

Корректировка параметров регрессионной модели на основе экспертной информации об изменении значимости предикторов на предыстории

С. И. Носков, Д. Е. Баянов

Иркутский государственный университет путей сообщения, г. Иркутск

Аннотация: В работе предложен алгоритмический способ учета при регрессионном моделировании сложных объектов любой природы наряду со статистической экспертной информацией об изменении значимости предикторов на предыстории. В качестве методов расчета неизвестных параметров модели использованы метод смешанного оценивания и непрерывная форма метода максимальной согласованности. Первый из них основан на возможном совмещении методов наименьших модулей и антиробастного оценивания, каждый из которых «работает» на «своей» подвыборке исходной выборки данных. Второй предназначен для повышения согласованности в изменении расчетных и заданных значений зависимой переменной. Реализация совместного использования при моделировании статистической и экспертной информации для этих методов сводится к решению соответствующих задач линейного программирования. Предложенный способ реализован при построении линейной регрессионной модели патентной активности в России.

Ключевые слова: регрессионная модель, экспертная информация, методы смешанного оценивания и максимальной согласованности, патентная активность.

Для цитирования: Носков С. И., Баянов Д. Е. Корректировка параметров регрессионной модели на основе экспертной информации об изменении значимости предикторов на предыстории // Вестник СибГУТИ. 2025. Т. 19, № 1. С. 11–19. <https://doi.org/10.55648/1998-6920-2025-19-1-11-19>



Контент доступен под лицензией
Creative Commons Attribution 4.0
License

© Носков С. И., Баянов Д. Е., 2025

Статья поступила в редакцию 09.04.2024;
принята к публикации 31.04.2024.

1. Введение

При построении регрессионных моделей сложных объектов различной природы иногда наряду со статистической используют еще и разнородную экспертную и дополнительную информацию. Так, в работе [1] при построении модели регрессии землепользования для исследования уровня непрерывного звукового давления в пиковые дневные периоды на скоростной автомагистрали долины Нью-Кланг в Шах-Аламе, Малайзия, используется дополнительная информация: различные показатели шума, включая скорость ветра, цифровая модель рельефа, тип землепользования (жилой дом, промышленный объект или природный заповедник), плотность жилой застройки, тип дороги (скоростная автомагистраль, основная или второстепенная) и интенсивность движения, средний уровень шума. В статье [2] модель логистической регрессии была применена для картирования территорий, подверженных оползням, в центральной части бассейна реки Заба на горных склонах Юго-западной провинции Западного Азербайджана в Иране. При этом

