

А. Н. Митрофанов, Ю. И. Жарков, С. А. Окладов

МЕТОДОЛОГИЯ ПОСТРОЕНИЯ МОДЕЛЕЙ ЦИФРОВЫХ ДВОЙНИКОВ ПРОЦЕССОВ ЭНЕРГООБЕСПЕЧЕНИЯ ТЯГИ ПОЕЗДОВ

Аннотация. Представлена методология построения моделей цифровых двойников процессов энергообеспечения тяги поездов, позволяющая произвести корректный подбор факторов и идентификацию параметров многофакторной модели для территориальных филиалов ОАО «РЖД». В качестве процессов энергообеспечения рассматривались расход электроэнергии на тягу, удельный расход электроэнергии, небаланс тягового энергообеспечения. В качестве независимых переменных модели исследовались процессы грузовых перевозок и процессов по их обеспечению: погрузка груза, грузооборот, скорость движения поездов, показатели использования электровагонов, вагонов и др. Согласно результатам процедуры валидации установлена возможность производить моделирование динамических реализаций процессов энергообеспечения тяги поездов с погрешностью не более 3,1–3,7 % от фактических значений.

Ключевые слова: цифровые двойники, экспресс-модели процессов, энергообеспечение, тяга поездов, корреляционная связь, кластерный анализ, верификация и валидация, многофакторная модель цифрового двойника.

Для цитирования: Митрофанов, А. Н. Методология построения моделей цифровых двойников процессов энергообеспечения тяги поездов / А. Н. Митрофанов, Ю. И. Жарков, С. А. Окладов // Вестник Ростовского государственного университета путей сообщения. – 2024. – № 2. – С. 82–99. – DOI 10.46973/0201-727X_2024_2_82.

Введение

Повышение эффективности работы инфраструктуры железнодорожных перевозок предусмотрено долгосрочной программой развития ОАО «РЖД» до 2025 года и на перспективу до 2030 года [1], а информационное сопровождение деятельности перевозок грузов и обеспечивающей её инфраструктуры предусмотрено программой «Цифровая железная дорога». Так, в разделе программы «Инновационное и технологическое развитие» [1] описаны основные направления инновационного развития, предусматривающие разработку и внедрение перспективных технических средств и технологий железнодорожной инфраструктуры, в том числе для хозяйств электрификации и электроснабжения дирекции Трансэнерго компании «Российские железные дороги». Методологической основой для разработки обозначенных инновационных технологий должны являться сформированные совокупность и последовательность методов, используемых исследователем при решении проблемы. К характерным методологически важным моментам при этом относятся постановка проблемы, предмета исследования, построение совокупности научно обоснованных методов, направленных на достижение цели исследования, а также проверка полученного результата с точки зрения его истинности, то есть оценки соответствия объекту исследования. При разработке инновационной технологии оценки состояния процессов энергообеспечения тяги поездов в качестве эффективного инструментария целесообразно использовать понятийный аппарат, используемый в ГОСТ Р 57700.37–2021 «Цифровые двойники изделий» и ГОСТ Р 57188–2016 «Численное моделирование физических процессов» [2, 3]. Согласно источнику [2] цифровой двойник изделия «представляет собой систему, состоящую из цифровой модели изделия и двусторонних информационных связей с изделием и с его составными частями». В свою очередь, «цифровая модель цифрового двойника представляет собой систему математических и компьютерных моделей, описывающую структуру, функциональность и поведение вновь разрабатываемого или эксплуатируемого изделия, для которой на основании результатов цифровых испытаний выполнена оценка соответствия предъявляемым к изделию требованиям».

Опираясь на положения, определенные в источнике [3], методология оценки состояния процессов энергообеспечения тяги поездов (далее – методология) должна базироваться на использовании в исследовании ряда методов, подходов и способов, последовательно выполняющих следующие процедуры:

– формулирование законов, связывающих основные объекты модели при описании объекта исследования;

- исследование математическими методами сформулированных законов взаимосвязей объектов модели задач, к которым приводит математическая модель;
- выбор и адаптация инструментария для реализации модели цифрового двойника исследуемого объекта;
- верификация модели цифрового двойника исследуемого объекта посредством подтверждения корректности решения математических выражений, описывающих исследуемые модели;
- валидация модели цифрового двойника посредством подтверждения адекватности математической модели исследуемому объекту в целом;
- анализ и модернизация модели цифрового двойника при накоплении данных об изучаемом объекте.

