

А. В. Караулова¹, М. П. Базилевский¹

¹ Иркутский государственный университет путей сообщения, г. Иркутск, Российская Федерация

ПРИМЕНЕНИЕ РЕГРЕССИОННОГО АНАЛИЗА ПРИ РЕШЕНИИ РЕАЛЬНЫХ ЗАДАЧ ТЕХНИЧЕСКОГО ХАРАКТЕРА

Аннотация. Статья посвящена обзору статей по вопросу применения регрессионного анализа в технической сфере современного мира. Рассмотрены следующие модели: модель, отражающая взаимосвязи характеристик обрабатываемого материала и технологических режимов лазерного упрочнения с глубиной поверхностного слоя; модель плотности имитационного шипика; модель длительности опорожнения ёмкости смесителя от остатков (выгрузки) при приготовлении зерновой смеси при производстве комбикормов; модель для вычисления температуры вспышки в закрытом тигле (ТВЗТ); модель влияния факторов условий эксплуатации на температуру поверхности автомобильного генератора; модель, используемая при разработке новой технологии высокообъемной комбинированной пряжи с использованием токов СВЧ; модель для мониторинга графиков движения поездов по критерию энергетической эффективности; модель, способствующая обеспечить точность формы отверстий при чистовой фрезерной обработке; модель для определения рациональных конструктивных параметров рифлёного валька. Почти все изученные модели являются нелинейными. На основе анализа рассмотренных источников был сделан вывод, что авторы при построении моделей практически не уделяют внимания их сложности.

Ключевые слова: анализ данных, регрессионная модель, эконометрика, моделирование, нелинейность.

A.V. Karaulova¹, M.P. Bazilevskiy¹

¹ Irkutsk State Transport University, Irkutsk, Russia

APPLICATION OF REGRESSION ANALYSIS IN SOLVING REAL TECHNICAL PROBLEMS

Abstract. The article is devoted to the review of articles on the application of regression analysis in the technical sphere of the modern world. Consider the following model: model that reflects the relationship of characteristics of the processed material and technological conditions of laser hardening depth of the surface layer; model density imitation lard; the duration model mixer from the residue (s) in the preparation of grain mixture in the production of feeds; the model for calculating the flash point in closed crucible (TWST); a model of the influence of operating conditions on the surface temperature of an automobile generator; a model used in the development of a new technology for high-volume combined yarn using microwave currents; a model for monitoring train schedules according to the criterion of energy efficiency; a model that helps to ensure the accuracy of the shape of holes during finishing milling; a model for determining rational design parameters of a grooved roll. Almost all the models studied are nonlinear. Based on the analysis of the sources considered, it was concluded that the authors do not pay much attention to their complexity when constructing models.

Keywords: data analysis, regression model, econometrics, modeling, nonlinearity.

Введение

Актуальной на сегодняшний день задачей является задача эффективной обработки и анализа статистических данных с целью извлечения из них неизвестных ранее полезных зна-

ний. Одним из инструментов анализа данных является регрессионный анализ [1–3] – статистический метод, позволяющий выявлять соотношения между зависимой переменной и одной или несколькими независимыми переменными. В результате применения этого анализа исследователь получает так называемую регрессионную модель, которую можно интерпретировать и использовать для получения прогнозных значений зависимой переменной. Регрессионные модели также часто отождествляются с эконометрическими моделями, являющимися центральным звеном науки эконометрики [4]. В последнем случае оценивание регрессии осуществляется по экономическим данным. Построению эконометрических моделей посвящено множество работ (см, например, [5–9]). Данная статья носит обзорный характер, а её целью является демонстрация того, что регрессионный анализ успешно применяется не только в экономике, но и при решении реальных задач технического характера.

Анализ литературных источников

Регрессионная модель в общем виде записывается так:

$$y_i = F(\alpha; x_{i1}; x_{i2}; \dots; x_{im}) + \varepsilon_i, \quad i = \overline{1, n}, \quad (1)$$

где $F(x_{i1}; x_{i2}; \dots; x_{im})$ – некоторая вещественная аппроксимирующая функция;

$\varepsilon_i, i = \overline{1, n}$ – ошибки аппроксимации;

α – вектор неизвестных параметров модели,

$y_i, i = \overline{1, n}$ – значения зависимой (объясняемой, выходной) переменной;

$x_{ij}, i = \overline{1, n}; j = \overline{1, m}$ – значения независимой (объясняющей, входной) переменной;

n – количество наблюдений;

m – количество независимых переменных.

