quad (1)$$

Здесь $\hat{y}_{i,k} = f_{i,k}(x_{i,k,j})$ — производственная функция; ε_i — нормальная случайная компонента (в простейшем случае $N(0, \sigma)$).

Пусть в один тот же период времени проведены измерения, на основании которых получены соответствующие значения результативных y_p , k и факторных признаков x_p , k , j . Будем полагать, что измерения проведены в условиях неопределенности, характеризуемые наличием множества метрологических требований $\{MX(i)\}$, множеств априорной информации A , ограничений и допущений O .

Тогда целесообразно результаты измерений представить в виде нечеткого числа, определяемого набором пар чисел $(h_{(\cdot),l}; p_{(\cdot),l})$, где $h_{(\cdot),l}$ — значение признака $(\cdot) = y_{i,k}, x_{i,k,j}$, соответствующее положению репера l на шкале с диапазоном динамических ограничений $(l = 1, \dots, L_p)$; $p_{(\cdot),l}$ — вероятность того, что значение $(\cdot) = h_{(\cdot),l}$. Число реперов L_r на числовой шкале задается пользователем. В случае использования лингвистической шкалы число реперов соответствует числу классов (9 классов = предельно ниже нормы, ..., норма, ..., предельно выше нормы) [27].

Интервал шкалы динамическими ограничениями (далее — ШДО) может быть задан как:

$$[(\cdot)_{\min} - \sigma_{(\cdot)}; (\cdot)_{\max} + \sigma_{(\cdot)}]. \quad (2)$$

Здесь $(\cdot)_{\min}^{\max}$ — минимальное и максимальное значения среди $y_{k,i}$ или $x_{k,i,j}$, которые могут быть определены или экспертами, или на основании множества измерений, проведенных ранее (в другие периоды времени $t = 1, \dots, T$); $\sigma_{(\cdot)}$ — корректировка интервала, рассчитываемая по формуле:

$$\sigma_{(\cdot)} = \left(\frac{(\cdot)_{\max} + (\cdot)_{\min}}{2} \right)^a / 3, \quad (3)$$

где a — обозначение априорной шкалы.

Отметим, что расширение шкалы обусловлено тем, что при появлении новых данных может возникнуть необходимость расширения диапазона, опреде-

ляемого формулой (2), что практически всегда происходит при изучении экономических явлений. Кроме того, при прогнозировании результатов и факторов их значения могут выйти за границы интервала без корректирующего параметра.

На практике будущие значения признаков обычно не выходят за пределы 3σ . При этом если факторы демонстрируют восходящую или нисходящую без аномалий (резких «всплесков») динамику, то для того, чтобы их значения находились в заданном априорном диапазоне, оказывается достаточным расширить его на величину σ , что и отражает формула (3).

Для каждого признака определяется норма — ожидаемое (плановое, нормативное) значение признака. Она задается априорно: экспертами, в соответствии со стандартами или в простейшем случае как среднее известных значений $(\cdot)$.

В более сложном случае для построения нормативов необходимо использовать преобразование мягких норм к средневзвешенным значениям для конкретного периода. Мягкие нормы для фактора зависят от условий (влияющих факторов), в которых осуществляется измерение. Так, например, норма объема ВРП может зависеть от стоимости основных фондов и численности занятых, если опираться на неоклассическую модель Кобба — Дугласа, представленную в виде производственной функции [33]. В данном исследовании будем ограничиваться нормой, вычисленной как среднее.

Длина интервалов для числовой шкалы определяется по формуле:

$$l_{(\cdot),h_l} = \left(\frac{(\cdot)_{\max} - (\cdot)_{\min} + 2 \cdot \sigma_{(\cdot)}}{L_r} \right). \quad (4)$$

Тогда:

$$h_{(\cdot),l} \in [(\cdot)_{\min} - \sigma_{(\cdot)} + l_{(\cdot),h_l} \cdot (l-1); (\cdot)_{\min} - \sigma_{(\cdot)} + l_{(\cdot),h_l} \cdot l]. \quad (5)$$

В случае, если $l = Lr$, то

$$h_{(\cdot),Lr} \in \left[\binom{(\cdot)}{\min} - \sigma_{(\cdot)} + l_{(\cdot),h_l} \cdot (L_r - 1); \binom{(\cdot)}{\max} + \sigma_{(\cdot)} \right].$$

В случае лингвистической шкалы рассчитываются отдельно длины интервалов, которые выше нормы и ниже нормы по соотношениям:

$$l_{(\cdot),h_l, \text{left}} = \left(\frac{\binom{(\cdot)}{\max} - \binom{(\cdot)}{\min}}{5} \right),$$

$$l_{(\cdot),h_l, \text{right}} = \left(\frac{\binom{(\cdot)}{\max} - \binom{(\cdot)}{\min}}{5} \right). \quad (6)$$

Вероятность априорная для $h_{(\cdot),l}$ в простейшем случае задается равновероятной как для числовой, так и для лингвистической шкалы:

$$P^a \left(H_{(\cdot),l} = p_{(\cdot),l} = \frac{1}{L} \right), \quad (7)$$

где $L = Lr$ для числовой шкалы и 9 для лингвистической шкалы; a — обозначение априорной вероятности; $H_{(\cdot),l}$ — набор гипотез или альтернативных решений (значение $(\cdot) = h_{(\cdot),l}$). Число реперов числовой шкалы может совпадать с числом классов лингвистической шкалы (традиционный вариант задания числа реперов).

В результате такого представления в соответствии с формулами (4)–(7), вероятность того, что значение случайной величины равно значению репера или попадает в один из девяти классов, будет равновероятным, а соответствующее распределение окажется равномерным.

В более сложном случае априорные вероятности можно задать в соответствии с известным (если он известен для данного фактора) распределением в соответствии с классическими типами распределений, разделенных на 25 классов, как показано в работе Прокопчиной [27].

Апостериорную вероятность факторов можно определить аналогичным байесовской формуле соотношением [27]:

$$p_{(\cdot),l}^{ap} = P^{ap} \left(H_{(\cdot),l} | S \right) = P^a \left(H_{(\cdot),l} \right) \times \times P \left(S | H_{(\cdot),l} \right) / \sum_{j=1}^J P^a \left(H_{(\cdot),l} \right) P \left(S | H_{(\cdot),l} \right), \quad (8)$$

где ap — апостериорная вероятность; S — событие совместного появления оценок $\tilde{h}_{(\cdot),l}$ для $(\cdot)$.

Следующая формула дает возможность вычислить условную вероятность (условие появления S события):

$$P \left(S | H_{(\cdot),l} \right) = \int_{-\infty}^{(\cdot)} f \left(x, h_{(\cdot),l}, \sigma_{(\cdot),norm} \right) dx, \quad (9)$$

где $f \left(x, h_{(\cdot),l}, \sigma_{(\cdot),norm} \right)$ — плотность нормального распределения с математическим ожиданием $h_{(\cdot),l}$ и дисперсией $(\sigma_{(\cdot),norm})^2 = (l_{(\cdot),h_l})^2$.

В результате таких действий будем иметь набор интервалов, реперов и вероятностей, соответствующих тому, что случайная величина (в данном случае результативный или факторные признаки) будет находиться в заданном интервале, а объем выборки будет равен числу интервалов (реперов). Такое представление формируется для каждого фактора и для каждого периода времени, в котором значение исходного измерения известно. В более сложном случае можно отойти от нормального распределения и использовать другие функции распределения вероятностей (например, Вейбулла, гамма и др.) в случае уверенности в обоснованности такого выбора.

Для получения выборки достаточной мощности (число наблюдений (измерений) должно быть порядка $N = 7000$ [27]) необходимо на каждом из интервалов сгенерировать $n_{(\cdot),l} = \left[N \cdot p_{(\cdot),l}^{ap} \right]$ значений (квадратные скобки означают, что от выражения нужно взять целую часть) с заданным распределением. Отметим, что в качестве базовой функции распре-

деления рекомендуется использовать треугольное распределение [28].

В результате будет получена выборка требуемого объема для $(\cdot)$ признаков, позволяющая методологически обоснованно переходить к построению эконометрических моделей (1). То есть для единичного (уникального) измерения оказывается возможным получить выборку достаточной мощности.

В дальнейшем сгенерированную выборку для случайной величины можно оценить на ее соответствие определенному типу распределения и его классу (25 классов), отображаемых на плоскости моментов Пирсона с разделяющимися границами [27]. Обоснованность такого подхода подтверждена его апробацией в работе Жуков и др. [28].

3.2. Последовательность реализации методологии

Этапы предлагаемой методологии можно описать следующим образом.

1. Сбор и обработка данных для выбранных элементов СЭС субъектов Российской Федерации, характеризующихся однотипными результативными и факторными признаками, выбранными в качестве переменных производственных функций. Стоимостные характеристики могут быть приведены к сопоставимому виду. Реализация данного этапа опирается на подтвержденную гипотезу об инвариантности процесса относительно моделей, скорректированных на инфляцию, и отраженную в работе Жукова и др. [33].

2. Представление данных в виде нечетких чисел (для каждого из признаков), включая процедуру задания априорных и апостериорных шкал с динамическими ограничениями и нормами для каждого из показателей. На данном этапе рассчитываются пары (значение, вероятность) для каждого признака (результативный или факторный) для конкретного пери-

ода времени. Наиболее интересна оценка последнего известного периода. Такое представление дает возможность из одного значения — результата измерения получить множества значений (для числового или лингвистического отображения), мощности которых соответствуют числу реперов или числу классов в рамках РБП.

3. Формирование выборки достаточного объема для выбранных факторных и результативных признаков для периода t . Данный этап опирается на методологию идентификации типа распределения по малым выборкам, примененную в авторском исследовании [28].