Таким образом, в качестве *предмета исследования* методологии может выступать процесс построения цифровых двойников процессов, присущих системе тягового электроснабжения (СТЭ) хозяйства электрификации и электроснабжения на филиалах ОАО «РЖД», включающий выполнение научно обоснованных методов, подходов и способов исследования.

Наличие в структурных подразделениях дирекции Трансэнерго ОАО «РЖД», осуществляющей управление хозяйством электрификации и электроснабжения, инструментария и программного обеспечения для реализации модели цифровых двойников СТЭ, позволит оценить режимы и повысить эффективность работы СТЭ на территориальных филиалах дирекции. При этом можно будет оценить прогнозные режимы работы СТЭ в условиях вариации показателей перевозочного процесса (например, массы, скорости поездов, межпоездного интервала и др.); хозяйств, обеспечивающих перевозки (например, рабочий парк электровозов, вагонов и др.); нагрузочной способности СТЭ (например, за счет вариации типов «усиливающих» элементов СТЭ) и т. д. Конечной целевой функцией использования моделей цифровых двойников процессов энергообеспечения тяги поездов является оценка ресурса СТЭ для повышения пропускной и провозной способности на участках железных дорог по условиям электроснабжения. Однако базовые положения методологии для построения моделей цифровых двойников процессов энергообеспечения тяги поездов к настоящему времени не нашли достаточного отражения в научной отраслевой литературе, в связи с чем её разработка является актуальной.

Цель настоящей статьи – обоснование и апробация методов, подходов, способов в составе методологии построения математических моделей цифровых двойников процессов энергообеспечения тяги поездов на территориальных филиалах ОАО «РЖД».

Методы и средства

Подход к автоматизации и управлению устройствами СТЭ на полигонах энергообеспечения тяги поездов по результатам моделирования во многом основан на научных исследованиях из области математической статистики и теории автоматического управления, например [4–7]. Исследования в области автоматизации и управления устройствами СТЭ проводились в ряде транспортных вузов России, в том числе и на кафедрах «Электроснабжение железнодорожного транспорта» в СамГУПСе и «Автоматизированные системы электроснабжения» в РГУПСе [8–12]. На основании ранее проведенных исследований авторов и приведенных выше обоснований в данной работе предложена методология построения моделей цифровых двойников процессов энергообеспечения тяги поездов, которая основана на использовании ряда методов, подходов и способов, включенных в 10 основных процедур:

1 Построение обобщенной структуры взаимодействия процесса энергообеспечения тяги поездов, реализуемой филиалом дирекции Трансэнерго на полигоне исследуемого филиала ОАО «РЖД», встроенного в структуру матричных процессов грузовых перевозок и обеспечивающих их процессов смежных хозяйств и дирекций. В качестве методов построения структуры модели могут использоваться метод и инструментарий для моделирования бизнес-процессов типа ARIS (Architecture of Integrated Information Systems) [13] или отечественная разработка семейства программных средств КОМПАС-3D [14].

2 Теоретическое обоснование подбора потенциально возможных для использования в цифровой модели ключевых показателей эффективности (деятельности) (КПЭ) процессов энергообеспечения тяги поездов, грузовых перевозок и обеспечивающих их процессов смежных хозяйств и дирекций. Используются методы теории «управления по целям» [15].

3 Формализация закономерностей, связывающих в структуре модели КПЭ процесса энергообеспечения тяги поездов с КПЭ процессов грузовых перевозок, а также с КПЭ процессов, их обеспечивающих. В качестве метода используется анализ последовательности технологических операций исследуемого объекта [16, 17].

4 Мониторинг и формирование цифровых баз данных КПЭ процессов энергообеспечения, грузовых перевозок и обеспечивающих их процессов смежных хозяйств и дирекций на основании технологических операций исследуемого объекта [18] при использовании метода агрегирования данных [19] и инструментария Big Data [20].

5 Оценка статистических показателей КПЭ процессов энергообеспечения, грузовых перевозок и обеспечивающих их процессов со стороны смежных хозяйств и дирекций [18, 19].