Рассмотрим далее конкретные приложения регрессионного анализа в технике.

В работе [10] построены три регрессионные модели, отражающие взаимосвязи характеристик обрабатываемого материала и технологических режимов лазерного упрочнения с глубиной поверхностного слоя. Лучшей из них оказалась степенная модель:

$$y = 0,03x_1^{0,38} x_2^{0,54} x_3^{0,69} x_4^{-0,61}, \quad (2)$$

где y – глубина закалки, мм;

x_1 – размер пятна лазера, мм;

x_2 – время закалки, с;

x_3 – коэффициент температурного расширения, °C⁻¹;

x_4 – скорость, м/с.

С использованием модели (2) было проведено ранжирование технологических факторов по степени значимости:

- 1) длительность воздействия лазерного луча x_2 ;
- 2) скорость вращения детали x_4 ;
- 3) радиус лазерного луча x_1 ;
- 4) теплофизические характеристики материала x_3 .

Работа [11] посвящена построению регрессионной модели плотности имитационного шпика. Шпик – это один из видов мясного сырья, использующихся в производстве колбас. Модели строились с использованием многочленов k -го порядка.

Наилучшие результаты показал аппроксимационный многочлен 3-й степени:

$$y = 666,16 - 9,353x_1 - 191,181x_2 - 0,5x_1^2 + 52,572x_2^2 + 30,317x_1x_2 + \\ + 0,0215x_1^3 - 3,031x_2^3 - 0,823x_1^2x_2 + 1,272x_1x_2^2, \quad (3)$$

где y – плотность имитационного шпика, г;

x_1 – время выдержки, час;

x_2 – количество содержания многофункциональной смеси, %.

С помощью модели (3) установлено, что наиболее рациональным, с точки зрения производства, является использование многофункциональной смеси с массовой долей от 3 до 5 %.

Статья [12] посвящена проблеме смешивания зерновой смеси при производстве комбикормов. Приготовление таких смесей осуществляется в основном смесителями разнообразной конструкции. В результате исследования была получена следующая регрессионная модель:

$$T = 306,9 + 206,9\varepsilon^{0,15} - 134,2n^{0,11} - 173,9Z^{0,51} + 289,0(\varepsilon \cdot n)^{-58,8} + 413,2(\varepsilon \cdot Z)^{-50,06} + 670,3(n \cdot Z)^{-0,17} + \frac{1285,45}{\varepsilon} - \frac{672,56}{n} - \frac{1460,2}{Z}, \quad (4)$$

где T – длительность опорожнения емкости смесителя от остатков (выгрузки) приготовленной смеси, с;

ε – степень заполнения емкости, %;

n – частота вращения, мин⁻¹;

Z – количество лопастей, шт.

С использованием модели (4) установлено, что рост степени заполнения увеличивает длительность опорожнения емкости смесителя, а повышение частоты вращения мешалки и количество её лопастей сокращают длительность опорожнения.

В статье [13] рассматривается задача построения зависимости температуры вспышки дизельного топлива в закрытом тигле от его измеренных нормативных параметров. Так, построена многомерная линейная регрессионная модель для вычисления температуры вспышки в закрытом тигле (ТВЗТ) по температуре кипения 10% и 20% фракциям дизельного топлива

$$T = 0,66t_0 + 0,19t_{10} - 0,08t_{20} - 74,29, \quad (5)$$

где T – значение ТВЗТ, °С;

t_0 – температура кипения, °С;

t_{10} – температура выкипания 10% фракции;

t_{20} – температура выкипания 20% фракции.

Полученные результаты показали, что тяжелые фракции дизельного топлива не влияют на ТВЗТ. Результат работы может быть использован при анализе похожих данных в области нефтехимии, например, для построения регрессионной модели взаимосвязанных свойств бензинов.