4. Построение моделей в виде эконометрических моделей, содержащих производственные функции (соотношение результатов и факторов) методами OLS или MLE. Этап включает классическое построение эконометрических моделей, где в качестве аппроксимирующей функции выступает производственная функция, например мультипликативная.

5. Статистическая оценка моделей. Предполагает расчет статистических характеристик с целью проверки адекватности и точности моделей, в том числе с использованием стандартных тестов, например критерия Фишера, Стьюдента и других.

В качестве базовых инструментальных средств применены программные платформы «ЭФРА» [34] и «Инфоаналитик 2.0», разработанной на основе более ранней версии [35].

«Инфоаналитик 2.0», с помощью которого реализуются этапы 2 и 3, предусматривает автоматический или ручной ввод данных после их предварительной обработки на этапе 1 в сконструированную иерархическую информационную модель. Иерархическая информационная модель отображается в виде структуры объектов, идентифицируемых на основе интегральных и частных факторов (атрибутов модели).

Следующий функционал комплекса содержит форму генерации априорных и апостериорных шкал в соответствии с формулами (3)–(9), где задается нижняя и верхняя границы, а также норма для каждого факторов. В случае интегрального фактора рекомендуется использовать шкалу от 0 до 10 и нормой 5.

Следующий блок программной платформы вычисляет динамику, интерпретацию динамики (выявление тенденций изменения) факторов, а также рассчитывает метрологические характеристики изменения факторов, в том числе точность, надежность, риск, энтропию априорной и апостериорной информации. Кроме того, предусмотрен функционал прогноза факторов с заданием порядка производной, числа точек (число периодов обучающей выборки) и шага прогноза, что дает возможность аналитику варьировать параметры прогнозных моделей.

Следующий блок — блок рекомендаций, включающий авторекомендации, позволяет на основании нечеткого логического вывода дать предложения по усилению или ослаблению влияющих на результативный или интегральный признак факторов, включенных в модель. Для генерации выборки используется дополнительный модуль байесовской математической статистики, куда подгружаются данные из числовой или лингвистической шкал, содержащие значения реперов, термы, соответствующие им вероятности, нижние и верхние границы интервалов в формате csv. Задается объем выборки. По данным на каждом из интервалов шкалы генерируется выборка с заданным числом значений (можно задать априорный тип распределения). По встроенному алгоритму определяется тип распределения полученной выборки. Результат генерации представляется в виде файла excel или json.

Созданные файлы формата excel загружаются в «ЭФРА» через соответствующий интерфейс программного комплекса и осуществляется четвертый и пятый этапы методологии. После получения данных предоставляется возможность строить эконометрические модели и прогноз. Из формы списка признаков выбирается результативный признак и передается в соответствующий виджет (y_i). В виджет факторов переносятся влияющие факторы (условия). Кроме того, может быть задан период оценки (период обучающей выборки). Для сгенерированных данных эту процедуру делать нецелесообразно. Выбирается функциональная форма (линейная, логарифмическая, экспоненциальная или мультипликативная). Далее при нажатии на кнопки осуществляются:

- 1) построение модели;
- 2) оценка качества и параметров модели (в том числе для нелинейных моделей) с помощью критериев Фишера с расчетом коэффициента детерминации и (t -статистика) Стьюдента;
- 3) загрузка дополнительных тестов ((по коэффициентам корреляции Спирмена) гетероскедастичность и (Фаррара — Глобера) мультиколлинеарность);
- 4) прогнозирование (на основании ожидаемых значения влияющих факторов, из файла или в ручном режиме);
- 5) формирование результатов в виде настраиваемого отчета, в котором предусмотрены в том числе вывод линеаризованных, нелинеаризованных и стандартизованных моделей; итоги тестирования, а также значения коэффициентов эластичности. Это дает возможность получить соответствующие модели с их статистической оценкой в удобной для аналитика форме.

Таким образом реализуются этапы представленной методологии исследования на основе «Инфоаналитик 2.0»

и «ЭФРА». В дальнейшем эти модели выступают в качестве нормативных моделей или моделей функционирования СЭЭС с целью определения результативных показателей (интегральных и частных) и построения оптимизационных моделей, дающих возможность получить количественно выраженные значения изменения влияющих факторов, улучшающих результаты функционирования СЭЭС, ее подсистем и элементов. Для этого в «ЭФРА» реализованы соответствующие программные модули с интерфейсной частью, где можно настраивать набор анализируемых показателей, а также выбирать класс оптимизационной модели и систему ограничений.

3.3. Информационная база

Информационную базу составили ежегодные сборники Росстата показателей социально-экономического развития регионов (2007–2022)¹, датасеты из единой межведомственной информационно-статистической базы ЕМИСС²; таблицы уровней инфляции (по данным Росстат)³. Использованы данные об объеме ВРП в Тульской области (2022)⁴ и структуре ВРП в этот же период⁵.

4. Результаты исследования

4.1. Подготовка данных

В рамках исследования отобраны производственные функции в виде степенных мультипликативных форм:

$$\hat{y}_{k,i} = a_{k,i,0} \cdot \prod_{j=1}^J x_{k,i,j}^{a_{k,i,j}}, \quad (10)$$

где $\hat{y}_{k,i}$ — результативный признак — ВРП по i разделу ОКВЭД для k области;

$x_{k,i,j}$ — факторы, характеризующие условия функционирования. В данном случае для Тульской области $k = 19$ и в дальнейшем этот индекс опущен.

При выборе факторов и функциональных форм моделей авторы руководствовались ранее проведенными исследованиями [30].

Для устранения влияния изменения цен стоимостные индикаторы корректировались на инфляцию и приводились к 2007 г. Кроме того, данные, представленные по ОКВЭД 2 (действует с 2017 г.), были приведены к ОКВЭД 1.

В частности, были выбраны факторы (буквенное обозначение соответствующий раздел по ОКВЭД 1 (в скобках ОКВЭД 2)):

1) А (А) (сельское хозяйство) — стоимость основных фондов (ОФ) $x_{1,1}$; численность занятых (средняя за год) (ЧЗ) $x_{1,2}$;

2) С (В) (добыча полезных ископаемых) — ОФ $x_{2,1}$; ЧЗ $x_{2,2}$;

3) D (С) (обрабатывающие производства) — ОФ $x_{3,1}$; ЧЗ $x_{3,2}$; инвестиции в основной капитал (И) $x_{3,3}$;

4) Е (Е, D) (производство электроэнергии, газа и воды) — ОФ $x_{4,1}$; ЧЗ $x_{4,2}$;

5) F (F) (строительство) — ОФ $x_{5,1}$; ЧЗ $x_{5,2}$;

6) G (G) (оптовая и розничная торговля) — ОФ $x_{6,1}$; ЧЗ $x_{6,2}$; И $x_{6,3}$;

7) H (I) (гостиницы и рестораны) — ЧЗ $x_{7,2}$; И $x_{7,3}$;

8) I (H, J) (транспорт и связь) — ОФ $x_{8,1}$; ЧЗ $x_{8,2}$; пассажирооборот (автобусы) общего пользования $x_{8,4}$ и отправление железнодорожным транспортом пассажиров $x_{8,5}$;

9) J (K) (финансовая деятельность) — ОФ (всего) x_1 ; среднегодовая численность населения (далее — ЧН) x_2 ;

10) K (L, M, N) (операции с недвижимым имуществом) — ЧЗ $x_{10,2}$; И $x_{10,3}$;

11) L (O) (государственное управление) — ЧН x_2 ; расходы на социальную

¹ <https://rosstat.gov.ru/folder/210/document/13204>

² <https://www.fedstat.ru/>

³ <https://уровень-инфляции.рф/таблицы-инфляции>

⁴ <https://realnoevremya.ru/attachments/1760>

⁵ <https://www.fedstat.ru/indicator/59450>

политику (консолидированный бюджет) $x_{11,2}$;

12) М (Р) (образование) — ЧЗ $x_{12,1}$; расходы на образование (консолидированный бюджет) $x_{12,2}$;

13) N (Q) (здравоохранение) — ЧН x_2 ; заболеваемость $x_{13,2}$;

14) О (R, S) (предоставление прочих услуг) — ЧН x_2 ; расходы на социальную политику (консолидированный бюджет) $x_{11,2}$.

4.2. Анализ результативных признаков по разделам

Динамика изменения результативных признаков для Тульской области за 2018–2022 гг. представлена в табл. 1.

Из данных табл. 1 видно, что для раздела «Сельское хозяйство» наблюдается увеличение объема ВРП, за исключением последних двух оцениваемых лет. При этом численность занятых уменьшалась, начиная с 2013 г., а стоимость

Таблица 1. Динамика изменения результативных признаков для Тульской области за 2018–2022 гг., млн руб.

Table 1. Dynamics of changes in performance indicators for the Tula region in 2018–2022, million rubles

№	Модель	1	2	3	4	5
1	A (A)	16695,520	21420,629	22757,904	21722,825	16327,694
2	C (B)	1113,035	447,340	1422,369	1448,188	1507,172
3	D (C)	125494,659	111155,695	113505,045	123964,920	110023,537
4	E (E, D)	11408,605	11578,718	11663,426	10716,594	7535,859
5	F (F)	14469,451	18236,481	12516,847	16219,709	12057,374
6	G (G)	28660,643	26052,116	24464,747	22591,738	21351,600
7	H (I)	2226,069	2315,744	1706,843	2027,464	1758,367
8	I (H, J)	19478,107	19683,821	19059,744	15640,434	17583,670
9	J (K)	834,776	868,404	853,421	579,275	502,391
10	K (L, M, N)	23930,245	40525,514	40110,805	41418,186	33157,778
11	L (O)	10852,088	12157,654	11947,900	10426,956	10047,812
12	M (P)	8069,501	8105,103	8249,740	7530,579	8038,249
13	N (Q)	11408,605	11868,186	13085,795	11295,869	7787,054
14	O (R, S)	3617,363	4052,551	3129,212	4054,927	3516,734

Примечание: А, ..., О — модель для раздела по ОКВЭД 1; () — обозначение по ОКВЭД 2; 1, ... 5, 2018–2022 гг.

