

Организация конкурса регрессионных моделей выгрузки вагонов на железнодорожном транспорте

С.И. Носков, Д.В. Пашков, Т.Т. Улыбин, А.Ю. Улыбина

Иркутский государственный университет путей сообщения

Аннотация: В работе представлено описание процедуры проведения конкурса регрессионных моделей на основе статистических данных для Восточно-Сибирской железной дороги. При этом предполагается построение множества аддитивных альтернативных вариантов модели с последующим выбором лучшего варианта на основе привлечения ряда критериев адекватности. В качестве выходной переменной модели выделена выгрузка вагонов, а входными переменными являются: средний вес брутто грузового поезда, случаи отказов технических средств 1 – 2 категории эксплуатационного характера, рабочий парк грузовых вагонов. Реализация конкурса моделей позволила построить свыше двухсот альтернативных вариантов, из которых была выбрана лучшая альтернатива при помощи методов многокритериального выбора, основанных в данном случае на непрерывном критерии согласованности поведения.

Ключевые слова: железнодорожный транспорт, математическая модель, регрессионный анализ, метод наименьших квадратов, конкурс моделей, критерии адекватности, многокритериальный выбор.

Введение

Успешное функционирование экономики России, как и любой другой развитой страны, во многом связано с работой железнодорожного транспорта. Она выражается широким спектром показателей, что открывает значительные возможности для проведения научных исследований в этой области, в том числе, с помощью методов математического моделирования. Можно отметить работы, посвященные: погрузке грузов [1], определенным инженерным решениям [2, 3], описанию конкретных моделей [4], экологической безопасности [5]. В ряде публикаций рассмотрены некоторые другие важные аспекты функционирования транспортных систем: оценка уязвимости грузовых железнодорожных сетей [6], пути либерализации европейского сектора высокоскоростных железных дорог [7], перемещение груза в тренажере порталного крана [8], производственно-коммерческая деятельность автотранспортного предприятия [9].

При регрессионном моделировании деятельности железнодорожного транспорта выделяют, в том числе, следующие факторы [10 – 12]:

- эксплуатационные показатели производительности: объем грузопогрузки/выгрузки, грузооборот, пассажирооборот и др.;
- аварийные показатели: инциденты, отказы, аварии и т.д.;
- управляющие факторы: техническая скорость, средняя нагрузка на вагон, рабочий парк и др.;
- внешние факторы влияния: внутренний валовый продукт страны, показатели импорта/экспорта и др.

Спецификации математических моделей при этом могут быть выражены как обычными линейными зависимостями, так и более сложными – кусочно-линейными, а часто и существенно нелинейными связями, что сильно влияет как на состав переменных, так и на качественные показатели адекватности модельных конструкций. Поиск возможных более гибких взаимосвязей регрессоров и построение качественных регрессионных моделей является одной из приоритетных задач регрессионного анализа. Поиск таких связей можно осуществить путем организации конкурса регрессионных моделей [13]. Его суть состоит в построении множества альтернативных вариантов регрессий с разным составом переменных и их целенаправленным перебором до получения лучшего варианта (или множества вариантов) с привлечением системы качественных критериев оценки адекватности моделей [14].

Постановка задачи

Воспользуемся алгоритмической схемой организации конкурса моделей для построения регрессионных соотношений на основе статистических данных Восточно-Сибирской железной дороги (ВСЖД) за 2015 – 2022 гг., представленных в табл. 1. Используем следующий состав переменных модели: Y – выгрузка вагонов (вагонов в среднем в сутки), x_1 –

средний вес брутто грузового поезда (тонн), x_2 – случаи отказов технических средств 1 – 2 категории эксплуатационного характера (час), x_3 – рабочий парк грузовых вагонов (вагонов в среднем в сутки).

В качестве критериев адекватности модели будем использовать следующие: R – критерий множественной детерминации, F – критерий Фишера, E – средняя относительная ошибка аппроксимации зависимой переменной, $N(a)$ – непрерывный критерий согласованности поведения (НКСП) [15], p – вектор значений оценок вкладов факторов для каждой независимой переменной [16].