В работе [14] исследуется зависимость изменения теплового состояния автомобильного генератора от климатических и эксплуатационных факторов. С помощью программы Table-Curve 3D была получена регрессионная модель влияния факторов условий эксплуатации на температуру поверхности генератора:

$$t_r = 51,0036 + 0,0131I_{\text{потр}}^2 - 0,0205v^{1,5} + 0,998t_b - 0,0014t_b^2, \quad (6)$$

где t_r – температура поверхности генератора, °С;

t_b – температура окружающего воздуха, °С;

$I_{\text{потр}}$ – ток потребителей, А;

v – скорость движения, км/ч.

С помощью регрессии (6) было установлено, что наибольшее влияние на температуру поверхности генератора оказывает температура воздуха окружающей среды. Получая информацию с датчика температуры поверхности генератора и сопоставляя её со значениями факторов модели (6), можно судить об исправности генератора.

В работе [15] разработана новая технология высокообъемной комбинированной пряжи с использованием токов СВЧ. В результате исследования была получена регрессионная модель

$$s = \frac{\tau \cdot P \cdot W}{\left((2,20 \cdot \tau + 193) \cdot (-0,197 \cdot P - 17) \cdot (0,515 \cdot 10^{-3} \cdot W - 5,08) \right)}, \quad (7)$$

где s – относительная усадка, %;

τ – время термообработки, с;

P – мощность излучения, Вт;

W – относительная влажность образцов до термообработки, %.

Анализ полученной зависимости (7) позволил сделать следующие выводы:

1) при одинаковых значениях режимных параметров процесса термообработки увеличение начальной влажности образцов приводит к повышению усадки;

2) одинаковых значений усадки при одном и том же времени обработки можно достичь различными комбинациями начальной влажности и мощности СВЧ излучения.

В статье [16] рассмотрены вопросы моделирования процессов электропотребления на тягу на участках с горным профилем пути. Имитационное моделирование проведено для серии вариантов с различными параметрами графика движения грузовых поездов и сохранением неизменными нормообразующих факторов. Структурная спецификация регрессионной модели имеет вид:

$$F(x) = a_{11} \cdot k_u + a_{12} \cdot k_{\text{умпак}} + a_{13} \cdot \frac{1}{1 + e^{-(\alpha - 0,5)}} + a_{14} \cdot n_{\text{ср}} + a_{15} \cdot \beta + a_{16} \cdot \gamma + a_{21} \cdot k_u^2 + a_{22} \cdot k_{\text{умпак}}^2 + a_{23} \cdot \frac{1}{1 + e^{-2(-\alpha + 0,5)}} + a_{24} \cdot n_{\text{ср}}^2 + a_{25} \cdot \beta^2 + a_{26} \cdot \gamma^2 + b, \quad (8)$$

где $F(x)$ – изменение объема электроэнергии на тягу;

k_u – коэффициент интенсивности для пакетной зоны графика движения;

α – относительное значение времени следования поездов в пакетах (пачках);

$k_{\text{умпак}}$ – коэффициент интенсивности для межпакетной зоны графика движения;

$n_{\text{ср}}$ – среднее количество поездов в пакете;

β – удельный вес остановок;

γ – удельный вес твердых ниток в графике;

$a_{ij}, i = \overline{1,2}; j = \overline{1,6}; b$ – коэффициенты модели.

Были найдены оптимальные оценки неизвестных параметров регрессии (8), что позволяет использовать её для мониторинга графиков движения поездов по критерию энергетической эффективности.

В работе [17] рассматривается задача обеспечения точности формы отверстий при чистовой фрезерной обработке на основе анализа процесса формообразования. В ней же было получено регрессионное уравнение, позволяющее определять значение погрешности врезания в зависимости от трех переменных:

$$\Delta_{\text{врез}} = 268,3452R_{\text{врез}} + 298,6397\alpha_{\text{врез}} + 291,1716F_{\text{врез}} - 1038,9263R_{\text{врез}} \cdot \alpha_{\text{врез}} - 997,953R_{\text{врез}} \cdot F_{\text{врез}} - 1089,0378\alpha_{\text{врез}} \cdot F_{\text{врез}} + 1875,0893R_{\text{врез}} \cdot \alpha_{\text{врез}} \cdot F_{\text{врез}}, \quad (9)$$

где $\Delta_{\text{врез}}$ – погрешность, возникающая в процессе врезания инструмента в заготовку, мкм;

$R_{\text{врез}}$ – радиус захода (врезания) фрезы;

$\alpha_{\text{врез}}$ – угол касательной к радиусу захода фрезы;

$F_{\text{врез}}$ – подача врезания.