были задействованы размытые данные о предрасположенности участков территории к оползням. Исследование [3] посвящено разработке двух новых процедур тестирования малых выборок на основе стандартного информационного критерия Шварца. В [4] изучается проблема составления комплексной карты подверженности оползням в окрестности города Мизунами, Япония, с целью уменьшить возможный ущерб населению и инфраструктуре. При этом привлечены методы построения логистической регрессии, двумерный статистический анализ и сплайновые модели многомерной адаптивной регрессии. В качестве дополнительной информации привлечены аэрофотоснимки, предоставленные NIED (Национальный научно-исследовательский институт наук о Земле и предотвращения стихийных бедствий, Япония). В работе [5] методы регрессионного моделирования используются при анализе производительности систем теплоэлектростанций (ТЭЦ). Построены модели индексов энергопотребления зданий, параметров проектирования и производительности. При этом привлечена информация об особенностях семи климатических зон. Среди типов изученных зданий выделены крупные гостиницы, больницы и амбулаторные учреждения здравоохранения. Выявлено, что использование накопителей тепловой энергии может существенно повысить комплексную производительность ТЭЦ. С помощью метода пошаговой регрессии для построения регрессионных моделей и определения наиболее значимых факторов изучена взаимосвязь между индексами спроса на энергию и проектными и эксплуатационными параметрами ТЭЦ. В [6] разработана двухуровневая модель регрессии независимых компонент (ICR) для многомерной спектроскопической калибровки. Для улучшения качества калибровки предлагается ансамблевая форма модели ICR. На первом уровне строятся различные подпространства на основе независимой компонентной декомпозиции исходного пространства данных. При этом путем определения связанного индекса для моделирования ICR выбираются наиболее важные с точки зрения исследователей переменные в каждом подпространстве, которые образуют второй уровень. Стратегия байесовского вывода позволяет получить вероятностную комбинацию результатов калибровки, полученных для разных подпространств. Для оценки эффективности предлагаемого подхода проводятся исследования на наборе эталонных спектральных данных. Статья [7] посвящена проблеме последовательной оценки коэффициентов регрессии в многомерной модели ошибок линейных ультраструктурных измерений, когда априори доступна дополнительная экспертная информация о коэффициентах регрессии. Такая дополнительная информация выражается в виде стохастических линейных ограничений. Используя эти ограничения, заданные априори, представлена методология для получения непротиворечивых оценок коэффициентов регрессии по двум типам дополнительной информации отдельно, а именно: ковариационная матрица ошибок измерения и матрица надежности, связанная с объясняющими переменными. Предполагается, что ошибки измерения не обязательно имеют нормальное распределение. Асимптотические свойства предлагаемых оценок выводятся и анализируются аналитически, а также численно с помощью метода Монте-Карло. В исследовании [8] разработаны региональные регрессионные модели для прогнозирования 13 перцентилей стока для групп бассейнов, сгруппированных на основе физических и климатических характеристик. Изучалось, как количество и информационное наполнение независимых переменных влияет на эффективность модели, а также сравнивались подходы к выбору переменных, основанные на данных с привлечением экспертных оценок. В [9] географически взвешенная регрессия была применена для получения пространственной температурной структуры в городе Вроцлав на юго-западе Польши. Полученные результаты сравнивались с моделями глобальной регрессии с использованием различных статистических процедур, включая скорректированный информационный критерий Акаике, коэффициент детерминации, дисперсионный анализ и индекс Морана. Было обнаружено, что учет пространственно интерполированных остатков приводит к значительному улучшению результатов интерполяции для обоих подходов.

2. Материалы и методы

Рассмотрим регрессионную модель (уравнение, зависимость):

$$y_k = \sum_{i=1}^m \alpha_i x_{ki} + \varepsilon_k, k = \overline{1, n}, \quad (1)$$

где y – объясняемая (выходная), а x_i – i -ая объясняющая (входная) переменная; α_i – i -ый подлежащий оцениванию параметр; ε_k – ошибки аппроксимации, k – номер наблюдения, n – число наблюдений (длина выборки).

Представим уравнение (1) в векторной форме:

$$y = X\alpha + \varepsilon, \quad (2)$$

где $y = (y_1, \dots, y_n)^T$, $\alpha = (\alpha_1, \dots, \alpha_m)^T$, $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T$, X – матрица ($n \times m$) с компонентами x_{ki} .

Заметим, что уравнение (1) имеет детерминированный характер.

Общеизвестно, что значения параметров α_i , $i = 1, m$, рассчитанные с помощью одного из методов регрессионного анализа, отражают средние тенденции во взаимосвязях между выходной и входными переменными, сложившиеся на предыстории исследуемого процесса (см., например, [10]).

Одним из основных направлений применения регрессионных моделей является прогнозирование на их основе будущих значений выходных переменных (см, в частности, [11–13]). При этом тенденции функционирования исследуемого объекта на периодах основания и упреждения прогноза могут не совпадать, что вызывает необходимость корректировки оценок параметров модели (1), как это сделано, например, в [14, 15] на основе учета специальным образом сформированной экспертной информации. Такая информация может наряду со статистической (т.е. выборкой (X, y)) иметь отношение к предыстории моделируемого процесса. Она, в частности, может быть связана с отражением в модели сравнительной значимости независимых переменных, как это продемонстрировано в [16] при использовании для поиска оценок параметров методов наименьших модулей (МНМ) и антиробастного оценивания (МАО).