Источник: Рассчитано авторами по данным^{12,3,4}.

¹ <https://rosstat.gov.ru/folder/210/document/13204>

² <https://уровень-инфляции.рф/таблицы-инфляции>

³ <https://realnoevremya.ru/attachments/1760>

⁴ <https://www.fedstat.ru/indicator/59450>

основных фондов увеличивалась до 2020 г. с последующим незначительным снижением в 2021 и 2022 гг. То есть наблюдалось замещение фактора численности занятых фактором основных фондов. Увеличение ВРП при снижении основных факторов (ОФ, ЧЗ) можно объяснить вероятным изменением технологий и структуры производства сельскохозяйственной продукции, которые не учитываются в модели. При этом прослеживается явная корреляция ВРП со стоимостью основных фондов и численностью занятых.

Для раздела «Добыча полезных ископаемых» в целом, по сравнению с 2007 г., наблюдается рост, с колебаниями к снижению объема в 2020 г. по отношению к 2019 г. и последующими ростом в 2021 и 2022 гг. При общем росте основных фондов в абсолютном значении в скорректированных величинах их стоимость 2021 и 2022 гг. уменьшилась, что и сказалось на объеме ВРП по разделу С. В этот же период снизилась и занятость. Неизменный рост, вероятно, связан с изменением структуры добычи ископаемых, основными из которых в Тульской области являются уголь, железные руды, гипс и поваренная соль. Кроме того, наблюдался рост добычи строительных камней, что отразилось на объеме ВРП.

Если проследить динамику для обрабатывающих производств (раздел D, скорректированные данные), то в 2022 г. по сравнению с 2020 г. наблюдалось снижение ВРП. Аналогичные изменения претерпели основные фонды и численность занятых по данному виду деятельности.

Раздел «Производство электроэнергии, газа и воды» продемонстрировал положительную динамику до 2020 г. В 2022 г., как и для раздела D, наблюдалось снижение ввиду того, что в Тульской области основными потребителями, например электроэнергии, относящейся к данному разделу, явля-

ются обрабатывающие производства. Аналогичное снижение характерно и для стоимости основных фондов.

Сектор строительства F имел разнонаправленную динамику с ростом в 2021 г. по отношению к 2020 г. и снижением в 2022 г. по отношению к 2021 г. При этом в 2021 и 2022 гг. по отношению к 2020 г. наблюдался рост числа занятых и резкое снижение стоимости основных фондов в скорректированных ценах.

Сохранилась тенденция снижения ВРП в сфере оптовой и розничной торговли (G) на фоне колебаний численности занятых и изменениями основных фондов. При этом пик стоимости основных фондов пришелся на 2020 г. Для розничной торговли это был проблемный год, связанный с произошедшим локдауном, что видно по снижению численности занятых. Снижение ВРП может быть связано также с «перегретостью» этого рынка.

Аналогичная ситуация снижения ВРП в 2020 г. характерна и для сферы гостиничных и ресторанных услуг, что вызвано одной из причин — локдауном, с последующим восстановлением в 2021 г. и снижением до уровня 2019 г. в 2022 г. Последнее изменение, вероятно, обусловлено политическими причинами, которые трудно формализовать и ввести в качестве факторов в модель. При этом рост инвестиций в 2021 и 2022 гг. не позволили исправить ситуацию в краткосрочном периоде.

ВРП по разделу I (транспорт и связь) также наблюдалось снижение с 2019 по 2021 г. и рост в 2022 г. по сравнению с 2021 г., но его уровень не достиг уровня даже 2020 г. Очевидно, это явилось последствиями 2020 г. на фоне роста численности занятых, уменьшения основных фондов, а также уменьшения объема перевозок автобусами и железнодорожным транспортом. Кроме того, рост в 2022 г. объема ВРП в данной сфере связан с изменением стоимости транспортных услуг

и услуг связи, темп роста которых оказался выше, чем уровень инфляции.

Для финансовой деятельности снижение в Тульской области наблюдалось с 2019 г., что связано с уменьшением объема выданных кредитов вследствие увеличения процентной ставки на фоне тенденций снижения численности населения и сокращения стоимости основных фондов (2020–2022) (ввиду невозможности их обновления).

Рост инвестиций не привел к положительному эффекту в области операций с недвижимым имуществом.

Для раздела «Финансовая деятельность», начиная с 2020 г., наблюдалось снижение объема ВРП с уменьшением темпа на фоне негативной тенденции изменения среднегодовой численности населения, так же как и расходов на социальную политику.

Аналогичные выводы можно сделать и для других видов деятельности.

Как видно из результатов анализа, за рассматриваемые периоды меняются

как результаты, так и факторы, причем ряд факторов (условий) ввиду объективных обстоятельств не могут быть включены в модель. Поэтому целесообразно рассматривать последний период оценки и строить модель конкретно для него. Для реализации предложенной методологии и построения моделей необходимо перейти к нечетким представлениям результативных и факторных признаков.

4.3. Представление данных по факторам в виде нечетких чисел

Для каждого из факторов были получены значения в виде нечетких чисел.

Для их расчета была использована программная платформа «Инфоаналитика 2.0» [35].

На рис. 1 представлено значение объема ВРП по разделу А на числовой и лингвистической шкалах для Тульской области в 2022 г. В качестве нормы использовано среднее значение за 2007–2022 гг., число реперов классов равно 9.

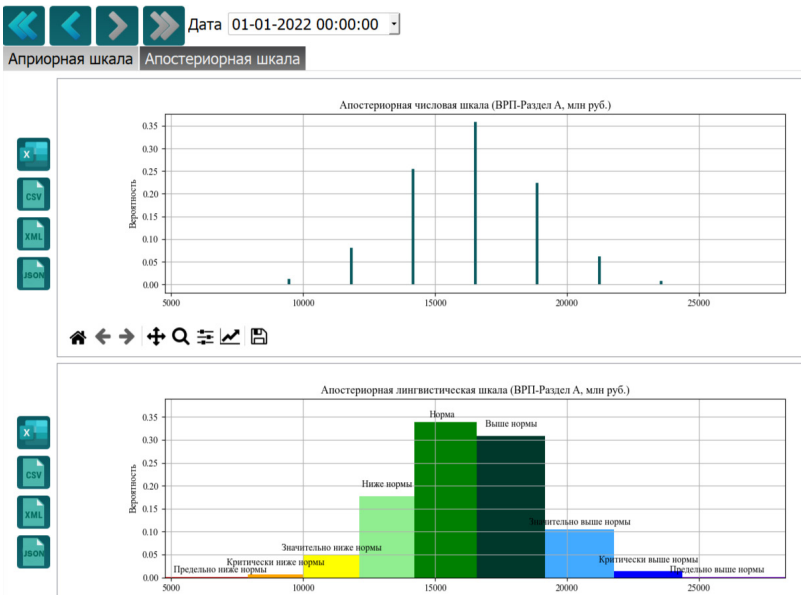


Рис. 1. Значение ВРП по разделу А на числовой и лингвистической шкалах для Тульской области в 2022 г.

Figure 1. The value of the GDP by region according to section A on numerical and linguistic scales for the Tula region in 2022

Таблица 2. Значения реперов и вероятности для объема ВРП по разделу А для Тульской области в 2022 г.

Table 2. Reference values and probabilities for the GDP by region according to section A for the Tula region in 2022

№	Значение репера	Терм	Вероятность	Нижняя граница	Верхняя граница
1	6869,584	Предельно ниже нормы	0,001	4767,469	7920,641
2	8971,699	Критически ниже нормы	0,007	7920,641	10022,756
3	11073,813	Значительно ниже нормы	0,049	10022,756	12124,871
4	13175,928	Ниже нормы	0,178	12124,871	14226,986
5	15278,043	Норма	0,339	14226,986	16576,537
6	17875,030	Выше нормы	0,309	16576,537	19173,524
7	20472,017	Значительно выше нормы	0,104	19173,524	21770,511
8	23069,004	Критически выше нормы	0,013	21770,511	24367,498
9	25665,991	Предельно выше нормы	0,001	24367,498	28262,978

Источник: рассчитано авторами.

Числовые значения реперов и вероятностей представлены в табл. 2.

4.4. Формирование выборки достаточного объема

На каждом из интервалов были сгенерированы значения результативных и факторных признаков общим объемом 7 000 наблюдений. В качестве базового

распределения использовано треугольное распределение.

Гистограммы сгенерированных выборок для ВРП, основных фондов и численности занятых по разделу А для Тульской области в 2022 г. представлены на рис. 2.

Фрагмент сгенерированных выборок представлен в табл. 3.