Статья [18] посвящена исследованию технологии повышения наполнительной способности перовой массы, заключающейся в прокатке стержней перьев рифленным валком на плоскости. Уравнение искомой регрессионной зависимости имеет вид

$$y = -13,519 - 13,914p^2 - 15,391b^2 - 24161,18h^2 - 444,444b \cdot h + 212,5h \cdot p - 12,667p \cdot b, \quad (10)$$

где y – показатель перовой массы, состоящей из деформированных перьев, %;

p – шаг рифления, мм;

b – ширина вершины рифли, мм;

h – зазор между вершиной рифли и поверхности, мм.

С помощью регрессионной модели (10) были определены рациональные конструктивные параметры рифлёного валка $p = 3,165$ мм, $b = 0,337$ мм, $h = 0,052$ мм, обеспечивающие повышение показателя y на $23 \pm 5,5$ %.

Заключение

Проведенный в этой работе обзор современных литературных источников подтверждает, что регрессионный анализ действительно является эффективным инструментом при решении не только экономических, но и технических задач. Анализируя рассмотренные регрессионные модели (2) – (10), можно сделать вывод, что только регрессия (5) является линейной. Остальные модели носят ярко выраженный нелинейный характер, что связано со сложностью функционирования обсуждаемых технологических процессов и систем. При этом во всех рассмотренных работах авторы ставят во главу угла минимизацию ошибок аппроксимации модели, игнорируя сложность полученной нелинейной зависимости. Поэтому справедливо возникает задача оценки не только точности регрессии, но и её сложности, а также проблема нахождения некоторого компромисса между её точностью и сложностью. Для решения этих задач можно использовать математический аппарат, разработанный в [19,20]. Исследованию этих вопросов будут посвящены дальнейшие работы авторов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Носков С.И. Технология моделирования объектов с нестабильным функционированием и неопределенностью в данных. – Иркутск: Облформпечать, 1996. – 321 с.
2. Draper N.R., Smith H. Applied regression Analysis, 3rd edition. – John Wiley & Sons, 1998. – 736 p.
3. Mendenhall W., Sincich T.T. A second course in statistics: regression analysis, 8th edition. – Pearson, 2019. – 848 p.
4. Базилевский М.П., Гефан Г.Д. Эконометрика (продвинутый уровень): лабораторный практикум. – Иркутск: ИрГУПС, 2016. – 76 с.
5. Носков С.И., Базилевский М.П. Построение регрессионных моделей с использованием аппарата линейно-булевого программирования: монография. – Иркутск, ИрГУПС, 2018. – 176 с.
6. Базилевский М.П. МНК-оценивание параметров специфицированных на основе функций Леонтьева двухфакторных моделей регрессии // Южно-Сибирский научный вестник. 2019. – № 2 (26). – С. 66-70.
7. Баенхаева А.В., Базилевский М.П., Носков С.И. Моделирование валового регионального продукта Иркутской области на основе применения методики множественного оценивания регрессионных параметров // Фундаментальные исследования. 2016. – № 10-1. – С. 9-14.
8. Баенхаева А.В., Базилевский М.П., Носков С.И. Выбор структурной спецификации регрессионной модели валового регионального продукта Иркутской области // Информа-

онные технологии и проблемы математического моделирования сложных систем. 2016. – № 16. – С. 31-38.

9. Носков С.И., Кириллова Т.К. Регрессионная модель оценки влияния рекреационной деятельности на социально-экономическое развитие территории // Вестник Иркутского государственного технического университета. 2013. – № 9 (80). – С. 24-28.

10. Чигиринский Ю.Л., Щепетнов И.А., Чигиринская Н.В. Регрессионная модель процесса лазерного упрочнения // Фундаментальные исследования. 2015. – № 6-2. – С. 306-310.