Таблица №1

Динамика статистических показателей ВСЖД

Год	y	x_1	x_2	x_3
2015	1856,2	3830	608,00	51112
2016	1903,62	3854	323,00	52901
2017	1901,14	3910	242,00	53521
2018	2020,99	3941	188,00	55614
2019	1983,04	3874	170,00	62723
2020	1962,48	3827	133,00	60953
2021	1974,17	3883	108,00	60085
2022	2140,79	4049	94,00	61349

По сформированной совокупности данных построим линейную модель при помощи обычного метода наименьших квадратов:

$$y = -1898.563 + 0.851x_1 + 0.004x_2 + 0.01x_3 \quad (1)$$

$$R = 0.931, F = 18.036, E = 0.887, N(a) = 64.381$$

$$p(x_1, x_2, x_3) = (85.075, 0.065, 14.860)$$

Модель (1) обладает высокими значениями критериев адекватности.

Отметим, что, в соответствии с оценками вкладов факторов, x_1 имеет

наибольшую значимость, превосходящую почти в 8 раз значение вклада для переменной x_3 , при этом значение x_2 пренебрежимо мало. Однако знаки оцененных параметров не соответствуют смыслу некоторых входящих в состав модели факторов. Обращает на себя внимание знак коэффициента при x_2 – очевидно, что отказы должны иметь отрицательное влияние на выходной показатель.

Попытаемся исправить ситуацию путем построения нелинейных аналогов модели (1) и выбора лучшего из них. Проведем конкурс регрессионных моделей, алгоритмическая схема которого описана в работе [17].

Результаты проведения конкурса моделей

При проведении конкурса были сформированы следующие ограничения на знаки параметров, значения критериев и возможные преобразования независимых переменных:

– независимые переменные должны влиять на зависимую переменную y следующим образом: x_1, x_3 – положительно, x_2 – отрицательно;

– $R \geq 0.8, F \geq 19, E \leq 15$;

– набор преобразований переменных: $G(x) = (x, x^2, 1/x, \ln x, \sqrt{x})$.

В результате проведения конкурса всего было построено 215 аддитивных по параметрам альтернативных вариантов модели, из них 36, удовлетворяющих указанным ограничениям. Окончательный выбор лучшего варианта производился по непрерывному критерию согласованности поведения.

Лучшие пять вариантов в порядке увеличения значения НКСП:

$$y = 878.664 + 0.0001x_1^2 + \frac{1785.165}{x_2} - \frac{27877884.434}{x_3} \quad (2)$$

$$R = 0.935, F = 19.188, E = 0.872, N(a) = 56.602$$

$$p(x_1^2, 1/x_2, 1/x_3) = (74.996, 0.841, 24.162)$$

$$y = -710.577 + 0.809x_1 + \frac{1928.885}{x_2} - \frac{27711306.481}{x_3} \quad (3)$$

$$R = 0.934, F = 18.853, E = 0.888, N(a) = 56.725$$

$$p(x_1, 1/x_2, 1/x_3) = (85.346, 0.534, 14.119)$$

$$y = -9572.822 + 101.199\sqrt{x_1} + \frac{2345.175}{x_2} + 475.696\ln x_3 \quad (4)$$