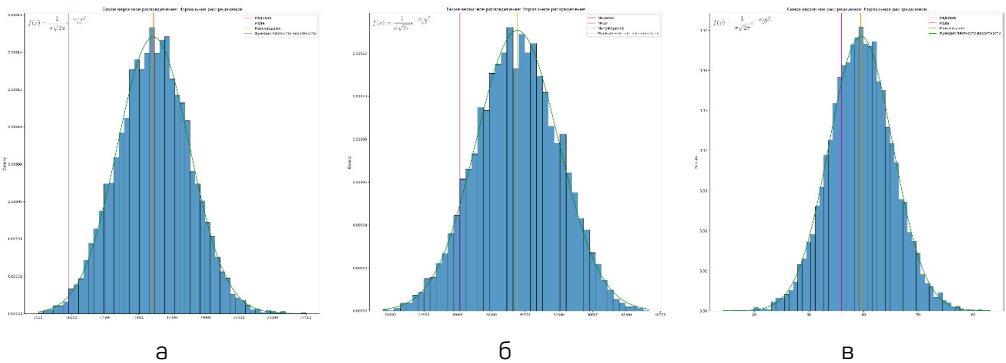


Рис. 2. Гистограммы сгенерированных выборок для ВРП (а), основные фонды (б) и численность занятых (в) по разделу А для Тульской области в 2022 г.

Figure 2. Histograms of the generated samples for GDP by region (a), Fixed Assets (б) and Employees (в) according to section A For the Tula region in 2022

Таблица 3. Фрагмент сгенерированных выборок ВРП, основных фондов и численности занятых по разделу А для Тульской области в 2022 г.

Table 3. Fragment of the generated samples for the GDP by region, Fixed Assets and Employees according to section A for the Tula region in 2022

№ п/п	Объем ВРП, млн руб.	Основные фонды, млн руб.	Численность занятых, тыс. чел.
1	5851,260	20434,485	28,334
2	7291,065	21860,154	27,685
3	7169,684	25732,702	26,452
4	7188,605	24592,437	25,468
5	8684,472	24886,831	26,409
6	8496,640	25623,197	26,093
7	8932,262	25377,953	26,954
8	8811,436	23745,217	24,385
9	8483,644	25401,627	25,415
10	9358,944	25372,247	28,355

Источник: рассчитано авторами.

4.5. Построение и тестирование моделей

Для каждого из разделов Тульской области по сгенерированным выборкам были построены соответствующие модели в форме (10). Для их расчета была использована программная платформа «ЭФРА» [34].

Спецификация и результаты тестирования моделей представлены в табл. 4.

Из данных табл. 4 видно, что коэффициенты модели значимы, а коэффициент детерминации значимо отличается

от 0. Кроме того, для моделей (за исключением модели раздела С) средняя относительная ошибка не превышает 5 %, что свидетельствует о их высокой точности.

В табл. 5 приведены значения стандартизованных (* — обозначение) коэффициентов моделей.

По данным табл. 5 можно судить о степени влияния факторов на результативный признак.

В табл. 6 представлены эластичности факторов моделей.

Таблица 4. Спецификация и результаты тестирования моделей для Тульской области по сгенерированным выборкам

Table 4. Specification and test results of models for the Tula region based on generated samples

№	Модель	a_0	a_1	a_2	a_3	a_4	R^2	MAPE, %
1	A (A)	2,243	0,758	0,287	—	—	0,834	4,594
2	C (B)	3,314	0,503	1,084	—	—	0,775	14,074
3	D (C)	30,023	0,195	0,933	0,111	—	0,881	3,638

Окончание табл. 4

№	Модель	a_0	a_1	a_2	a_3	a_4	R^2	$MAPE$, %
4	E (E, D)	6,623	0,349	1,044	—	—	0,909	3,734
5	F (F)	55,498	0,211	1,025	—	—	0,890	3,950
6	G (G)	12,722	0,137	0,694	0,042	—	0,895	3,902
7	H (I)	209,067	0,742	0,502	—	—	0,836	6,256
8	I (H, J)	520,708	0,137	0,473	0,054	0,052	0,902	2,732
9	J (K)	2,783	0,265	0,699	—	—	0,832	3,660
10	K (L, M, N)	480,674	0,154	0,456	—	—	0,843	4,744
11	L (O)	0,142	1,116	0,293	—	—	0,887	3,632
12	M (P)	27,460	0,647	0,292	—	—	0,844	3,140
13	N (Q)	0,012	0,760	1,164	—	—	0,816	6,882
14	O (R, S)	0,108	0,803	0,440	—	—	0,807	4,820

Примечание: a_i ($i = 0, \dots, 4$) — коэффициенты модели (для коэффициентов оценка осуществлялась по критерию Стьюдента, $p\text{-value} < 0,001$); A, ..., O — модель для раздела по ОКВЭД 1; () — обозначение по ОКВЭД 2; R^2 — коэффициент детерминации (значимость определялась по критерию Фишера, $p\text{-value} < 0,001$); $MAPE$ — средняя относительная ошибка выборки.

Источник: рассчитано авторами.

Таблица 5. Значения коэффициентов стандартизованных моделей для Тульской области по сгенерированным выборкам

Table 5. Coefficient values of standardized models for the Tula region based on the generated samples

№ п/п	Модель	a_1^*	a_2^*	a_3^*	a_4^*
1	A (A)	0,635	0,304	—	—
2	C (B)	0,357	0,509	—	—
3	D (C)	0,193	0,591	0,186	—
4	E (E, D)	0,324	0,651	—	—
5	F (F)	0,281	0,687	—	—
6	G (G)	0,194	0,642	0,139	—
7	H (I)	0,468	0,481	—	—
8	I (H, J)	0,213	0,482	0,128	0,165
9	J (K)	0,386	0,562	—	—
10	K (L, M, N)	0,463	0,507	—	—
11	L (O)	0,706	0,271	—	—
12	M (P)	0,598	0,362	—	—

Окончание табл. 5

№ п/п	Модель	a_1^*	a_2^*	a_3^*	a_4^*
13	N (Q)	0,353	0,560	—	—
14	O (R, S)	0,523	0,418	—	—

Примечание: a_i^* ($i = 1, \dots, 4$) — стандартизованные коэффициенты модели; А, ..., О — модель для раздела по ОКВЭД; () — обозначение по ОКВЭД2.

Источник: рассчитано авторами.

Таблица 6. Значения эластичностей факторов моделей для Тульской области по сгенерированным выборкам, %

Table 6. Elasticity values of models for the Tula region based on the generated samples, %

№	Модель	E_1	E_2	E_3	E_4
1	A (A)	0,811	0,107	—	—
2	C (B)	0,646	0,162	—	—
3	D (C)	0,203	0,408	0,095	—
4	E (E, D)	0,414	0,373	—	—
5	F (F)	0,177	0,420	—	—
6	G (G)	0,173	0,449	0,037	—
7	H (I)	0,226	0,205	—	—
8	I (H, J)	0,158	0,123	0,037	0,042
9	J (K)	0,372	0,523	—	—
10	K (L, M, N)	0,044	0,362	—	—
11	L (O)	0,885	0,327	—	—
12	M (P)	0,290	0,342	—	—
13	N (Q)	0,621	0,877	—	—
14	O (R, S)	0,718	0,554	—	—

Примечание: a_i^* ($i = 1, \dots, 4$) — стандартизованные коэффициенты модели; А, ..., О — модель для раздела по ОКВЭД; () — обозначение по ОКВЭД2.

Источник: рассчитано авторами.

По данным табл. 6 можно судить, насколько изменится результативный признак при изменении фактора на 1 %.

Рассмотрим раздел А. Среднемесячная заработная плата по сельскому хозяйству в Тульском регионе в 2022 г. была 49188,3 руб.

Увеличение численности занятых (в 2022 г. — 39,5 тыс. чел.) на 1 % составит 395 человек. Тогда общие среднегодовые затраты будут 233,2 млн руб. В этом случае прирост ВРП (коэффициент эластичности равен 0,107) будет насчитывать 1747,1 млн руб.

Изменение стоимости основных фондов на 1 % (в 2022 г. — 37347,8 млн руб.) приведет к увеличению основных фондов на 373,5 млн руб., а ВРП на 13241,8 млн руб. Если увеличить численность занятых в стоимостном выражении на величину стоимости основных фондов (то есть в 1,6 раза), то изменение ВРП будет в объеме 2798,5 млн руб., то есть полученный эффект почти в 4,7 раза меньше, чем при увеличении стоимости основных фондов.

Это означает, что производителю выгоднее модернизировать собственное производство, сокращая численность занятых в данной сфере. Такую стратегию хозяйствующих субъектов можно проследить по динамике сокращения численности занятых, начиная с 2013 г. Это подтверждает ранее высказанное предположение о замещении основными фондами трудовых ресурсов на фоне увеличения объема ВРП в рассматриваемые периоды.

Кроме того, заработная плата по разделу А составляет 84,7 % от среднемесячной заработной платы по разделу D (обрабатывающие производства), который является основным и приоритетным видом деятельности для Тульской области. Это также не стимулирует экономически активное население работать в данной сфере.

По разделу D эластичность для фактора занятости (0,408) существенно выше (почти в два раза), чем эластичность стоимости основных фондов. Согласно данным табл. 5, ситуация аналогична. Низкая эластичность 0,095 характерна для инвестиций. Вероятно, такой результат получен вследствие того, что в Тульском регионе обрабатывающая промышленность наукоемка и инвестиции дают эффект с запозданием, что не учитывает предложенная модель, хотя влияние инвестирования безусловно.

Еще одной возможной причиной такого результата может служить и управление инвестициями в данной сфере, что требует корректировки инвестиционной политики в области. В 2022 г. по разделу D заработная плата в среднем за месяц составила 58047,7 руб., численность занятых — 161,1 тыс. чел. Изменение ЧЗ на 1 % приведет к увеличению численности на 1610 чел. и затрат со стороны предприятия на (без учета страховых взносов) 1122,2 млн руб., а также увеличению объема ВРП на 44889,6 млн руб.