11. Наумов В.А., Коржавина Ю.Н., Шибeko А.Г., Сингаев В.И., Альшевский Д.Л. Регрессионная модель плотности имитационного шпика // Известия КГТУ. 2018. – № 49. – С. 145-153.

12. Борисова М.В., Титов А.Ю., Новиков В.В., Коновалов В.В. Регрессионная модель производительности опорожнения тихоходного смесителя // Вестник Башкирского государственного аграрного университета. 2019. – № 2 (50). – С. 103-108.

13. Максимова А. Ю., Иванова А. А., Лозинский Н. С. Регрессионная модель для прогнозирования температуры вспышки дизельного топлива в закрытом тигле // Информатика и кибернетика. 2019. – № 4 (18). – С.5-13.

14. Пузаков А.В., Осаулко Я.Ю. Исследование влияния эксплуатационных факторов на тепловое состояние автомобильного генератора // Вестник Московского автомобильно-дорожного государственного технического университета (МАДИ). 2018. – № 1 (52). – С. 16-23.

15. Куландин А.С. Коган А.Г. Производство высокообъемной комбинированной пряжи с использованием токов СВЧ. // Моделирование в технике и экономике: сборник материалов Междунар. науч.-практич. конф. / ВГТУ, - Витебск, 2016. – с.101-103.

16. Незевак В. Л. Моделирование процессов электропотребления на тягу при изменении параметров графика движения поездов на электрифицированных участках с III-м и IV-м типом профиля пути // Современные технологии. Системный анализ. Моделирование. – 2019. – Т. 61, № 1. – С. 156–166.

17. Стельмаков В.А., Никитенко А.В., Давыдов В.М., Гимадеев М.Р. Обеспечение точности формы при чистовой обработке отверстий методом фрезерования // Информационные технологии XXI века: сборник научных трудов. 2017. – С. 502-509.

18. Чертов Ю.Е., Галатов К.С., Молев М.Д. Регрессионный анализ процесса деформации стержня пера рифленным валком // Современные проблемы науки и образования. 2015. – № 2-3. – С. 49.

19. Базилевский М.П. Критерии нелинейности многофакторных квазилинейных регрессий // Молодежь и наука: актуальные проблемы фундаментальных и прикладных исследований: сборник материалов II Всероссийской национальной научной конференции студентов, аспирантов и молодых ученых. 2019. С. 210-213.

20. Базилевский М.П. Критерии нелинейности квазилинейных регрессионных моделей // Моделирование, оптимизация и информационные технологии. – 2018. – Т. 6. – № 4 (23). – С. 185–195.

REFERENCES

1. Noskov S. I. Technology of modeling objects with unstable functioning and uncertainty in data. - Irkutsk: Oblinformpechat, 1996, 321 p

2. Draper N.R., Smith H. 1998. Applied regression Analysis, 3rd edition. John Wiley & Sons, 736p.

3. Mendenhall W., Sincich T.T. A second course in statistics: regression analysis, 8th edition. – Pearson, 2019. – 848 p.

4. Bazilevsky M. P., Gefan G. D. Econometrics (advanced level): laboratory session. Irkutsk [Irkutsk State Transport University], 2016. – 76 p.