6 Идентификация параметров взаимосвязи КПЭ, исследуемых для включения в математические модели цифровых двойников процессов энергообеспечения тяги поездов. Выявление «приоритетных» КПЭ, значительно влияющих на энергообеспечение тяги поездов методами корреляционного [4, 5] и кластерного анализов [21].

7 Построение комплекса однофакторных экспресс-моделей цифровых двойников СТЭ и их верификация по критерию физической и аналитической обоснованности. В качестве методов используются метод планирования эксперимента [22] и разведочный анализ данных (exploratory data analysis, EDA) [23]. Для верификации моделей применимы методы подхода формальной верификации: «проверка моделей» (model checking) и «логический вывод» (logical inference) [24].

8 Построение динамических реализаций КПЭ процессов энергообеспечения тяги поездов и «приоритетных» КПЭ процессов грузовых перевозок и обеспечивающих их процессов смежных хозяйств и дирекций, а также оценка их адекватности по критерию стационарности. В качестве метода используется «теория случайных процессов» [4].

9 Построение многофакторных моделей цифровых двойников КПЭ процессов энергообеспечения, верификация, устранение мультиколлинеарности «приоритетных» КПЭ процессов грузовых перевозок и обеспечивающих их процессов смежных хозяйств и дирекций. В качестве способа построения моделей используется множественная регрессия [25], пакеты прикладных программ Statistica [26], Microsoft Excel [27].

10 Валидация построенных моделей цифровых двойников процессов энергообеспечения посредством их сопоставления с фактическими процессами энергообеспечения тяги поездов на участках ОАО «РЖД». В качестве метода валидации могут применяться статистические методы перекрестной проверки (кросс-валидация) [28, 29].

Результаты и обсуждения

Реализация *процедуры 1* методологии по построению обобщенной структуры взаимодействия процесса энергообеспечения тяги поездов дирекцией Трансэнерго на одном из территориальных филиалов ОАО «РЖД» представлена на рис. 1. Данная структура, построенная согласно рекомендациям [14], представляет собой строку матричного представления перевозочных процессов и процессов его обеспечивающих, предложенных авторами и развитых в работах [16–18].

В схеме рис. 1 управление процессом грузовых перевозок осуществляется Территориальной дирекцией управления движения (ТДУД). Энергообеспечение процесса грузовых перевозок осуществляется Территориальной дирекцией энергообеспечения (ТЭ – Трансэнерго). К хозяйствам и дирекциям, обеспечивающим грузовые перевозки, относятся территориальные дирекции: Территориальная дирекция инфраструктуры (ТДИ), включающая хозяйство пути (П), вагонное хозяйство (В), хозяйство сигнализации и связи (Ш); Территориальная дирекция тяги (ТДТ) и некоторые другие. На схеме рис. 1 восходящими стрелками от процесса грузовых перевозок к блокам обеспечивающих дирекций передается информация о потребностях в ресурсах перевозочного процесса. Нисходящими стрелками от блоков обеспечивающих дирекций к процессу грузовых перевозок указаны ресурсы, которые фактически привносятся в процесс грузовых перевозок. Уровень привносимого ресурса от каждой дирекции целесообразно оценивать, как в абсолютных значениях, так и в «условных» КПЭ хозяйства, отнесенных к КПЭ процессов грузовых перевозок, которые курируются дирекцией ТДУД. При этом каждый показатель из вектора показателей, например, КПЭ ТЭ/ТДУД, будет иметь конкретный физический смысл. Так, очевидно, что для роста показателя ТДУД грузооборота от дирекции ТЭ потребуется увеличение показателя энергообеспечения. Отношение же данных показателей позволит оценить показатель эффективности расхода энергообеспечения. В предложенной структуре предполагается наличие обмена информацией непосредственно между дирекциями, обеспечивающими грузовые перевозки, например, между Трансэнерго и ТДТ – дирекцией тяги.