С другой стороны, увеличение стоимости основных фондов на 1 % даст дополнительные затраты на сумму 1813,4 млн руб. и, соответственно, ВРП на величину 22334,8 млн руб., что в 2,0 раза меньше, чем при увеличении численности занятых. Если уравнивать затраты на трудовые ресурсы и основные фонды, увеличив численность на 1,6 % (тогда затраты составят 1813,4 млн руб.), то прирост ВРП по разделу D составит 72541,1 млн руб., или в 3,2 раза больше, чем при изменении стоимости ОФ. Учитывая страховые выплаты (примем их равными 30 %), затраты на увеличение численности на 1 % составят 1458,8 тыс. руб., а при выравнивании затрат на ОФ и численность занятых даст прирост ВРП на 55800,8 млн руб., что в 2,5 раза выше. Следовательно, и в этом случае предприятиям обрабатывающей промышленности выгодней увеличивать трудовые ресурсы, чем обновлять или модернизировать основные фонды. Такая стратегия, за исключением 2021 г., прослеживалась с 2017 г.

Поскольку есть потенциал и стимул со стороны предприятий обрабатывающих производств, можно корректировать и уровень заработной платы, например до уровня заработной платы в финансовой сфере, где он выше почти в 1,3 раза выше. Однако невнимание к модернизации и обновлению основных фондов

со стороны органов управления может дать негативный эффект в среднесрочной перспективе ввиду возможного отставания в использовании современных средств производства.

Согласно данным табл. 6, для раздела G (оптовая и розничная торговля) прослеживается аналогичная разделу D ситуация по эластичности (0,173 — ОФ, 0,449 — ЧЗ, 0,037 — И). Увеличение численности занятых на 1 % приведет к увеличению затрат (среднемесячная заработная плата в 2022 г. составила 42928,5 тыс. руб., или 79 % от уровня по разделу D) на 838,4 млн руб. (с учетом страховых выплат), тогда как для основных фондов (в 2022 г. — 12682,9 млн руб.) — на 126,8 млн руб. В свою очередь такое изменение даст прирост ВРП соответственно на 9586,9 млн руб. и 3693,8 млн руб.

Если воспользоваться аналогичным уравниванием затрат (для численности занятых затраты в 6,6 раз больше, чем стоимость основных фондов) и уменьшить изменение численности занятых до 0,2 %, то увеличение ВРП составит 1450,2 млн руб., что в 2,5 раза меньше, чем при корректировке основных фондов.

Следовательно, предприятиям оптовой и розничной торговли целесообразно увеличивать стоимость основных фондов и уменьшать численность занятых. Последняя тенденция как раз и наблюдалась в 2022 г., по сравнению с 2021 г.

Однако увеличение стоимости основных фондов может привести к менее ожидаемому результату со снижением темпов роста, что прежде всего связано с возможным насыщением данной сферы деятельности. Действительно, наблюдавшийся в последние годы рост количества торговых центров, оптовых баз и других торговых точек привел к перераспределению потребителей среди новых структур и оттоку из точек про-

даж, существовавших ранее в силу ограниченной и постоянно уменьшающейся численности населения. А рост потребления замедляется инфляционными процессами на фоне незначительного роста реальных денежных доходов населения согласно официальной статистике¹. Эластичность по инвестициям также незначительна и составила в 2022 г. 0,037.

Аналогичные рассуждения можно провести и для остальных видов деятельности, осуществляемых в Тульской области.

5. Обсуждение

Результаты, рассмотренные в данной статье и посвященные применению комбинации эконометрического моделирования и нечеткого моделирования на основе байесовских интеллектуальных измерений к изучению экономической подсистемы СЭЭС, дают возможность методологически обоснованно, в силу наличия достаточного объема выборки (согласно [27]), использовать методы математической статистики для построения корректных регрессионных моделей. Полученные соотношения и правомерность их использования подтверждаются работами Huning & Wahl [6] и Zhukov et al. [30].

В отличие, например, от исследований D'Andrea [13] и Duan et al. [14], генерация выборки осуществлялась на нечетких данных, но с учетом конкретных условий функционирования региона, что расширяет возможности применения байесовского подхода при изучении экономических процессов в сложных системах.

Полученные модели для объема ВРП по разделам ОКВЭД обладают статистически значимыми R^2 и низкой MAPE. Таким образом, возможно мето-

¹ https://rosstat.gov.ru/storage/mediabank/Urov_11subg-nm.xlsx

дологически обоснованное построение эконометрических моделей по данным для отдельного субъекта Российской Федерации (а именно, рассматривается экономическая подсистема Тульской области) в определенный период времени, что подтверждает поставленную гипотезу.

Кроме того, полученные модели дают возможность оценить степень влияния факторов на результативный признак, в частности при преобразовании моделей к стандартизованному виду (табл. 2) или вычислении соответствующих эластичностей.

Превалирующее влияние основных фондов по отношению к другим факторам, включенным в модели, на результативный признак характерно для разделов А (А); численности занятых — для С (В), D (С), E (Е, D), F (F), G (G), I (H, J), M (P); инвестиций — K (L, M, N); среднегодовой численности населения — J (K), L (O), O (R, S). За счет замещения факторов моделей можно добиться заданных (плановых) значений результативных признаков, что с точки зрения управления представляет собой регулирование процессов в рамках социально-экономического развития территорий, таких, например, как модернизация основных фондов, корректировка политики занятости, демографическая политика, перераспределение расходов в консолидированном бюджете области и т. п. Например, Тульская область обладает промышленной специализацией (доля в ВРП в 2022 г. — 43,8 %). Однако степень влияния основных фондов существенно ниже, чем численности занятых. Это может означать, что необходимо улучшать и модернизировать ОФ в ближайшем будущем.

На втором месте по доле в ВРП Тульской области стоит раздел К (операции с недвижимым имуществом; предоставление услуг; профессиональная,

научная и техническая деятельность по ОКВЭД2) (13,2 %). Здесь соотношение между коэффициентами численности занятых (0,463) и инвестициями (0,507) близко к единице 0,921. Отметим, что это единственный вид деятельности, в котором влияние инвестиций выше относительно других факторов. Следовательно, можно сделать вывод о заинтересованности институциональных единиц области в развитии данного направления.

На третьем месте по доле находится раздел G (торговля) (8,5 %), где значения стандартизованных коэффициентов соответственно равны 0,194 (основные фонды), 0,642 (численность занятых) и 0,139 (инвестиции). Такой результат (наибольшее влияние у занятых) ожидаем вследствие специфики данного вида деятельности. Значения коэффициентов для основных фондов и инвестиций сравнимы между собой, их отношение к численности занятых составляет 0,302 и 0,217 соответственно.

Четвертое место занимает раздел I (транспорт и связь) (доля ВРП 7,0 %). Соответствующие стандартизованные коэффициенты равны 0,213 — основные фонды, 0,482 — численность занятых, 0,128 — пассажирооборот автобусами, а для отправления пассажиров РЖД — 0,165. И по этому разделу превалирует влияние численности занятых.

На пятом месте находится раздел А (6,5 %). В отличие от остальных видов деятельности, наибольшее влияние оказывает наличие основных фондов. Если посмотреть на тенденцию изменения численности занятых, то она неуклонно уменьшается, начиная с 2013 г. Одной из причин может являться низкий уровень заработной платы по сравнению с другими сферами деятельности.

Таким образом, характер влияния факторов на результативный признак

(при анализе стандартизованных коэффициентов) может дать ответ на особенности функционирования элементов экономической подсистемы. Кроме того, сравнение с параметрами моделей (например, по датасетам других областей (Центральный федеральный округ)) по кроссекционной или панельной (объединенной по пространству и во времени) выборкам даст возможность оценить соответствие функционирования субъектов Российской Федерации общерегиональным тенденциям. Полученные таким образом модели могут выступать в качестве нормативных моделей, представленных, например, в [33, 34].

Следует отметить наличие ограничений на построенные модели, касающиеся того, что они могут быть использованы только для Тульской области. Для других субъектов Российской Федерации необходимо строить собственные модели по собственным данным. Кроме того, разработанные модели могут быть применены для поиска количественных значений, необходимых для улучшения результатов функционирования экономической подсистемы в рамках оптимизационных моделей, аналогичных моделям в работе Жукова [34] лишь на краткосрочный период без введения дополнительных предположений.

Работа по преодолению таких ограничений будет проведена в дальнейших исследованиях.

6. Заключение

Данное исследование устраняет методологический пробел при построении эконометрических моделей по данным, характеризующимся ограниченным набором, неполнотой и нечеткостью, что, по сути, дает возможность корректно и методологически обосновано моделировать функциони-

рование сложных систем в условиях неопределенности, характерных для социо-эколого-экономических систем. По его результатам можно сформулировать следующие выводы.

1. Подтверждена гипотеза относительно возможности методологически обоснованного построения эконометрических моделей по данным для отдельного субъекта Российской Федерации в определенный период времени в условиях неполноты и нечеткости данных.

2. Предложены этапы, позволяющие объединить методологии эконометрического моделирования и байесовских интеллектуальных измерений применительно к социо-эколого-экономической системе, ее подсистемам (в частности, экономическим подсистемам) и элементам.

3. Проведена апробация представленной методологии на примере экономической подсистемы Тульской области, в результате которой были построены модели связи объема ВРП для 14 разделов ОКВЭД 1 по данным за 2022 г., обладающие высоким качеством и статистически значимыми параметрами.

4. Построенные модели могут служить инструментом анализа степени воздействия влияющих факторов на результативный признак с целью дальнейшего регулирования социально-экономической политики Тульской области в рамках разрабатываемых программ развития.

