

ОСОБЕННОСТИ ПОСТРОЕНИЯ СТОХАСТИЧЕСКОЙ МОДЕЛИ ОЦЕНКИ ПАРАМЕТРА НЕПРОИЗВОДИТЕЛЬНОГО ПОТЕРЯ

Дилмурод Баходирович Бутунов

PhD, доцент кафедры «Организация движения на транспорте», Ташкентский
государственный транспортный университет (Узбекистон),

dilmurodpgups@mail.ru

Шинполат Мансуралиевич Суюнбаев

Кандидат технических наук, и.о. профессор кафедры «Организация движения
на транспорте», Ташкентский государственный транспортный университет
(Узбекистон),

shinbolat_84@mail.ru

Муслима Джалаловна Ахмедова

Старший преподаватель кафедры «Организация движения на транспорте»,
Ташкентский государственный транспортный университет (Узбекистон),

muslimaakhmedova@mail.ru

АННОТАЦИЯ

В данной статье определена методика ввода параметров непроизводительных потерь времени при нормировании времени нахождения транзитных вагонов с переработкой на основе их планового нормирования. В результате появляется возможность совершенствовать методы нормирования времени нахождения вагона, а также рационально организовать и управлять работой по переработке вагонов на сортировочной станции. Разработан алгоритм для определения и установления значений параметра непроизводительной потери времени при переработке вагонов на сортировочной станции. Применение этого алгоритма позволяет устанавливать связь между значениями параметра непроизводительных потерь и величинами группы факторов на основе их анализа.

Ключевые слова: Сортировочная станция, непроизводительная потеря, время нахождения вагонов на станции, стохастическое моделирование, коэффициент корреляция.

FEATURES OF BUILDING A STOCHASTIC MODEL FOR ESTIMATING THE PARAMETER OF NONPRODUCTIVE LOSS

ABSTRACT

This article defines a methodology for entering parameters of unproductive time losses when standardizing the time spent by transit wagons with processing on the basis of their planned rate setting. As a result, it becomes possible to improve the methods of standardizing the time spent by a wagon, as well as to rationally organize and manage the work on processing wagons at the sorting station. An algorithm has been developed for determining and establishing the values of the parameter of unproductive loss of time when processing wagons at a sorting station. The use of this algorithm makes it possible to establish a relationship between the values of the parameter of non-productive losses and the values of a group of factors on the basis of their analysis.

Keywords: Sorting station, non-productive loss, time spent by wagons at the station, stochastic modeling, correlation coefficient.

ВВЕДЕНИЕ

Непроизводительные потери (далее в тексте – потерь) в работе любой сортировочной станции возникает из-за недостатков организации и управления вагонопотоков, нарушений планово-учетной дисциплины, слабые взаимодействия со смежными подразделениями, участвующими в одном технологическом процессе [1-3]. Поэтому определение причины возникновения потерь при переработке вагонопотоков, а также своевременно разработать меры по его сокращению в работе станции становится важной проблемой.

Как отмечено в научные работы [2, 3] причины потерь на сортировочные станции разделяется на две группы: явные и скрытые.

На практике явные потери возникновения легко физически увидеть при переработке вагонопотоков на сортировочной станции или отследить в отчетных данных. А, скрытые потери присутствуют в значительно большем количестве, но являются скрытыми, т.е. вред от явных потерь величина скрытых потерь менее очевидна и, если не проводить специальных мероприятий она становится неизбежной частью технологического процесса.

В этой связи определение и сокращение, а также моделирование и оценки такие потери при переработке вагонопотоков является актуальной проблемой работы сортировочной станции.

ЛИТЕРАТУРНЫЙ АНАЛИЗ И МЕТАДОЛОГИЯ

Одним из эффективных методов моделирование непроизводительных потерь в работе сортировочной станции является стохастическое моделирование.

Статистические данные о выполнениях время нахождения транзитных вагонов с переработкой на сортировочной станции ($t_{\text{пер}}$) [1, 3] и величинах потерь в их работе [1-4] позволяют определить величину потерь ($t_{\text{нп}}$) применив математический аппарат стохастического моделирования [5].