Поставим задачу отражения в модели экспертной информации относительно изменения значимости (вкладов) группы независимых переменных на предыстории процесса для случая, когда оценивание неизвестных параметров модели (1) производится с помощью методов смешанного оценивания (МСО) [17] и максимальной согласованности (ММС) [18]. Дадим их краткое описание.

МСО в простейшей его трактовке является своего рода симбиозом МНМ и МАО, отражая их отличительные качества: первый не реагирует на выбросы в данных, второй, напротив, к ним сильно тяготеет. При его применении всё множество номеров наблюдений $P = \{1, 2, \dots, n\}$ разбивается на два непересекающихся подмножества P_1 и P_2 таким образом, что на подвыборке с номерами наблюдений из P_1 МСО «работает» как МНМ, а на подвыборке с номерами из P_2 – как МАО. При этом, естественно, должны выполняться соотношения:

$$P = P_1 \cup P_2, P_1 \cap P_2 = \emptyset.$$

Реализация МСО сводится к следующей задаче линейного программирования (ЛП):

$$\sum_{i=1}^m \alpha_i x_{ki} + u_k - v_k = y_k, k \in P \quad (3)$$

$$u_k + v_k - r \leq 0, k \in P_2, \quad (4)$$

$$u_k \geq 0, v_k \geq 0, k \in P, \quad (5)$$

$$\sum_{k \in P_1} (u_k + v_k) / |P_1| + r + \delta \sum_{k \in P_2} (u_k + v_k), \quad (6)$$

где $|P_1|$ – число элементов (мощность) в множестве P_1 , а δ – малое положительное число.

Применение ММС направлено на достижение максимальной согласованности в изменении фактических (заданных) y_k и расчетных (вычисленных по модели) $\hat{y}_k = y_k - \varepsilon_k$

значений зависимой переменной на предыстории. При этом дискретная форма ММС предполагает решение задачи оптимизации

$$\sum_{k=1}^{n-1} \sum_{s=k+1}^n \text{sign}((y_k - y_s)(\hat{y}_k - \hat{y}_s)) \rightarrow \max, \quad (7)$$

где

$$\text{sign}(a) = \begin{cases} 1, & a \geq 0 \\ 0, & \text{в противном случае,} \end{cases}$$

а непрерывная (НММС) – задачи

$$\sum_{k=1}^{n-1} \sum_{s=k+1}^n l_{ks} \rightarrow \min, \quad (8)$$

где

$$l_{ks} = \begin{cases} |\hat{y}_k - \hat{y}_s|, & (y_k - y_s)(\hat{y}_k - \hat{y}_s) < 0 \\ 0, & \text{в противном случае} \end{cases}$$

Реализация НММС также сводится к задаче ЛП следующим образом. Рассчитаем величины ω_{ks} , $k = \overline{1, n-1}$, $s = \overline{k+1, n}$ по правилу:

$$\omega_{ks} = \begin{cases} 1, & y_k - y_s > 0 \\ -1, & y_k - y_s < 0 \\ 0, & y_k - y_s = 0. \end{cases}$$

Тогда задача (8) сводится к задаче ЛП с ограничениями (3), (5),

$$\omega_{ks} \sum_{i=1}^m \alpha_i (x_{ki} - x_{si}) + l_{ks} \geq 0, \quad k = \overline{1, n-1}, s = \overline{k+1, n}, \quad (9)$$

$$l_{ks} \geq 0 \quad (10)$$

и целевой функцией

$$\rho \sum_{k=1}^n (u_k + v_k) + (1 - \rho) \sum_{k=1}^{n-1} \sum_{s=k+1}^n l_{ks} \rightarrow \min, \quad (11)$$

где число $\rho \in (0, 1)$ задает уровень компромисса между МНМ и НММС.