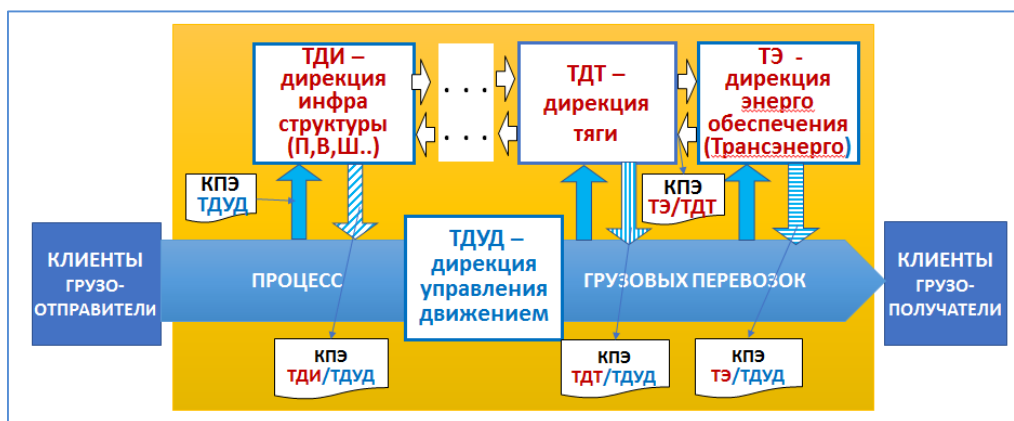


Рис. 1. Структурная схема взаимодействия основного процесса грузовых перевозок с обеспечивающими процессами территориальных дирекций для учета информационных потоков в цифровых двойниках

Реализация *процедуры 2* методологии по подбору КПА для включения их в качестве переменных в модель цифровых двойников процессов энергообеспечения тяги поездов была впервые апробирована авторами в работах [16–18]. В процедуре использовались рекомендации, представленные в [15]. Для получения наиболее информативных показателей процессов энергообеспечения тяги поездов со стороны СТЭ целесообразно использовать показатели, измеряемые на тяговых подстанциях (ТП). К данным показателям относятся в первую очередь:

- 1 Расход электроэнергии на тягу поездов по показаниям на ТП, кВтч (*Расход $W_{т\text{тп}}$*).
- 2 Удельный расход электроэнергии на тягу поездов по показаниям на ТП, кВтч / 10 тыс. ткм брутто (*Уд. Расход $W_{т\text{уд}}$*).
- 3 Небаланс электроэнергии, % (*Небаланс $W_{\text{неб}}$*).

При этом показатель *Уд. Расход $W_{т\text{уд}}$* является показателем эффективности использования электроэнергии, расходуемой на тягу поездов, и традиционно вычисляется как отношение абсолютного значения *Расход $W_{т\text{тп}}$* , отнесенное к объему выполняемой тонно-километровой работы брутто (A_t) на исследуемом участке.

Показатель *Небаланс $W_{\text{неб}}$* формируется как разница между показателями расхода тягового электропотребления по показаниям ТП и показателями тягового электропотребления на подвижном составе. На рис. 2 представлены реализации помесячных значений данных фактических процессов в СТЭ совместно с показателем работы A_t за период 46 кварталов (138 месяцев).

На рис. 2 видно наличие схожести характера реализации *Расход $W_{т\text{тп}}$* (тренды 2 и 4) и реализации A_t (тренды 1 и 2), что позволяет использовать при построении моделей цифровых двойников процессов СТЭ методы идентификации, а также статистического и корреляционного анализа [4, 5]. Из анализа характера реализации показателя *Уд. Расход $W_{т\text{уд}}$* виден её периодический и сезонный характер. При этом минимальные, а значит, и наиболее эффективные их значения приходятся на летние месяцы. Причем периоды данного мультипликативного показателя не совпадают с периодами колебаний показателей – *Расход $W_{т\text{тп}}$* и A_t , используемых для его традиционной оценки. Данное обстоятельство говорит о наличии в схеме формирования *Уд. Расход $W_{т\text{уд}}$* и других дополнительных эксплуатационных показателей, требующих идентификации и включения в модели цифровых двойников СТЭ.

К показателям качества КПА процессов грузовых перевозок и обеспечивающих их процессов смежных хозяйств и дирекций в обобщенной структурной схеме (см. рис. 1) следует отнести около 20 эксплуатационных показателей производственно-хозяйственной деятельности филиала ОАО «РЖД». Конкретный перечень КПА будет представлен ниже в процедуре 6.