$$R = 0.933, F = 18.522, E = 0.89, N(a) = 57.761$$

$$p(\sqrt{x_1}, 1/x_2, \ln x_3) = (54.948, 0.213, 44.839)$$

$$y = 887.791 + 0.0001x_1^2 - \frac{30767380.075}{x_3} \quad (5)$$

$$R = 0.934, F = 35.609, E = 0.882, N(a) = 59.963$$

$$p(x_1^2, 1/x_3) = (74.398, 25.602)$$

$$y = 991.701 + 0.00009x_1^2 - 70.837\ln x_2 \quad (6)$$

$$R = 0.892, F = 20.677, E = 1.106, N(a) = 112.107$$

$$p(x_1^2, \ln x_2) = (76.241, 23.759)$$

Анализ итогов конкурса указывает на достигнутое соответствие знаков параметров смыслу регрессоров и высокие значения аппроксимационных характеристик. Несмотря на то, что уравнение (2) обладает большим значением НКСП, в качестве победителя может быть представлена любая из полученных альтернатив, в зависимости от приоритетов, установленных исследователем. Например, в уравнении (6) виден рост значения вклада x_2 , а модель (5) обладает большим значением критерия Фишера, чем вариант (2), что может послужить решающим фактором при выборе модели, особенно, если учесть незначительное влияние переменной x_2 в моделях (1) – (4).

Заключение

В работе описаны результаты проведения конкурса регрессионных моделей с заданными характеристиками, полученных на основе статистической информации Восточно-Сибирской железной дороги за 2015 – 2022 гг. Организация вычислений выполнена с помощью соответствующей вычислительной технологии, что позволило построить 215 альтернативных

вариантов модели. Произведен краткий анализ этих результатов, указывающий на возможность изменения характера влияния независимых факторов на выходную переменную и повышение значений критериев адекватности при использовании более гибких нелинейных связей между переменными.

Литература

1. Голик В. Ю. Сравнительный анализ погрузки грузов на железнодорожном транспорте России и США // Микроэкономика, 2010, № 6, С. 93-99.

2. Petrushov V. V., Zaruba E. R., Poganyuk L. I., Rutsкая E. O., Martynov M. S. Increase of efficiency of railway junction with application of intelligent model module of railway station // Metallurgical and Mining Industry, 2017, No. 4, pp. 36-41.

3. Bobrovskiy V., Kozachenko D., Dorosh A., Demchenko E., Bolvanovska T., Kolesnik A. Probabilistic Approach for the Determination of Cuts Permissible Braking Modes on the Gravity Humps // Transport Problems, 2016, Vol. 11, No. 1, pp. 147-155.

4. Прищепа М. В. Математическое обеспечение распределенной системы диагностирования устройств железнодорожной автоматики и телемеханики // Инженерный вестник Дона, 2007, № 2. URL: ivdon.ru/ru/magazine/archive/n2y2007/27.

5. Кузина Е. Л. Основы управления развитием эколого-экономической системы природопользования на железнодорожном транспорте // Инженерный вестник Дона, 2011, № 3. URL: ivdon.ru/ru/magazine/archive/n3y2011/501.

6. Bababeik M., Nasiri M. M., Khademi N., Chen A. Vulnerability evaluation of freight railway networks using a heuristic routing and scheduling optimization model // Transportation, 2019, Vol. 46, pp. 1143–1170.

7. Ruiz-Rúa, A., Palacín, R. Towards a liberalised European high speed railway sector: Analysis and modelling of competition using Game Theory // European Transport Research Review, 2013, Vol. 5, pp. 53–63.

8. Файзрахманов Р. А., Тютюных А. А., Полевщиков И. С. Математическое, алгоритмическое и программное обеспечение задачи расчета управляющих воздействий на основе нечеткой логики при моделировании перемещения груза в тренажере порталного крана // Инженерный вестник Дона, 2022, №1. URL: ivdon.ru/ru/magazine/archive/n1y2022/7384.

9. Кравцов Д. В. Моделирование динамики экономических показателей производственно-коммерческой деятельности автотранспортного предприятия с помощью маржинального анализа // Инженерный вестник Дона, 2010, № 4. URL: ivdon.ru/ru/magazine/archive/n4y2010/279.

10. Hong WT, Clifton G, Nelson JD. Railway accident causation analysis: Current approaches, challenges and potential solutions // Accident Analysis & Prevention, 2023, Vol. 186. URL: doi.org/10.1016/j.aap.2023.107049.

11. Серикова О. Ю., Мартышев А. В., Якимова Л. Д. Оценка эффективности системы мотивации персонала как фактора обеспечения качества работы на примере предприятия железнодорожного транспорта // International Journal of Advanced Studies, 2021, Т. 11, № 2, С. 81-99.