3. Формализация экспертных высказываний

Экспертные высказывания могут иметь как упрощенный, так и весьма сложный смысловой характер. Пусть, например, исследователь на основе своих профессиональных знаний и опыта может сформулировать свои суждения относительно некоторых независимых переменных следующим образом: значимость i -ой переменной, выраженная ее вкладом в правую часть модели (1), на наблюдениях с номерами из индексного множества R^1 должна превышать значимость j -ой переменной в h^1 раз, а на наблюдениях из множества R^2 – в h^2 раз. При этом $R^1 \subset P$, $R^2 \subset P$, $R^1 \cap R^2 = \emptyset$.

Пусть также, следуя [16] и учитывая закономерности функционирования анализируемого процесса или объекта, заранее известно направление влияния каждой переменной x_i на выходной фактор y , что позволяет назначить числа b_i , $i \in \{1, 2, \dots, m\}$ по правилу:

$$b_i = \begin{cases} 1, & \text{если с ростом значений } i\text{-ой переменной значение } y \text{ возрастает} \\ -1, & \text{если с ростом значений } i\text{-ой переменной значение } y \text{ убывает} \\ 0, & \text{если направление влияния } i\text{-ой переменной на } y \text{ может быть любым.} \end{cases}$$

Тогда приведенные выше вербальные экспертные высказывания относительно сравнительной значимости одних и тех же переменных на различных участках выборки данных могут быть формализованы путем задания следующих линейных ограничений на параметры модели (1):

$$\alpha_j \bar{x}_i^1 \geq h^1 \alpha_j \bar{x}_j^1, \quad (12)$$

$$\alpha_j \bar{x}_i^2 \geq h^2 \alpha_j \bar{x}_j^2, \quad (13)$$

$$b_j \alpha_j \geq 0, j \in \{1, 2, \dots, m\} \quad (14)$$

Здесь \bar{x}_j^1 и \bar{x}_j^2 – средние значения j -ой переменной на индексных множествах R^1 и R^2 соответственно:

$$\bar{x}_j^1 = \frac{1}{|R^1|} \sum_{k \in R^1} x_{kj},$$

$$\bar{x}_j^2 = \frac{1}{|R^2|} \sum_{k \in R^2} x_{kj}$$

Таким образом, применение МСО и НММС на основе совмещения статистической и заданной указанным образом экспертной информации приводит к задачам ЛП соответственно (3)–(5), (12)–(14), (6) и (3), (5), (9), (10), (12)–(14), (11). Обратим внимание, что размерность этих задач является, как правило, вполне приемлемой для выборок, соответствующих реальным объектам.

4. Результаты

Применим описанный способ учета экспертной информации при моделировании для построения регрессионной модели патентной активности в России на основе ретроспективной информации за 2010–2020 гг. [19]. Введем следующие обозначения:

y – число отечественных патентных заявок на промышленные образцы, шт.;

x_1 – используемые передовые производственные технологии, ед.;

x_2 – внутренние текущие затраты на фундаментальные исследования, млн руб.

Будем строить линейную регрессионную двухфакторную модель со свободным членом:

$$y_k = \alpha_0 + \alpha_1 x_{k1} + \alpha_2 x_{k2} + \varepsilon_k, k = \overline{1, 11} \quad (15)$$

с помощью МСО, НММС и описанного выше способа корректировки оценок параметров по отношению к этим методам.

Зададим необходимую дополнительную и экспертную информацию следующим образом:

$$P_1 = \{1, \dots, 5\}, P_2 = \{6, \dots, 11\},$$

$$\rho = 0.5,$$

$$R^1 = \{1, 2, 3\}, R^2 = \{9, 10, 11\},$$

$$b = (1, 1),$$

$$h^1 = 4, h^2 = 6.$$

Для построения трех версий модели (15) воспользуемся размещенной в Интернете в свободном доступе весьма эффективной программой решения задач ЛП LPsolve (см., в частности, [20–22]).