Построение соотношений между значениями $t_{\text{нп}}$ и величинами потери в работе станции $t_{b_1}, t_{b_2}, \dots, t_{b_{15}}$ и $t_{q_1}, t_{q_2}, \dots, t_{q_{34}}$ [2, 6] с помощью статистического обобщения закономерностей изменения их характеристик лежит в основе стохастического моделирования параметра $t_{\text{нп}}$.

На основе работы [7-9] определены этапы построения модели оценки параметра $t_{\text{нп}}$.

1. Массив статистических данных о выполнении $t_{\text{пер}}$ на станции и величинах причины потерь.

В рамках задачи разработки модели взаимосвязи величины параметра $t_{\text{нп}}$ и причин потерь $t_{b_1}, t_{b_2}, \dots, t_{b_{15}}$ и $t_{q_1}, t_{q_2}, \dots, t_{q_{34}}$ влияющих на значение $t_{\text{пер}}$, моделируемая совокупность статистических данных должна содержать ряд значений параметра $t_{\text{нп}}$ и причин потерь $t_{b_1}, t_{b_2}, \dots, t_{b_{15}}$ и $t_{q_1}, t_{q_2}, \dots, t_{q_{34}}$. При этом параметр $t_{\text{нп}}$ является результирующим признаком, так как под его действием изменяется причина потерь.

Составления совокупности данных значений $t_{\text{нп}}$ и причинных групп P_z обеспечивается наличием данных отчётов ДО-24 ВЦ (отчет о работе сортировочных станций) и ежесуточных справок об ожиданиях грузовых вагонов при переработке вагонов на станции.

На основе [2] формируется ряд посуточных значений параметра $t_{\text{нп}}$

Объем выборок зависит от числа влияющих причины, учитываемых в модели. Минимальный объем выборки n_{\min} определяется соотношением [7]:

$$n_{\min} = (6 \dots 8)m, \quad (1)$$

где m – число включаемых причины.

Минимальный объем выборки при 6 рассмотренных группах причины составит

$$n_{\min} = (6 \div 8)6 = 36 \div 48.$$

Построение модели оценки параметра $t_{\text{нн}}$ производится на основе статистических данных, взятых минимум за 2 предшествующих месяца.

Для определения аналитического выражения зависимости между $t_{\text{нн}}$ и P_z применяется метод регрессионного анализа. При решении задачи необходимо учесть одновременное влияние 6 причинных признаков, то используется множественная регрессия.

Используемый регрессионный анализ дает возможность получить уравнение регрессии (формализованное выражение), т.е. связи значений $t_{\text{нн}}$ с величинами причинных групп P_z .

2. Анализ моделируемой совокупности

Для корректного использования стохастического метода к моделированию связей – однородность совокупности [7, 9, 10].

Для получения реальной модели требуется совпадение численных характеристик связей в резерве всех наблюдений. Варьирование предложенных величин может происходить в пределах однозначной определенности качественной стороны явлений [11].

Неоднородность совокупности возникает из-за значительной вариации значений признака, так называемых «аномальных» наблюдений [12, 13]. Для их выявления используется правило 3-х сигм: величины признака признаются аномальными, если его отклонение от выборочной более чем в 3 раза превышает среднеквадратическое отклонение [12]:

$$|x_i - \bar{x}| > 3\sigma, \quad (2)$$

где x_i – значение признака;

σ – среднеквадратическое отклонение значений признака;

\bar{x} – среднее значение признака.

Численная оценка однородности рядов данных причинных групп P_z может проводиться по коэффициенту вариации Var [7, 10, 9, 14].

$$Var = \frac{\sigma}{\bar{x}} \quad (3)$$

Выборка считается однородной когда [15] $Var \leq 0,33$.

3. Построение модели оценки параметра $t_{\text{нн}}$.

Построение модели оценки [15] начинается с выбора класса модели для описания влияния причины потерь на величину параметра $t_{\text{нн}}$.