Проведенное исследование вносит определенный вклад в развитие моделирования функционирования сложных систем — социо-эколого-экономических систем, за счет комбинирования методов эконометрики с методологией байесовских интеллектуальных измерений с целью методологически обоснованного построения моделей по единичным датасетам и измерениям, которые обладают специфическими условиями и ограниче-

ниями. Это определяет теоретическую значимость работы.

Дальнейшее применение моделей заключается в их использовании в оптимизационных моделях, дающих возможность получить количествен-

но выраженные изменения факторов, необходимые для обеспечения устойчивого и сбалансированного развития области, что представляет собой практическую значимость проведенного исследования.

Список использованных источников

1. Akperov I., Kurianov N. The assessment of regional economic potential based on the methodology of fuzzy set theory // *Procedia Computer Science*. 2017. Vol. 120. Pp. 372–375. <https://doi.org/10.1016/j.procs.2017.11.252>
2. Zadeh L. A. The Concept of Linguistic Variable and its Application to Approximate Reasoning // *Information Sciences*. 1975. Vol. 8, Issue 3. Pp. 199–249. [https://doi.org/10.1016/0020-0255\(75\)90036-5](https://doi.org/10.1016/0020-0255(75)90036-5)
3. Zolfani S. H., Krishankumar R., Pamucar D., Gorçün O. F. The potentials of the Southern & Eastern European countries in the process of the regionalization of the global supply chains using a q-rung orthopair fuzzy-based integrated decision-making approach // *Computers & Industrial Engineering*. 2022. Vol. 171. 108405. <https://doi.org/10.1016/j.cie.2022.108405>
4. Huang H., Tian Y., Tao Zh. Multi-rule combination prediction of compositional data time series based on multivariate fuzzy time series model and its application // *Expert Systems with Applications*. 2024. Vol. 238. 121966. <https://doi.org/10.1016/j.eswa.2023.121966>
5. Ouifak H., Idri A. Application of neuro-fuzzy ensembles across domains: A systematic review of the two last decades (2000–2022) // *Engineering Applications of Artificial Intelligence*. 2023. Vol. 124. 106582. <https://doi.org/10.1016/j.engappai.2023.106582>
6. Huning T. R., Wahl F. The fetters of inheritance? Equal partition and regional economic development // *European Economic Review*. 2021. Vol. 136. 103776. <https://doi.org/10.1016/j.eurocorev.2021.103776>
7. Kutty A. A., Kucukvar M., Onat N. C., Ayvaz B., Abdella G. M. Measuring sustainability, resilience and livability performance of European smart cities: A novel fuzzy expert-based multi-criteria decision support model // *Cities*. 2023. Vol. 137. 104293. <https://doi.org/10.1016/j.cities.2023.104293>
8. Ruben R. B., Rajendran C., Ram R. S., Kouki F., Alshahrani H. M., Assiri M. Analysis of barriers affecting Industry 4.0 implementation: An interpretive analysis using total interpretive structural modeling (TISM) and Fuzzy MICMAC // *Heliyon*. 2023. Vol. 9, Issue 12. e22506. <https://doi.org/10.1016/j.heliyon.2023.e22506>
9. Ziesmer J., Jin D., Mukashov A., Henning Ch. Integrating fundamental model uncertainty in policy analysis A Bayesian averaging approach combining CGE-models with metamodeling techniques // *Socio-Economic Planning Sciences*. 2023. Vol. 87. 101591. <https://doi.org/10.1016/j.seps.2023.101591>
10. Forrester A, Sobester A, Keane A. *Engineering Design via Surrogate Modelling: A Practical Guide*. John Wiley & Sons, 2008. <https://doi.org/10.1002/9780470770801>
11. Myers R. H., Montgomery D., Vining G. G., Robinson T. J. *Generalized Linear Models: With Applications in Engineering and the Sciences*. Second Edition. John Wiley & Sons, 2012. <https://doi.org/10.1002/9780470556986>
12. Iacopini M., Poon A., Rossini L., Zhu D. Bayesian mixed-frequency quantile vector autoregression: Eliciting tail risks of monthly US GDP // *Journal of Economic Dynamics & Control*. 2023. Vol. 157. 104757. <https://doi.org/10.1016/j.jedc.2023.104757>
13. D'Andrea S. Are there any robust determinants of growth in Europe? A Bayesian Model Averaging approach // *International Economics*. 2022. Vol. 171. Pp. 143–173. <https://doi.org/10.1016/j.inteco.2022.06.001>

14. Duan K., Lan F., Zhao Y., Huang Y. Housing networks in urban China: A panel VAR model with Bayesian stochastic search // *Cities*. 2023. Vol. 140. 104400. <https://doi.org/10.1016/j.cities.2023.104400>
15. Canoz I., Kalkavan H. Forecasting the dynamics of the Istanbul real estate market with the Bayesian time-varying VAR model regarding housing affordability // *Habitat International*. 2024. Vol. 148. 103055. <https://doi.org/10.1016/j.habitatint.2024.103055>
16. Salles Th.T, Nogueira D. A., Beijo L. A., Silva L. F. Bayesian approach and extreme value theory in economic analysis of forestry projects // *Forest Policy and Economics*. 2019. Vol. 105. Pp. 64–71. <https://doi.org/10.1016/j.forpol.2019.05.021>
17. Debnath B., Shakur M. S., Bari A. B.M.M., Karmaker Ch.L. A Bayesian Best-Worst approach for assessing the critical success factors in sustainable lean manufacturing // *Decision Analytics Journal*. 2023. Vol. 6. 100157. <https://doi.org/10.1016/j.dajour.2022.100157>
18. Klein N., Herwartz H., Kneib Th. Modelling regional patterns of inefficiency: A Bayesian approach to geoadditive panel stochastic frontier analysis with an application to cereal production in England and Wales // *Journal of Econometrics*. 2020. Vol. 214, Issue 2. Pp. 513–539. <https://doi.org/10.1016/j.jeconom.2019.07.003>
19. Szafranek K., Szafranski G., Leszczyńska-Pachecna A. Inflation returns. Revisiting the role of external and domestic shocks with Bayesian structural VAR // *International Review of Economics and Finance*. 2024. Vol. 93. Pp. 789–810. <https://doi.org/10.1016/j.iref.2024.03.054>
20. Mozdzen A., Cremaschi A., Cadonna A., Guglielmi A., Kastner G. Bayesian modeling and clustering for spatio-temporal areal data: An application to Italian unemployment // *Spatial Statistics*. 2022. Vol. 52. 100715. <https://doi.org/10.1016/j.spasta.2022.100715>
21. Elhorst J. P. The dynamic general nesting spatial econometric model for spatial panels with common factors: Further raising the bar // *Review of Regional Research*. 2022. Vol. 42. Pp. 249–267. <https://doi.org/10.1007/s10037-021-00163-w>
22. Noia A. D., Mastrantonio G., Lasinio G. J. Bayesian size-and-shape regression modelling // *Statistics and Probability Letters*. 2024. Vol. 204. 109928. <https://doi.org/10.1016/j.spl.2023.109928>
23. Aivazian S. A. Quality of Life and Living Standards Analysis: An Econometric Approach. De Gruyter, 2016. <https://doi.org/10.1515/9783110316254>
24. Potapenko A., Popov A., Vorontsov K. Interpretable probabilistic embeddings: Bridging the gap between topic models and neural networks // *Proceedings of 6th Conference Artificial Intelligence and Natural Language. AINL 2017*. Edited by A. Filchenkov, L. Pivovarova, J. Žižka. Communications in Computer and Information Science. Vol. 789. Springer, Cham, 2018. Pp. 167–180. https://doi.org/10.1007/978-3-319-71746-3_15
25. Averkin A., Yarushev S. Deep neural networks in semantic analysis // *Proceedings of 10th International Conference on Theory and Application of Soft Computing, Computing with Words and Perceptions — ICSCCW-2019*. Edited by R. A. Aliev, J. Kacprzyk, W. Pedrycz, M. Jamshidi, M. B. Babanli, F. M. Sadikoglu. Advances in Intelligent Systems and Computing. Vol. 1095. Springer, Cham, 2020. Pp. 846–853. https://doi.org/10.1007/978-3-030-35249-3_112
26. Бормотов И. В., Жуков Р. А., Руднева И. Д., Спиридонова К. А. Ценностный мир молодежи: состояние и перспективы в условиях экономических санкций // *Научные труды Вольного экономического общества России*. 2023. Т. 243, № 5. С. 34–50. <https://doi.org/10.38197/2072-2060-2023-243-5-34-50>
27. Прокопчина С. В. Проектный менеджмент в условиях неопределенности на основе регуляризирующего байесовского подхода // *Мягкие измерения и вычисления*. 2023. Т. 69, № 8. С. 50–68. <https://doi.org/10.36871/26189976.2023.08.004>
28. Жуков Р. А., Прокопчина С. В., Гиниатов И. А., Николина Е. М. Применение библиотеки «Байесовская математическая статистика» в программном комплексе «Инфоинтегратор» // *Мягкие измерения и вычисления*. 2022. Т. 54, № 5. С. 99–108. <https://doi.org/10.36871/2618-9976.2022.05.008>