В результате получим:

1. МСО:

$$y_k = -2111.1 + 0.0159 x_{k1} + 0.0084 x_{k2} + \varepsilon_k, k = \overline{1, 11}. \quad (16)$$

2. НММС:

$$y_k = -1150.8 + 0.0109 x_{k1} + 0.0095 x_{k2} + \varepsilon_k, k = \overline{1, 11}. \quad (17)$$

3. МСО (после корректировки):

$$y_k = -2247.7 + 0.0171 x_{k1} + 0.0073 x_{k2} + \varepsilon_k, k = \overline{1, 11}. \quad (18)$$

4. НММС (после корректировки):

$$y_k = -1362.7 + 0.0128 x_{k1} + 0.0084 x_{k2} + \varepsilon_k, k = \overline{1, 11}. \quad (19)$$

Таким образом, учет экспертных высказываний привел к некоторому росту значения параметра α_1 в модели (18) по сравнению с моделью (16) и, соответственно, к уменьшению значения параметра α_2 в модели (19) по сравнению с (17).

5. Обсуждение и выводы

В работе представлен алгоритмический способ учета при регрессионном моделировании сложных объектов экспертной информации об изменении значимости предикторов на предыстории. В качестве методов вычисления неизвестных параметров модели использованы метод смешанного оценивания и непрерывная форма метода максимальной согласованности. Первый из них в простейшей версии основан на совмещении методов наименьших модулей и антиробастного оценивания, каждый из которых «работает» на «своей» подвыборке исходной выборки данных. Вторым предназначен для повышения согласованности в изменении расчетных и заданных значений зависимой переменной.

Формализовано экспертное высказывание относительно изменения значимости независимых переменных в виде линейных ограничений на параметры модели.

Реализация совместного использования при моделировании статистической и экспертной информации при применении обоих методов сводится к решению соответствующих задач линейного программирования.

Предложенный способ реализован при построении линейной регрессионной модели патентной активности в России.

Литература

1. *Adulaimi A. A. A. Traffic Noise Modelling Using Land Use Regression Model Based on Machine Learning, Statistical Regression and GIS // Energies, Basel, 2021. V. 14. № 16. 5095 p.*
2. *Shahabi H., Hashim M., Ahmad B. B. Remote sensing and GIS-based landslide susceptibility mapping using frequency ratio, logistic regression, and fuzzy logic methods at the central Zab basin // Iran. Environmental Earth Sciences, 2015. V. 73. pp. 8647–8668.*
3. *Nosek, Konrad. Schwarz Information Criterion Based Tests for a Change-Point in Regression Models // Statistical papers, Berlin, Germany, 2010. V. 51. pp. 915–929.*
4. *Wang, Liang-Jie. Landslide Susceptibility Mapping in Mizunami City, Japan: A Comparison between Logistic Regression, Bivariate Statistical Analysis and Multivariate Adaptive Regression Spline Models // Catena, Giessen, 2015. V. 135. pp. 271–282.*
5. *Yang, G., Zheng, C. Y., Zhai, X. Q. Influence analysis of building energy demands on the optimal design and performance of CCHP system by using statistical analysis. // Energy and Buildings, 2017. V. 153. pp. 297–316.*
6. *Zheng, J., Song, Z. Two-level independent component regression model for multivariate spectroscopic calibration // Chemometrics and Intelligent Laboratory Systems, 2016. V. 155. pp. 160–169.*
7. *Shalabh, Garg, G., Misra, N. Consistent estimation of regression coefficients in ultrastructural measurement error model using stochastic prior information // Statistical Papers, Berlin, Germany, 2010. V. 51. pp. 717–748.*
8. *Fouad G., Skupin A., Tague C. L. Regional regression models of percentile flows for the contiguous United States: Expert versus data-driven independent variable selection // Journal of Hydrology. Regional Studies, 2018. V. 17. pp. 64–82.*
9. *Szymanowski, M., Kryza, M. Local regression models for spatial interpolation of urban heat island—an example from Wrocław // Theoretical and Applied Climatology, SW Poland, 2012. V. 108. pp. 53–71.*
10. *Дрейнер Н., Смит С. Прикладной регрессионный анализ. М.: Диалектика. 912 с.*
11. *Сизяков Н. П., Шестопалова О. Л. Прогнозирование соответствия характеристик космических средств предъявляемым требованиям на основе использования нечеткой регрессионной модели // Информация и космос. 2010. № 1. С. 133–135.*
12. *Бойко Н. С., Лошаков А. В. Прогнозирование показателей безопасности полётов с учётом внедрения управленческого решения на основе регрессионных моделей //*