Этот этап включает в себя следующие задачи:

3.1. *Уточнение перечня группы причин, включаемых в модель.* Основным этапом построения множественной регрессии – отбор и последующее включение в него причинных признаков. Данной регрессия может дать хороший результат при моделировании только в том случае, если влиянием других причины. В то же время неоправданное увеличение количества причины приводит к уменьшению достоверности результатов [8, 11].

В работе [6] было выделено шесть групп причин потерь, определяющих величина параметра $t_{\text{нн}}$ и подлежащих в модель. В регрессионном анализе каждого причин проверяется с помощью формальных статистических методов [15] и устанавливается тесной линейной корреляционной зависимости между ними. Эта связи определяется с помощью коэффициента корреляции [9]. Коэффициент корреляции между причинами P_z и P_k определяется по формулу:

$$r_{P_z P_k} = \frac{1}{n-1} \frac{\sum_{t=1}^n (P_{zt} - \bar{P}_z)(P_{kt} - \bar{P}_k)}{\sigma_{P_z} \sigma_{P_k}}, \quad (4)$$

где $\sigma_{P_z}, \sigma_{P_k}$ – средние квадратические отклонения переменных причины, соответственно P_z и P_k .

Причины не должны быть интеркоррелированы, которые включаемые в модель, т.е. между причинами имеется высокая корреляция, тогда нельзя определить их изолированное влияние и параметры уравнения регрессии оказываются неинтерпретируемыми. Тогда две переменные явно коллинеарны, если $r_{P_z P_k} > 0,7$ [7, 9, 10].

Когда более 2 группа причина связаны между собой зависимостью, наряду с парной коллинеарностью может иметь место мультиколлинеарность.

Мультиколлинеарности выделяют следующие [11]:

- повышение дисперсий оценок параметров;
- снижение значений t – статистик для параметров;
- принятие неустойчивых оценок параметров модели и их дисперсий;
- получения неверного знака коэффициента регрессии.

Для оценки мультиколлинеарности причины может применяться значения рассчитаться $Det|R|$ матрицы парных коэффициентов корреляции $r_{P_z P_k}$

$$R = \begin{bmatrix} r_{P_1 P_1} & r_{P_1 P_2} & \dots & r_{P_1 P_6} \\ r_{P_2 P_1} & r_{P_2 P_2} & \dots & r_{P_2 P_6} \\ \dots & \dots & \dots & \dots \\ r_{P_6 P_1} & r_{P_6 P_2} & \dots & r_{P_6 P_6} \end{bmatrix} = \begin{bmatrix} 1 & r_{P_1 P_2} & \dots & r_{P_1 P_6} \\ r_{P_2 P_1} & 1 & \dots & r_{P_2 P_6} \\ \dots & \dots & \dots & \dots \\ r_{P_6 P_1} & r_{P_6 P_2} & \dots & 1 \end{bmatrix} \quad (5)$$

Когда межпричинной корреляции $Det|R|$ ближе к нулю определитель матрицы считается сильнее мультиколлинеарность между причинами и результатами множественной регрессии.

При оценке статистической значимости мультиколлинеарности причины может быть применен тот факт, что значения $\left[n - 1 - \frac{1}{6}(2m + 5) \lg Det|R| \right]$ считает доверительное распределение χ^2 с $df = \frac{1}{2} p(p - 1)$ степенями свободы.

Гипотеза H_0 о независимости изменчивых, $Det|R|=1$. Если фактическое величине χ^2 превосходит критическое $\chi^2_{факт} > \chi^2_{табл(df, \alpha)}$, то гипотеза H_0 отклоняется и в результате ультиколлинеарность является доказанной [10].

Критериев включения причины в модель считается уровень их влияния на признак, рассчитываемая парной корреляции $r_{t_{не}P_z}$ [9]:

$$r_{t_{не}P_z} = \frac{1}{n-1} \frac{\sum_{t=1}^n (t_{неt} - \bar{t}_{не})(P_{zt} - \bar{P}_z)}{\sigma_{t_{не}} \sigma_{P_z}}, \quad (6)$$

Величина коэффициента изменяются в интервале $[-1; 1]$, чем ближе $|r_{t_{не}P_z}|$ к 1, тем связь теснее.