29. Zhong F., Chen R., Luo X., Song X., Ullah A. Assessing regional resilience in China using a sustainable livelihoods approach: Indicators, influencing factors, and the relationship with economic performance // *Ecological Indicators*. 2024. Vol. 158. 111588. <https://doi.org/10.1016/j.ecolind.2024.111588>
30. Zhukov R. A., Kuznetsov G. V., Fomicheva I. V., Myasnikova E. B., Vasina M. V., Tsigler M. V. A Model of socio-ecological and economic system: the Tula region of the Russian Federation // *Journal of Environmental Management and Tourism*. 2019. Vol. 10. No. 7. Pp. 1539–1558. [https://doi.org/10.14505/jemt.v10.7\(39\).12](https://doi.org/10.14505/jemt.v10.7(39).12)
31. Andrews I., Stock J. H., Sun L. Weak Instruments in Instrumental Variables Regression: Theory and Practice // *Annual Review of Economics*. 2019. Vol. 11. Pp. 727–753. <https://doi.org/10.1146/annurev-economics-080218-025643>
32. Макаров В. Л., Айвазян С. А., Афанасьев М. Ю., Бахтизин А. Р., Нанавян А. М. Моделирование развития экономики региона и эффективность пространства инноваций // *Форсайт*. 2016. Т. 10, № 3. С. 76–90. <https://doi.org/10.17323/1995-459X.2016.3.76.90>
33. Жуков Р. А., Плинская М. А., Манохин Е. В. Оценка функционирования регионов на основе производственных функций с приведенными стоимостными факторами // *Journal of Applied Economic Research*. 2023. Т. 22, № 3. С. 657–682. <https://doi.org/10.15826/vestnik.2023.22.3.027>
34. Жуков Р. А. Подход к оценке функционирования иерархических социально-экономических систем и принятию решений на базе программного комплекса «ЭФРА» // *Бизнес-информатика*. 2020. Т. 14, № 3. С. 82–95. <https://doi.org/10.17323/2587-814X.2020.3.82.95>
35. Жуков Р. А., Прокопчина С. В., Бормотов И. В., Манохин Е. В., Руднева И. Д. Модель ценностного мира российской молодежи на основе байесовских байесовских интеллектуальных измерений // *МИР (Модернизация. Инновации. Развитие)*. 2024. Т. 15, № 1. С. 96–114. <https://doi.org/10.18184/2079-4665.2024.15.1.96-114>

ИНФОРМАЦИЯ ОБ АВТОРАХ

Жуков Роман Александрович

Доктор экономических наук, доцент, научный сотрудник Тульского филиала Финансового университета при Правительстве Российской Федерации, г. Тула, Россия (300012, г. Тула, ул. Оружейная, 1а); ORCID <https://orcid.org/0000-0002-2280-307X> e-mail: pluszh@mail.ru

Прокопчина Светлана Васильевна

Доктор технических наук, профессор кафедры системного анализа в экономике Финансового университета при Правительстве Российской Федерации, г. Москва, Россия (125167, Москва, Ленинградский просп., 49/2); ORCID <https://orcid.org/0000-0001-5500-2781> e-mail: svprokopchina@mail.ru

Плинская Мария Александровна

Магистрант Тульского филиала Финансового университета при Правительстве Российской Федерации, г. Тула, Россия (300012, г. Тула, ул. Оружейная, 1а); ORCID <https://orcid.org/0000-0002-1307-0935> e-mail: plinskaya@gmail.com

Желункина Мария Анатольевна

Магистрант Тульского филиала Финансового университета при Правительстве Российской Федерации, г. Тула, Россия (300012, г. Тула, ул. Оружейная, 1а); ORCID <https://orcid.org/0009-0006-3129-2749> e-mail: maria202001@yandex.ru

БЛАГОДАРНОСТИ

Исследование выполнено за счет гранта Российского научного фонда № 24-28-20020, <https://rscf.ru/project/24-28-20020/> и Тульской области.

ДЛЯ ЦИТИРОВАНИЯ

Жуков Р. А., Прокопчина С. В., Плинская М. А., Желуница М. А. Моделирование функциональных связей региональных экономических систем по малым выборкам на основе байесовских интеллектуальных измерений // Journal of Applied Economic Research. 2024. Т. 23, № 3. С. 721–750. <https://doi.org/10.15826/vestnik.2024.23.3.029>

ИНФОРМАЦИЯ О СТАТЬЕ

Дата поступления 26 мая 2024 г.; дата поступления после рецензирования 18 июня 2024 г.; дата принятия к печати 22 июня 2024 г.

Modeling of Functional Relationships of Regional Economic Systems Based on Small Samples Based on Bayesian Intelligent Measurements

Roman A. Zhukov¹  , Svetlana V. Prokopchina² , Maria A. Plinskaya¹ ,
Maria A. Zhelunitsina¹ 

¹Tula Branch of The Financial University under the Government
of the Russian Federation,
Tula, Russia

²The Financial University under the Government of the Russian Federation,
Moscow, Russia

 pluszh@mail.ru

Abstract. Regional economies, which are economic subsystems of a socio-ecological-economic system, are characterized by rapidly changing conditions affecting their functioning. Therefore, traditional approaches based on long-term trends and the construction of appropriate models are not enough for modeling and need to be over updated, including through combination with other approaches that take into account the volatility of conditions. Open sources datasets are fraught with uncertainty (different sources may have different values for the same indicator). So, some indicators are being constantly adjusted: changes are introduced into datasets from the previous periods; evaluation periods are not sufficient for model construction that would be methodologically valid from the point of view of mathematical statistics. Additionally, the results of functioning depend on influencing factors as well as measurement results. These problems lead to the need to resort to probabilistic estimates and the incompleteness and fuzziness dataset, as well as to obtain samples of sufficient volume from low-power data. The purpose of the study is to model the relationship between the volume of gross domestic product and the factors characterizing the functioning of the economic subsystem of the regions, under the conditions of uncertainty and limited data in a given period of time. The case study is constructed on The Tula region dataset. The working hypothesis is that it is possible to construct a methodologically correct econometric models using the dataset of a specific region of the Russian Federation over a certain period of time. We applied the methods of econometric modeling and the methodology of Bayesian intelligent measurements, as well as the methodology for generating a sufficient amount of data on small samples. For the Tula region and dataset for 2022, nonlinear models for the GDP by region volumes (14 sections of the NACE) were built. This made it possible to test the presented methodology based on a combination of econometric and Bayesian approaches to modeling the complex systems functioning, with conclusions having both theoretical and practical value for the region's sustainable development.

Key words: gross domestic product by region; production function; socio-ecological-economic system; model; Bayesian intellectual measurements; small sample.

JEL C10, C43, P25, R15, R11

References

1. Akperov, I, Kurianov, N. (2017). The assessment of regional economic potential based on the methodology of fuzzy set theory. *Procedia Computer Science*, Vol. 120, 372–375. <https://doi.org/10.1016/j.procs.2017.11.252>
2. Zadeh, L.A. (1975). The Concept of Linguistic Variable and its Application to Approximate Reasoning. *Information Sciences*, Vol. 8, Issue 3, 199–249. [https://doi.org/10.1016/0020-0255\(75\)90036-5](https://doi.org/10.1016/0020-0255(75)90036-5)

3. Zolfani, S.H., Krishankumar, R., Pamucar, D., Gorçün, O.F. (2022). The potentials of the Southern & Eastern European countries in the process of the regionalization of the global supply chains using a q-rung orthopair fuzzy-based integrated decision-making approach. *Computers & Industrial Engineering*, Vol. 171, 108405. <https://doi.org/10.1016/j.cie.2022.108405>
4. Huang, H., Tian, Y., Tao, Zh. (2024). Multi-rule combination prediction of compositional data time series based on multivariate fuzzy time series model and its application. *Expert Systems with Applications*, Vol. 238, 121966. <https://doi.org/10.1016/j.eswa.2023.121966>
5. Ouifak, H., Idri, A. (2023). Application of neuro-fuzzy ensembles across domains: A systematic review of the two last decades (2000–2022). *Engineering Applications of Artificial Intelligence*, Vol. 124, 106582. <https://doi.org/10.1016/j.engappai.2023.106582>
6. Huning, T.R., Wahl, F. (2021). The fetters of inheritance? Equal partition and regional economic development. *European Economic Review*, Vol. 136, 103776. <https://doi.org/10.1016/j.euroecorev.2021.103776>
7. Kutty, A.A., Kucukvar, M., Onat, N.C., Ayvaz, B., Abdella, G.M. (2023). Measuring sustainability, resilience and livability performance of European smart cities: A novel fuzzy expert-based multi-criteria decision support model. *Cities*, Vol. 137, 104293. <https://doi.org/10.1016/j.cities.2023.104293>
8. Ruben, R.B., Rajendran, C., Ram, R.S., Kouki, F., Alshahrani, H.M., Assiri, M. (2023). Analysis of barriers affecting Industry 4.0 implementation: An interpretive analysis using total interpretive structural modeling (TISM) and Fuzzy MICMAC. *Heliyon*, Vol. 9, Issue 12, e22506. <https://doi.org/10.1016/j.heliyon.2023.e22506>
9. Ziesmer, J., Jin, D., Mukashov, A., Henning, Ch. (2023). Integrating fundamental model uncertainty in policy analysis A Bayesian averaging approach combining CGE-models with meta-modeling techniques. *Socio-Economic Planning Sciences*, Vol. 87, 101591. <https://doi.org/10.1016/j.seps.2023.101591>
10. Forrester, A., Sobester, A., Keane, A. (2008). *Engineering Design via Surrogate Modelling: A Practical Guide*. John Wiley & Sons. <https://doi.org/10.1002/9780470770801>
11. Myers, R.H., Montgomery, D., Vining, G.G., Robinson, T.J. (2012). *Generalized Linear Models: With Applications in Engineering and the Sciences*. Second Edition. John Wiley & Sons. <https://doi.org/10.1002/9780470556986>
12. Iacopini, M., Poon, A., Rossini, L., Zhu, D. (2023). Bayesian mixed-frequency quantile vector autoregression: Eliciting tail risks of monthly US GDP. *Journal of Economic Dynamics & Control*, Vol. 157, 104757. <https://doi.org/10.1016/j.jedc.2023.104757>
13. D'Andrea, S. (2022). Are there any robust determinants of growth in Europe? A Bayesian Model Averaging approach. *International Economics*, Vol. 171, 143–173. <https://doi.org/10.1016/j.inteco.2022.06.001>
14. Duan, K., Lan, F., Zhao, Y., Huang, Y. (2023). Housing networks in urban China: A panel VAR model with Bayesian stochastic search. *Cities*, Vol. 140, 104400. <https://doi.org/10.1016/j.cities.2023.104400>
15. Canoz, I., Kalkavan, H. (2024). Forecasting the dynamics of the Istanbul real estate market with the Bayesian time-varying VAR model regarding housing affordability. *Habitat International*, Vol. 148, 103055. <https://doi.org/10.1016/j.habitatint.2024.103055>
16. Salles, Th.T, Nogueira, D.A., Beijo, L.A., Silva, L.F. (2019). Bayesian approach and extreme value theory in economic analysis of forestry projects. *Forest Policy and Economics*, Vol. 105, 64–71. <https://doi.org/10.1016/j.forpol.2019.05.021>
17. Debnath, B., Shakur, M.S., Bari, A.B.M.M., Karmaker, Ch.L. (2023). A Bayesian Best-Worst approach for assessing the critical success factors in sustainable lean manufacturing. *Decision Analytics Journal*, Vol. 6, 100157. <https://doi.org/10.1016/j.dajour.2022.100157>
18. Klein, N., Herwartz, H., Kneib, Th. (2020). Modelling regional patterns of inefficiency: A Bayesian approach to geoadaptive panel stochastic frontier analysis with an application to