5. Noskov S. I., Bazilevsky M. P. Building regression models using linear-Boolean programming: monograph. – Irkutsk [Irkutsk State Transport University], 2018. – 176 p.
6. Bazilevsky M. P. MNK-estimation of parameters of two-factor regression models specified on the basis of Leontiev functions // South Siberian scientific Bulletin, 2019, No 2 (26). pp. 66-70.
7. Baenkhaveva A.V., Bazilevsky M. P., Noskov S. I. Modeling of the gross regional product of the Irkutsk region based on the application of the method of multiple estimation of regression parameters. Fundamental research, 2016, No 10-, pp. 9-14.
8. Baenkhaveva A.V., Bazilevsky M. P., Noskov S. I. Choice of structural specification of the regression model of the gross regional product of the Irkutsk region // Information technologies and problems of mathematical modeling of complex systems, 2016, No 16, pp. 31-38.
9. Noskov S. I., Kirillova T. K. Regression model for assessing the impact of recreational activities on the socio-economic development of the territory // Bulletin of Irkutsk state technical University, 2013, No 9 (80), pp. 24-28.
10. Chigirinsky Yu. L., Shchepetnov I. A., Chigirinskaya N. V. Regression model of laser hardening process // Fundamental research, 2015, No 6-2, pp. 306-310.
11. Naumov V. A., Korzhavina Yu. N., Shibeko A. G., Singaev V. I., Alyshevsky D. L. a Regression model of the density of imitation lard // Proceedings of KGTU, 2018, No 49, pp. 145-153.
12. Borisova M. V., Titov A. Yu., Novikov V. V., Konovalov V. V. Regression model of low-speed mixer emptying performance // Bulletin of the Bashkir state agrarian University, 2019, No 2 (50), pp. 103-108.
13. Maximova A. Yu., Ivanova A. A., Lozinsky N. S. Regression model for predicting the flash point of diesel fuel in a closed crucible // Informatics and Cybernetics, 2019, Vol. 18, No. 4, pp. 5-13.
14. Puzakov A.V., Osaulko Ya. Yu. Investigation of the influence of operational factors on the thermal state of an automobile generator // Bulletin of the Moscow automobile and road state technical University (MADI), 2018, No 1 (52), pp. 16-23.
15. Kulandin A. S. Kogan A. G. Production of high-volume combined yarn using microwave currents. // Modeling in engineering and Economics: collection of materials international. scientific-practical Conf. / VGTU, Vitebsk, 2016, pp. 101-103.
16. Nezevak V. L. Modeling of power consumption for traction if you change the settings of the schedule of trains on electrified sections III-th and IV-th type of the profile path // *Sovremennyye tekhnologii. Sistemnyy analiz. Modelirovaniye [Modern Technologies. System Analysis. Modeling]*, 2019, No. 1 (61), pp. 156–166.
17. Stelmakov V. A., Nikitenko A.V., Davydov V. M., Gimadeev M. R. Providing precision forms for finishing holes by milling // Information technologies of the XXI century: collection of scientific papers, 2017, pp. 502-509.
18. Chertov Yu. E., Galatov K. S., Molev M. D. Regression analysis of the process of deformation of the pen rod by a grooved roll // Modern problems of science and education, 2015, No 2-3, p. 49.
19. Bazilevsky M. P. criteria of non-linearity of multi-factor quasilinear regression // Youth and science: actual problems of fundamental and applied research: collection of materials of the II all-Russian national scientific conference of students, postgraduates and young scientists, 2019, pp. 210-213.
20. Bazilevsky M. P. Criteria for non-linearity of quasilinear regression models // Modeling, optimization and information technologies, 2018, Vol. 6, No 4 (23), pp. 185-195.

Информация об авторах

Анна Витальевна Караулова – аспирант кафедры «Информационные системы и защита информации», Иркутский государственный университет путей сообщения, г. Иркутск, e-mail: anuta_kav@mail.ru

Михаил Павлович Базилевский – к. т. н., доцент, доцент кафедры «Математика», Иркутский государственный университет путей сообщения, г. Иркутск, e-mail: mik2178@yandex.ru

Authors

Anna Vitalievna Karaulova – postgraduate student, the Subdepartment of Information Systems and Information Security, Irkutsk State Transport University, Irkutsk, e-mail: anuta_kav@mail.ru

Mikhail Pavlovich Bazilevskiy – Ph. D. in Engineering Science, Associate Professor, the Subdepartment of Mathematics, Irkutsk State Transport University, Irkutsk, e-mail: mik2178@yandex.ru

Для цитирования

Караулова А.В., Базилевский М.П. Применение регрессионного анализа при решении реальных задач технического характера // «Молодая наука Сибири»: электрон. науч. журн. – 2020. – №3(9). Режим доступа: <http://mnv.ircgups.ru/toma/39-2020> . – Загл. с экрана. – Яз. рус., англ. (дата обращения: 20.10.2020)

For citations

Karaulova A.V., Bazilevskiy M.P. Application of regression analysis in solving real technical problems // *The electronic scientific journal "Young science of Siberia"*, 2020, no. 3(9), 2020. [Accessed 20/10/20]