3.2. *Выбор уравнения регрессии.* Уравнения $t_{нп}$, выражающая зависимость среднего $t_{нп}$ от значений групп P_z имеет вид

$$t_{нп} = f\{P_1, P_2, \dots, P_6\} + \varepsilon,$$

В ходе работе проведено исследование по выбору вида уравнений множественной регрессии по методике. В результате исследования определена линейная регрессия, и сформулированная их модель

$$t_{нп} = a + b_1P_1 + b_2P_2 + \dots + b_6P_6 + \varepsilon \quad (7)$$

где b_g – коэффициенты «чистой» регрессии, $g = 1 \div 6$. a – независимый коэффициент.

3.3. *Оценка коэффициентов модели параметра $t_{нп}$.*

Для этого используем метод наименьших квадратов, которому следует выбрать следующие значения параметров a и b_g , вследствие чего сумма квадратов отклонений фактических значений результативного признака $t_{нп}$ от теоретических значений $t_{нп} = f\{P_{1i}, P_{2i}, \dots, P_{6i}\}$ незначительна, т.е.

$$S = \sum_{i=1}^n (\bar{t}_{\text{нн}i} - t_{\text{нн}i})^2 = \min,$$

С учетом (7) значения S считается функцией неизвестных параметров a и b_g

$$S = \sum_{i=1}^n (t_{\text{нн}i} - a - b_1 P_1 - b_2 P_2 - \dots - b_6 P_6)^2$$

Наилучшие величины параметров a и b_g удовлетворяют условиям

$$\frac{\partial S}{\partial a} = 0, \quad \frac{\partial S}{\partial b_1} = 0, \quad \frac{\partial S}{\partial b_2} = 0, \quad \dots \quad \frac{\partial S}{\partial b_6} = 0$$

После этого для определения a и b_g получим следующие уравнения

$$\begin{aligned} \frac{\partial S}{\partial a} &= -2 \sum_{i=1}^n (t_{\text{нн}i} - a - b_1 P_1 - b_2 P_2 - \dots - b_6 P_6), \\ \frac{\partial S}{\partial b_1} &= -2 b_1 \sum_{i=1}^n (t_{\text{нн}i} - a - b_1 P_1 - b_2 P_2 - \dots - b_6 P_6), \\ &\dots \\ \frac{\partial S}{\partial b_6} &= -2 b_6 \sum_{i=1}^n (t_{\text{нн}i} - a - b_1 P_1 - b_2 P_2 - \dots - b_6 P_6), \end{aligned}$$

откуда получим систему искомых уравнений

$$\begin{aligned} \sum t_{\text{нн}} &= n \cdot a + b_1 \sum P_1 + b_2 \sum P_2 + \dots + b_6 \sum P_6; \\ \sum t_{\text{нн}} P_1 &= a \sum P_1 + b_1 \sum P_1^2 + b_2 \sum P_2 P_1 + \dots + b_6 \sum P_6 P_1; \\ &\dots \\ \sum t_{\text{нн}} P_6 &= a \sum P_6 + b_1 \sum P_1 P_6 + b_2 \sum P_2 P_6 + \dots + b_6 \sum P_6^2 \end{aligned} \quad (8)$$

При решении полученные системы (8) удобно описать с помощью матричных обозначений

$$B = \begin{bmatrix} a \\ b_1 \\ \dots \\ b_6 \end{bmatrix}, \quad t_{\text{ин}} = \begin{bmatrix} t_{\text{ин}1} \\ t_{\text{ин}2} \\ \dots \\ t_{\text{ин}n} \end{bmatrix}, \quad P = \begin{bmatrix} 1 & P_{11} & \dots & P_{61} \\ 1 & P_{12} & \dots & P_{62} \\ 1 & \dots & \dots & \dots \\ 1 & P_{1n} & \dots & P_{6n} \end{bmatrix},$$

где B – матрица-столбец $(6+1 \times 1)$ из коэффициентов a и b_g ; $t_{\text{ин}}$ – матрица-столбец $(n \times 1)$ исходных значений параметра $t_{\text{ин}}$; P – матрица $(6+1 \times n)$ исходных значений групп причин потерь P_k , где первый столбец из единиц можно учитывать как значения «фиктивной» переменной, соответствующей коэффициенту a .