cereal production in England and Wales. *Journal of Econometrics*, Vol. 214, Issue 2, 513–539. <https://doi.org/10.1016/j.jeconom.2019.07.003>

19. Szafranek, K., Szafranski, G., Leszczyńska-Paczesna, A. (2024). Inflation returns. Revisiting the role of external and domestic shocks with Bayesian structural VAR. *International Review of Economics and Finance*, Vol. 93, 789–810. <https://doi.org/10.1016/j.iref.2024.03.054>

20. Mozdzen, A., Cremaschi, A., Cadonna, A., Guglielmi, A., Kastner, G. (2022). Bayesian modeling and clustering for spatio-temporal areal data: An application to Italian unemployment. *Spatial Statistics*, Vol. 52, 100715. <https://doi.org/10.1016/j.spasta.2022.100715>

21. Elhorst, J.P. (2022). The dynamic general nesting spatial econometric model for spatial panels with common factors: Further raising the bar. *Review of Regional Research*, Vol. 42, 249–267. <https://doi.org/10.1007/s10037-021-00163-w>

22. Noia, A.D., Mastrantonio, G., Lasinio, G.J. (2024). Bayesian size-and-shape regression modelling. *Statistics and Probability Letters*, Vol. 204, 109928. <https://doi.org/10.1016/j.spl.2023.109928>

23. Aivazian, S.A. (2016). *Quality of Life and Living Standards Analysis: An Econometric Approach*. De Gruyter. <https://doi.org/10.1515/9783110316254>

24. Potapenko, A., Popov, A., Vorontsov, K. (2018). Interpretable probabilistic embeddings: Bridging the gap between topic models and neural networks. *Proceedings of 6th Conference Artificial Intelligence and Natural Language. AINL 2017*. Edited by A. Filchenkov, L. Pivovarova, J. Žižka. Communications in Computer and Information Science, Vol. 789. Springer, Cham, 167–180. https://doi.org/10.1007/978-3-319-71746-3_15

25. Averkin, A., Yarushev, S. (2020). Deep neural networks in semantic analysis. *Proceedings of 10th International Conference on Theory and Application of Soft Computing, Computing with Words and Perceptions — ICSCCW-2019*. Edited by R. A. Aliev, J. Kacprzyk, W. Pedrycz, M. Jamshidi, M. B. Babanli, F. M. Sadikoglu. Advances in Intelligent Systems and Computing, Vol. 1095. Springer, Cham, 846–853. https://doi.org/10.1007/978-3-030-35249-3_112

26. Bormotov, I.V., Zhukov, R.A., Rudneva I. D., Spiridonova K. A. (2023). Youth World of Values: Status and Trends under Economic Sanctions. *Scientific Works of the Free Economic Society of Russia*, Vol. 243, No. 5, 34–50. (In Russ.). <https://doi.org/10.38197/2072-2060-2023-243-5-34-50>

27. Prokopchina, S.V. (2023). Project management in conditions of uncertainty based on the regularizing Bayesian approach. *Soft Measurements and Computing*, Vol. 69, No. 8, 50–68. (In Russ.). <https://doi.org/10.36871/26189976.2023.08.004>

28. Zhukov, R.A., Prokopchina, S.V., Giniatov, I.A., Nikolina, E.M. (2022). Application of The Bayesian Mathematical Statistics Library in the Infointegrator Software Package. *Soft Measurement and Computing*, Vol. 54, No. 5, 99–108. (In Russ.). <https://doi.org/10.36871/2618-9976.2022.05.008>

29. Zhong, F., Chen, R., Luo, X., Song, X., Ullah, A. (2024). Assessing regional resilience in China using a sustainable livelihoods approach: Indicators, influencing factors, and the relationship with economic performance. *Ecological Indicators*, Vol. 158, 111588. <https://doi.org/10.1016/j.ecolind.2024.111588>

30. Zhukov, R.A., Kuznetsov, G.V., Fomicheva, I.V., Myasnikova, E.B., Vasina, M.V., Tsigler, M.V. (2019). A Model of socio-ecological and economic system: the Tula region of the Russian Federation. *Journal of Environmental Management and Tourism*, Vol. 10, No. 7, 1539–1558. [https://doi.org/10.14505/jemtv10.7\(39\).12](https://doi.org/10.14505/jemtv10.7(39).12)

31. Andrews, I., Stock, J.H., Sun, L. (2019). Weak Instruments in Instrumental Variables Regression: Theory and Practice. *Annual Review of Economics*, Vol. 11, 727–753. <https://doi.org/10.1146/annurev-economics-080218-025643>

32. Makarov, V., Ayvazyan, S., Afanasyev, M., Bakhtizin, A., Nanavyan, A. (2016). Modeling the Development of Regional Economy and an Innovation Space Efficiency. *Foresight and STI Governance*, Vol. 10, No. 3, 76–90. (In Russ.). <https://doi.org/10.17323/1995-459X.2016.3.76.90>

33. Zhukov, R.A., Plinskaya, M.A., Manokhin, E.V. (2023). Assessment of the Regions Functioning Based on Production Functions with the Above Cost Factors. *Journal of Applied Economic Research*, Vol. 22, No. 3, 657–682. (In Russ.). <https://doi.org/10.15826/vestnik.2023.22.3.027>

34. Zhukov, R.A. (2020). An approach to assessing the functioning of hierarchical socio-economic systems and decision-making based on the EFRA software package. *Business Informatics*, Vol. 14, No. 3, 82–95. (In Russ.). <https://doi.org/10.17323/2587-814X.2020.3.82.95>

35. Zhukov, R.A., Prokopchina, S.V., Bormotov I. V., Manokhin E. V., Rudneva I. D. (2024). The model of the value system of Russian youth based on Bayesian intellectual measurements. *MIR (Modernization. Innovation. Research)*, Vol. 15, No. 1, 96–114. (In Russ.). <https://doi.org/10.18184/2079-4665.2024.15.1.96-114>

INFORMATION ABOUT AUTHORS

Roman Aleksandrovich Zhukov

Doctor of Economics, Associate Professor, Researcher of The Financial University under the Government of the Russian Federation (Tula Branch), Tula, Russia (300012, Tula, Oruzheynaya street, 1a); ORCID <https://orcid.org/0000-0002-2280-307X> e-mail: pluszh@mail.ru

Svetlana Vasilievna Prokopchina

Doctor of Technical Sciences, Professor, Department of System Analysis in Economics, Financial University under the Government of the Russian Federation, Moscow, Russia (125167, Moscow, Leningradsky Prospekt, 49/2); ORCID <https://orcid.org/0000-0001-5500-2781> e-mail: svprokopchina@mail.ru

Maria Aleksandrovna Plinskaya

Master Student, The Financial University under the Government of the Russian Federation (Tula Branch), Tula, Russia (300012, Tula, Oruzheynaya street, 1a); ORCID <https://orcid.org/0000-0002-1307-0935> e-mail: plinskaya@gmail.com

Maria Anatolievna Zhelunitsina

Master Student, The Financial University under the Government of the Russian Federation (Tula Branch), Tula, Russia (300012, Tula, Oruzheynaya street, 1a); ORCID <https://orcid.org/0009-0006-3129-2749> e-mail: maria202001@yandex.ru

ACKNOWLEDGMENTS

The study was carried out at the expense of a grant from the Russian Science Foundation № 24-28-20020, <https://rscf.ru/project/24-28-20020/> and Tula region.

FOR CITATION

Zhukov, R.A., Prokopchina, S.V., Plinskaya, M.A., Zhelunitsina, M.A. (2024). Modeling of Functional Relationships of Regional Economic Systems Based on Small Samples Based on Bayesian Intelligent Measurements. *Journal of Applied Economic Research*, Vol. 23, No. 3, 721–750. <https://doi.org/10.15826/vestnik.2024.23.3.029>

ARTICLE INFO

Received May 26, 2024; Revised June 18, 2024; Accepted June 22, 2024.