- Вестник Ульяновского государственного технического университета. 2022. № 2 (98). С. 74-76.
13. *Геращенко И. П.* Методы прогнозирования в регрессионных и адаптивных моделях при анализе динамических рядов // Математические структуры и моделирование. 2000. № 5. С. 140-154.
 14. *Головченко В. Б., Носков С. И.* Выбор класса линейной по параметрам регрессии на основе экспертных высказываний // Кибернетика и системный анализ. 1992. № 5. С.109-115.
 15. *Головченко В. Б., Носков С. И.* Комбинирование прогнозов с учетом экспертной информации // Автоматика и телемеханика. 1992. № 11. С.109-117.
 16. *Носков С. И.* Построение линейной регрессии с учетом экспертной информации относительно сравнительной значимости переменных // Вестник Технологического университета. 2021. Т. 24. № 2. С. 83-86.
 17. *Носков С. И.* Метод смешанного оценивания параметров линейной регрессии: особенности применения // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. 2021. № 1. С. 126-132.
 18. *Носков С. И.* Метод максимальной согласованности в регрессионном анализе // Известия Тульского государственного университета. Технические науки. 2021. № 10. С. 380-385.
 19. *Носков С. И., Пашков Д. В.* Реализация конкурса регрессионных моделей эффективности интеллектуальной деятельности // Электронный сетевой политематический журнал «Научные труды КубГТУ». 2022. № 6. С. 40–51.
 20. *Шипицына Р. Е., Витвицкий Е. Е.* Сравнение удобства использования программных продуктов при решении транспортной задачи линейного программирования: LPSolve IDE и Microsoft Excel // В сборнике: Образование. Транспорт. Инновации. Строительство. Сборник материалов V Национальной научно-практической конференции. Омск, 2022. С. 250-254.
 21. *Арсланов М. З.* Математические модели задачи об упаковке единичных квадратов // Проблемы информатики. 2015. № 4 (29). С. 5-13.
 22. *Первун О. Е.* Оптимизация и исследование задач линейного программирования средствами приложения R // Информационно-компьютерные технологии в экономике, образовании и социальной сфере. 2022. № 4 (38). С. 87-92.

Носков Сергей Иванович

д.т.н., профессор кафедры информационных систем и защиты информации, Иркутский государственный университет путей сообщения (ИрГУПС, 664074, г. Иркутск, ул. Чернышевского, д. 15), тел. +7 395 2638 322, e-mail: sergey.noskov.57@mail.ru, ORCID ID: 0000-0003-4097-2720.

Баянов Даниил Евгеньевич

магистрант кафедры информационных систем и защиты информации, Иркутский государственный университет путей сообщения (ИрГУПС, 664074, г. Иркутск, ул. Чернышевского, д. 15), e-mail: danya.bayanov.2001@mail.ru, ORCID ID: 0009-0007-5315-3042.

Авторы прочитали и одобрили окончательный вариант рукописи.

Авторы заявляют об отсутствии конфликта интересов.

Вклад соавторов: Каждый автор внес равную долю участия как во все этапы проводимого теоретического исследования, так и при написании разделов данной статьи.

Correction of regression model parameters based on expert information on changes in the significance of predictors in the background

Sergei I. Noskov, Daniel E. Bayanov

Abstract: The paper proposes an algorithmic method of accounting for regression modeling of complex objects of any nature, along with statistical expert information on changes in the significance of predictors in the background. The method of mixed estimation and the continuous form of the maximum consistency method are used as methods for calculating unknown parameters of the model. The first of them is based on a combination of the methods of smallest modules and anti-robust estimation, each of which "works" on "its own" subsample of the original data sample. The second one is designed to increase consistency in changing the calculated and set values of the dependent variable. The implementation of the joint use of statistical and expert information in modeling for these methods is reduced to solving the corresponding linear programming problems. The proposed method is implemented in the construction of a linear regression model of patent activity in Russia.