Тогда, система (8) примет вид

$$(P'P)B = P't_{\text{ин}}, \quad (9)$$

где P' – транспонированная матрица P . Матрица $P'P$ является неособенной квадратной размерности $(6+1 \times 6+1)$ при условии, что столбцы матрицы P линейно независимы.

Решение системы (9) определяется соотношением

$$B = (P'P)^{-1} P't_{\text{ин}}, \quad (10)$$

После получения коэффициентов уравнения множественной регрессии a и b_g оценивается их точность. Определяемые оценки коэффициентов регрессии зависят от используемой выборки значений переменных P_k и $t_{\text{ин}}$ и является случайными величинами. Для характеристики точности полученных оценок можно использовать стандартные ошибки коэффициентов регрессии.

4. *Проверка адекватности.* Применения модели оценки параметра $t_{\text{ин}}$ глубокое значение имеет ее адекватность. Регрессионный анализ взаимосвязи групп причины потерь и P_z параметра $t_{\text{ин}}$ проводится для ограниченной по объёму совокупности, коэффициенты уравнения регрессии могут быть изменены действием случайных причин. Проверка адекватности модели оценки параметра $t_{\text{ин}}$ производится на основе F – критерия Фишера.

Оценка качества уравнения регрессии с использованием коэффициента детерминации R^2 . Он представляет собой отношение объясненной части $D(\hat{t}_{\text{нн}})$ дисперсии переменной $t_{\text{нн}}$ ко всей дисперсии $D(t_{\text{нн}})$

$$R^2 = \frac{D(\hat{t}_{\text{нн}})}{D(t_{\text{нн}})} \text{ или } R^2 = \frac{\sum_{i=1}^n (\hat{t}_{\text{нн}i} - \bar{t}_{\text{нн}})^2}{\sum_{i=1}^n (t_{\text{нн}i} - \bar{t}_{\text{нн}})^2}, \quad (11)$$

где $D(t_{\text{нн}}) = \frac{1}{n} \sum (t_{\text{нн}i} - \bar{t}_{\text{нн}})^2$, $D(\hat{t}_{\text{нн}}) = \frac{1}{n} \sum (\hat{t}_{\text{нн}i} - \bar{t}_{\text{нн}})^2$, $D(e) = D_{\text{ост}} = \frac{1}{n} \sum (\hat{t}_{\text{нн}i} - t_{\text{нн}i})^2$,

Коэффициент R^2 принимает значения $0 \leq R^2 \leq 1$. Когда выше величина R^2 , тогда считается, что данная модель лучше и согласуется с данными наблюдений. Исходя из этого, делается вывод о правильности уравнения регрессии.

РЕЗУЛЬТАТЫ

На основе исследования разработана модели оценки параметра потерь $t_{\text{нн}}$. Модель включает в себя следующие:

1. Сбор исходных данных.

Для этого выбирается базового периода анализа и на основе суточного плана-графика работы сортировочной станции составляется таблица значений условно-постоянных факторов. На основе обработки отчетов ДО-24 ВЦ и ежесуточных справок об ожиданиях грузовых вагонов при переработке вагонов на станции формируется таблица посуточных значений среднего времени нахождения транзитного вагона с переработкой и значений причин потерь [2]. Используя полученную таблицу, рассчитываются значения параметра $t_{\text{нн}}$ по [2] и групп причин потерь P_z [2].

2. Анализ моделируемой совокупности.

Для этого на основе правила трех сигм (2) производится исключение «аномальных» значений, нахождение коэффициента вариации значений параметра $t_{\text{нн}}$ и значений $\sum_{i=1}^l P$ (3), а также определяется закон распределения

значений $t_{\text{нн}}$ и $\sum_{i=1}^l P$.