12. Дегтярева О. С. Учет инфляционного фактора в оценке финансовых результатов организаций железнодорожного транспорта // Сибирская финансовая школа, 2005, № 3(56), С. 37-39.

13. Носков С. И., Потороченко Н. А. Диалоговая система реализации конкурса регрессионных зависимостей // Управляющие системы и машины, 1992, № 3-4, С. 111-116.

14. Носков С. И. Построение свертки критериев адекватности регрессионных моделей // Модели, системы, сети в экономике, технике, природе и обществе, 2022, № 1(41), С. 73-81.

15. Носков, С. И. Применение непрерывного критерия согласованности поведения при построении регрессионных моделей // Известия Тульского государственного университета. Технические науки, 2021, № 6, С. 74-78.

16. Носков С. И., Врублевский И. П. Анализ регрессионной модели грузооборота железнодорожного транспорта // Вестник транспорта Поволжья, 2020, № 1(79), С. 86-90.

17. Пашков, Д. В, С. И. Носков. Реализация конкурса регрессионных моделей эффективности интеллектуальной деятельности // Электронный сетевой политематический журнал "Научные труды КубГТУ", 2022, № 6, С. 40-51.

References

1. Golik V. Yu. Mikroekonomika. 2010, № 6, pp. 93-99.
 2. Petrushov V. V., Zaruba E. R., Poganyuk L. I., Rutsкая E. O., Martynov M. S. Metallurgical and Mining Industry, 2017, No. 4, pp. 36-41.
 3. Bobrovskiy V., Kozachenko D., Dorosh A., Demchenko E., Bolvanovska T., Kolesnik A. Probabilistic Transport Problems. 2016. Vol. 11, No. 1. pp. 147-155.
 4. Prishchepa M. V. Inzhenernyj vestnik Dona, 2007, № 2. URL: ivdon.ru/ru/magazine/archive/n2y2007/27.
 5. Kuzina E. L. Inzhenernyj vestnik Dona, 2011, № 3. URL: ivdon.ru/ru/magazine/archive/n3y2011/501.
 6. Bababeik M., Nasiri M. M., Khademi N., Chen A. Transportation, 2019, Vol. 46, pp. 1143–1170.
 7. Ruiz-Rúa, A., Palacín, R. European Transport Research Review, 2013, Vol. 5, pp. 53–63.
-



8. Fayzrakhmanov R. A., Tyutyunykh A. A., Polevshchikov I. S. Inzhenernyj vestnik Dona, 2022, № 1. URL: ivdon.ru/ru/magazine/archive/n1y2022/7384.
9. Kravtsov D. V. Inzhenernyj vestnik Dona, 2010, № 4. URL: ivdon.ru/ru/magazine/archive/n4y2010/279.
10. Hong WT, Clifton G, Nelson JD. Accident Analysis & Prevention, 2023, Vol. 186. URL: doi.org/10.1016/j.aap.2023.107049.
11. Serikova O. Yu., Martyshev A. V., Yakimova L. D. International Journal of Advanced Studies, 2021, T. 11, № 2, pp. 81-99.
12. Degtyareva O. S. Sibirskaya finansovaya shkola, 2005, № 3(56), pp. 37-39.
13. Noskov S. I., Potorochenko N. A. Upravlyayushchie sistemy i mashiny, 1992, № 3-4, pp. 111-116.
14. Noskov S. I. Modeli, sistemy, seti v ekonomike, tekhnike, prirode i obshchestve, 2022, № 1(41), pp. 73-81.
15. Noskov S. I. Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki, 2021, № 6, pp. 74-78.
16. Noskov S. I., Vrublevskiy I. P. Vestnik transporta Povolzh'ya, 2020, № 1(79), pp. 86-90.
17. Pashkov, D. V., Noskov S. I. Elektronnyj setevoy politematicheskij zhurnal "Nauchnye trudy KubGTU", 2022, № 6, pp. 40-51.