Keywords: regression model, expert information, methods of mixed assessment and maximum consistency, patent activity.

For citation: Noskov S. I., Bayanov D. E. Correction of regression model parameters based on expert information on changes in the significance of predictors in the background. *Vestnik SibGUTI*, 2025, vol. 19, no. 1, pp. 11-19. <https://doi.org/10.55648/1998-6920-2025-19-1-11-19>.



Content is available under the license
Creative Commons Attribution 4.0
License

© S. I. Noskov, D. E. Bayanov, 2025

The article was submitted: 09.04.2024;
accepted for publication 22.05.2024.

References

1. *Adulaimi A. A. A.* Traffic Noise Modelling Using Land Use Regression Model Based on Machine Learning, *Statistical Regression and GIS // Energies*, Basel, 2021, vol. 14, no. 16, 5095 p.
2. *Shahabi H., Hashim M., Ahmad B. B.* Remote sensing and GIS-based landslide susceptibility mapping using frequency ratio, logistic regression, and fuzzy logic methods at the central Zab basin // *Iran. Environmental Earth Sciences*, 2015, vol. 73, pp. 8647–8668.
3. *Nosek, Konrad.* Schwarz Information Criterion Based Tests for a Change-Point in Regression Models // *Statistical papers*, Berlin, Germany, 2010, vol. 51, pp. 915–929.
4. *Wang, Liang-Jie et al.* Landslide Susceptibility Mapping in Mizunami City, Japan: A Comparison between Logistic Regression, Bivariate Statistical Analysis and Multivariate Adaptive Regression Spline Models // *Catena*, Giessen, 2015, vol. 135, pp. 271–282.
5. *Yang, G., Zheng, C. Y., Zhai, X. Q.* Influence analysis of building energy demands on the optimal design and performance of CCHP system by using statistical analysis. // *Energy and Buildings*, 2017, vol. 153, pp. 297–316.
6. *Zheng, J., Song, Z.* Two-level independent component regression model for multivariate spectroscopic calibration // *Chemometrics and Intelligent Laboratory Systems*, 2016, vol. 155, pp. 160–169.
7. *Shalabh, Garg, G., Misra, N.* Consistent estimation of regression coefficients in ultrastructural measurement error model using stochastic prior information // *Statistical Papers*, Berlin, Germany, 2010, vol. 51, pp. 717–748.
8. *Fouad G., Skupin A., Tague C. L.* Regional regression models of percentile flows for the contiguous United States: Expert versus data-driven independent variable selection // *Journal of Hydrology. Regional Studies*, 2018, vol. 17, pp. 64-82.
9. *Szymanowski, M., Kryza, M.* Local regression models for spatial interpolation of urban heat island—an example from Wrocław // *Theoretical and Applied Climatology*, SW Poland, 2012, vol. 108, pp. 53–71.