Когда значений меньше 0,33 можно считать нормальной. Это означает что значения потерь $t_{\text{нн}}$ и $\sum_{i=1}^I P$ распределены по нормальному закону.

3. Оценка связи между группами причины потерь и параметрами потерь

Для этого определяются коэффициенты межфакторной корреляции по формуле (4) и составляется корреляционная матрица (5) и проверяются на условие $r_{P_z P_k} < 0,7$. Если условие выполняется, то причин считаются неколлинеарными. После оценивается мультиколлинеарность факторов путем вычисления определения корреляционной матрицы $Det|R|$.

Затем определяются коэффициенты парной корреляции $r_{t_{\text{нн}} \sum_{i=1}^I P}$ между $t_{\text{нн}}$ и $\sum_{i=1}^I P$ по формуле (6). В рамках определенной задачи коэффициенты $r_{t_{\text{нн}} P_z}$ необходимо принимать положительными величинами; коэффициент $r_{t_{\text{нн}} \sum_{i=1}^I P}$ значение близкое к единице. При неположительном величинах $r_{t_{\text{нн}} P_z}$ или ближе к $r_{t_{\text{нн}} P_z} < 0,1$ выполняется повторная (экспертная) оценка роли данной группы причин. Если в результате группа признается неважной, тогда исключается из модели.

4. Оценка коэффициентов уравнения.

Для этого определяются величины коэффициентов уравнения методом наименьших квадратов (10) и рассчитывается коэффициент R^2 (11), F – критерий Фишера.

После этого оценивается достоверность коэффициентов уравнения a и b_g с помощью доверительных интервалов. Если он попадает в ноль, тогда коэффициент считается незначимым и исключается из модели. Исходя из этого, построится новое уравнение множественной регрессии и снова выполняется оценка значимости всех оставшихся коэффициентов регрессии. Когда все регрессионные коэффициенты значимы, тогда процесс исключения факторов считается законченным.

Путем постановки полученных значений коэффициентов a и b_g в (7) выводится уравнение регрессии.

5. Проверка адекватности параметра потерь $t_{\text{нн}}$.

Проверка адекватности модели оценки параметра потерь $t_{\text{нп}}$ выполняется на основе коэффициента R^2 . Чем больше величины R^2 , тем лучше данная модель.

Если модель признается адекватной, то она применяется для расчета величина потерь $t_{\text{нп}}$ на прогнозируемый период.

Разработанный алгоритм определения параметра $t_{\text{нп}}$ представлен на рисунке 1.



Рис. 1. Блок-схема алгоритма оценки параметра $t_{\text{нп}}$ в работе сортировочной станции: А – процедура исключения из моделируемой совокупности группы причины потерь P_z , коэффициент при которой незначим; В – процедура преобразования модели

ЗАКЛЮЧЕНИЕ

В статье подробно рассмотрены этапы построения стохастической модели применительно к задаче определения значения параметра $t_{\text{нпв}}$ в зависимости от размеров причины потерь, возникающих при переработке вагонов на сортировочной станции.

Для установления аналитического выражения стохастической зависимости между параметрами $t_{\text{нпв}}$ и факторными группами P_z может использоваться метод множественного регрессионного анализа.

Применение разработанной модели для расчета нормы времени нахождения транзитного с переработкой вагона на сортировочной станции позволяет:

- рационально и обоснованно установить норму времени нахождения вагона;
- анализировать возможное его отклонение от планового значения;
- оценивать значимость влияния на величину $t_{\text{пер}}$ отдельных групп факторов P_z и устанавливать необходимость проведения мероприятий по минимизации этого влияния с целью сокращения причины потерь при переработке вагонопотоков.