10. *Draper N., Smith H.* Prikladnoj regressionnyj analiz [Applied regression analysis]. M.: Dialektika. 912 p.
11. *Sizjakov N. P., Shestopalova O. L.* Prognozirovanie sootvetstvija harakteristik kosmicheskikh sredstv predjavljaemym trebovanijam na osnove ispolzovanija nechetkoj regressionnoj modeli [Forecasting the compliance of the characteristics of space assets with the requirements based on the use of a fuzzy regression model]. Informacija i kosmos. 2010, no. 1, pp. 133-135.
12. *Bojko N. S., Loshakov A. V.* Prognozirovanie pokazatelej bezopasnosti poljotov s uchjotom vnedrenija upravlencheskogo reshenija na osnove regressionnyh modelej [Forecasting flight safety indicators, taking into account the implementation of a management solution based on regression models]. Vestnik Ul'janovskogo gosudarstvennogo tehničeskogo universiteta. 2022, no. 2 (98), pp. 74-76.
13. *Gerashhenko I. P.* Metody prognozirovanija v regressionnyh i adaptivnyh modeljah pri analize dinamičeskikh rjadov [Forecasting methods in regression and adaptive models in the analysis of dynamic series]. Matematicheskie struktury i modelirovanie. 2000, no. 5, pp. 140-154.
14. *Golovchenko V. B., Noskov S. I.* Vybor klassa linejnoj po parametram regressii na osnove jekspertnyh vyskazyvanij [Choosing a linear regression class based on expert statements]. Kibernetika i sistemnyj analiz. 1992, no. 5, pp. 109-115.
15. *Golovchenko V. B., Noskov S. I.* Kombinirovanie prognozov s uchjotom jekspertnoj informacii [Combining forecasts with expert information]. Avtomatika i telemekhanika. 1992, no. 11, pp. 109-117.
16. *Noskov S. I.* Postroenie linejnoj regressii s uchjotom jekspertnoj informacii odnositelno sravnitelnoj znachimosti peremennyh [The construction of a linear regression taking into account expert information on the comparative significance of variables]. Vestnik Tehnologičeskogo universiteta. 2021, Vol. 24, no. 2, pp. 83-86.
17. *Noskov S. I.* Metod smeshannogo ocenivanija parametrov linejnoj regressii: osobennosti primenenija [The method of mixed estimation of linear regression parameters: application features]. Vestnik Voronezhskogo gosudarstvennogo universiteta. Serias: Sistemnyj analiz i informacionnye tehnologii. 2021, no. 1, pp. 126-132.
18. *Noskov S. I.* Metod maksimalnoj soglasovannosti v regressionnom analize [The method of maximum consistency in regression analysis]. Izvestija Tul'skogo gosudarstvennogo universiteta. Tehničeskije nauki. 2021, no. 10, pp. 380-385.
19. *Noskov S. I., Pashkov D. V.* Realizacija konkursa regressionnyh modelej jeffektivnosti intellektual'noj dejatel'nosti [Implementation of the competition for regression models of intellectual activity efficiency]. Jelektronnyj setevoj politematicheskiy zhurnal «Nauchnye trudy KubGTU». 2022, no. 6, pp. 40–51.
20. *Shipicyna R. E., Vitvickij E. E.* Sravnenie udobstva ispol'zovanija programmnyh produktov pri reshenii transportnoj zadachi linejnogo programmirovanija: LPSolve IDE i Microsoft Excel [Comparison of the usability of software products in solving the transport problem of linear programming: LPSolve IDE and Microsoft Excel]. Obrazovanie. Transport. Innovacii. Stroitel'stvo. Sbornik materialov V Nacional'noj nauchno-praktičeskoi konferencii. Omsk, 2022, pp. 250-254.
21. *Arslanov M. Z.* Matematicheskie modeli zadachi ob upakovke edinichnyh kvadratov [Mathematical models of the problem of packing unit squares]. Problemy informatiki. 2015, no. 4 (29), pp. 5-13.
22. *Pervun O. E.* Optimizacija i issledovanie zadach linejnogo programmirovanija sredstvami prilozhenija R [Optimization and research of linear programming problems by means of the R application]. Informacionno-komp'juternye tehnologii v jekonomike, obrazovanii i social'noj sfere. 2022, no. 4 (38), pp. 87-92.

Noskov Sergei Ivanovich

Dr. of Sci. (Engineering), Professor; Professor of the Department of Information systems and information protection, Irkutsk State Transport University (IrGUPS, 664074, Irkutsk, Chernyshevskogo str., 15), e-mail: sergey.noskov.57@mail.ru, ORCID ID: 0000-0001-8148-1843.

Bayanov Daniel Evgenyevich

undergraduate student of the Department of Information systems and information protection, Irkutsk State Transport University (IrGUPS, 664074, Irkutsk, Chernyshevskogo str., 15), e-mail: danya.bayanov.2001@mail.ru, ORCID ID: 0009-0007-5315-3042.