REFERENCES

1. Dilmurod B. Butunov. *Estimation of inefficient losses in railroad yard operation* / D.B. Butunov, A.G. Kotenko // Emperor Alexander I St. Petersburg State Transport University, 2018, Volume 15, Issue 4, Pages 498-512. (<https://cyberleninka.ru/article/n/otsenka-neproizvoditelnyh-poter-v-rabote-sortirovochnoy-stantsii>)
2. Suyunbayev, Sh.M. and Butunov, D.B. (2019) “Development of classification of the reasons of losses in the work sorting stations” *Journal of Tashkent Institute of Railway Engineers*: Vol. 15: Iss. 2, Article 23. Available at: (<https://uzjournals.edu.uz/tashiit/vol15/iss2/23>)

3. Dilmurod Butunov Baxodirovich, Sokijon Khudayberganov Kobiljonovich, Shinpolat Suyunbaev Mansuralievich. Modeling of unproductive losses in the operation of a sorting station / European Journal of Molecular & Clinical Medicine, 2020, Volume 7, Issue 2, Pages 277-290. (https://ejmcm.com/article_2070.html).
4. Butunov, D.B. (2019) “*Development of economic and mathematical model of calculation of expenses at processing of cars at sorting station,*” Journal of Tashkent Institute of Railway Engineers: Vol. 15: Iss. 3, Article 19. Available at: <https://uzjournals.edu.uz/tashiit/vol15/iss3/19>
5. Орлов В.Н. Калькуляция и анализ себестоимости железнодорожных перевозок / В.Н. Орлов, А.С. Чудов. – М.: Транспорт, 1967. – 288 с.
6. Butunov, D.B. (2020) “*Substantiation of the input of the parameters of the unprofitable loss of time when norming the time of the duration of the wagons on the sorting station*” Journal of Tashkent Institute of Railway Engineers: Vol. 16: Iss. 3, Article 16. (<https://uzjournals.edu.uz/cgi/viewcontent.cgi?article=1191&context=tashiit>)
7. Вентцель Е.С. Теория вероятностей: учебник / Е.С. Вентцель. – 11-е изд., стер. – М.: КНОРУС, 2010. – 664 с. (https://www.centrmag.ru/catalog/product/teoriya_veroyatnostey_uchebnik_11_e_izd_/)
8. Иванов О.Б. Формирование единой риск-ориентированной системы внутреннего аудита и контроля в холдинге «РЖД» // Железнодорожный транспорт. 2009. №1 – С. 19-26. (<https://elibrary.ru/item.asp?id=17738419>)
9. Шапкин И.Н. Технология и управление перевозками на железнодорожном транспорте: (опыт, теория и практика переходного периода). М.: Желдориздат, 2003. – 527 с.
10. Айвазян С.А. Прикладная статистика. Основы эконометрики: учеб. для вузов. В 2 т. Т. 1: Теория вероятностей и прикладная статистика 2-е изд., испр. / С.А. Айвазян, В.С. Мхитарян. – М.: ЮНИТИ, 2001. – 656 с. (<https://ecsocman.hse.ru/data/2010/11/09/1214795494/Книга.pdf>)
11. Иберла К. Факторный анализ / Пер. с нем. В.М. Ивановой; Предисл. А.М. Дуброва. М.: Статистика, 1980. – 398 с.
12. Ветухов Е.А. Определение уровня загрузки станций методом моделирования их работы на ЭЦВМ / Е.А. Ветухов, Е.А. Сотников // Железнодорожный транспорт. – 1969. №7. – С. 34-37.

13. Лагутин М.Б. Наглядная математическая статистика: Учебное пособие. – 2-е изд., испр. – М.: БИНОМ. Лаборатория знаний, 2009. – 472 с. (<http://www.1variant.ru/content/uchebniki/matematika/650.pdf>)
14. Котенко А.Г. Методология риск-ориентированного планирования качественных показателей эксплуатационной работы железных дорог: автореф. дис... док. техн. наук: 05.22.08 / Котенко А.Г. – СПб., 2014. – 34 с. (<https://elibrary.ru/item.asp?id=30407302>)
15. Гоголева А.В. Систематизация, исследование причин невыполнения графика движения поездов и оценка их влияния на участковую скорость / А.В. Гоголева, Г.М. Грошев, В.И. Ковалёв, А.Г. Котенко // Известия ПГУПС. – 2012. №2 (31). – С. 5-11. (<https://elibrary.ru/item.asp?id=20213860>)