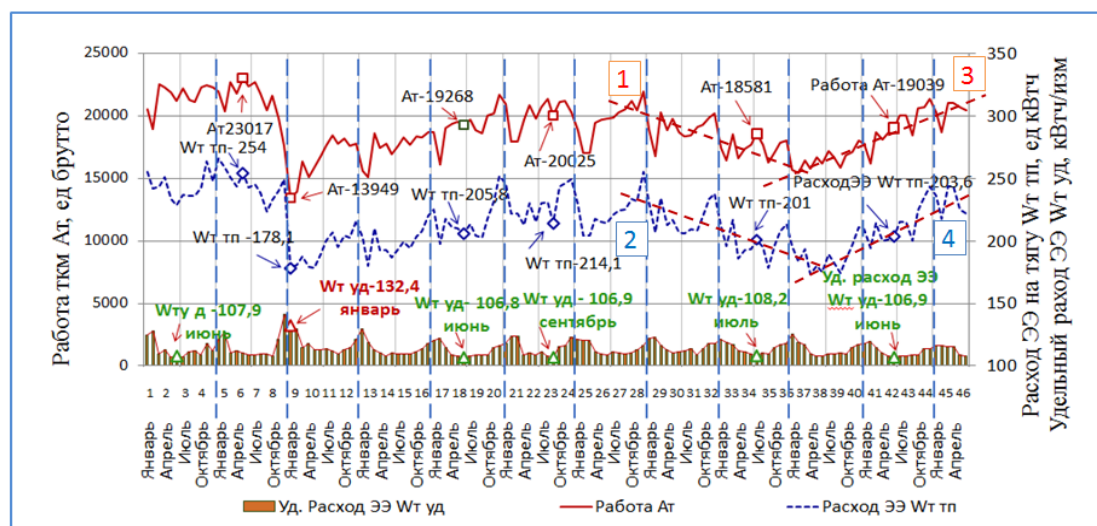


Рис. 2. Реализации фактических значений показателей процессов в СТЭ:
Расход $W_{тп}$, Уд. Расход $W_{уд}$ и тонно-километровой работы A_t brutto

Реализация *процедуры 3* методологии по формализации закономерностей, связывающих КПЭ процесса энергообеспечения тяги поездов с КПЭ процессов грузовых перевозок и обеспечивающих их процессов смежных хозяйств и дирекций на основании [16–18], может быть представлена в виде операторной формы записи:

$$X_{ТЭ}^* = F_{ТЭ}(X_{ТДУД}, X_{ТДИ}, X_{ТДТ}) + \xi_{ТЭ}, \quad (1)$$

где $X_{ТЭ}^*$ – модельная оценка вектора выходного процесса энергообеспечения тяги поездов для блока схемы ТЭ; $X_{ТДУД}$, $X_{ТДИ}$, $X_{ТДТ}$, $X_{ТЭ}$ – векторы наблюдаемых выходных фактических показателей блоков схемы ТДУД, ТДИ, ТДТ, ТЭ; $F_{ТЭ}$ – функционал взаимосвязи между выходными показателями блока ТЭ и выходными показателями блоков ТДУД, ТДИ, ТДТ, ТЭ, полученных методами идентификации [4, 5]; $\xi_{ТЭ}$ – векторы стохастической составляющей цифровой модели из-за неучтенных факторов.

Реализация *процедуры 4* методологии основана на мониторинге и формировании цифровых баз данных актуальных КПЭ процессов согласно рекомендациям методов [18, 19] и формата, представленного в табл. 1.

Таблица 1

Вид фрагмента цифровой базы КПЭ процессов грузовых перевозок и энергообеспечения тяги поездов

№ набл (квартал)	№ КПЭ						...	19
	1	2	3	4	5	6		
	Погрузка	Грузо-оборот Тарифный	Грузо-оборот Экспл	Работа ткм A_t	Расход ЭЭ $W_{тп}$	Участковая Скорость $V_{уч}$		Производи- тельность вагона
1	5315	12800	11909	20556	255	42		14682,0
2	5884	14380	13601	22269	251	40		13620,0
...								
45	5961	15708	12361	20320	236	45		6400,0
46	5972	16054	13064	21028	243	44		6762,0

Для этого целесообразно организовать мероприятия по подбору, выгрузке и трансформации данных КПЭ из доступных для исследователей железнодорожных систем управления, например, АСОУП, ИОММ, ГИД-Урал, АСУ-Т, АСУ-Э, ТХО-125, ДУ и др. Количество наблюдаемых значений КПЭ (длительность выборки), включаемых в базу данных, определяется поставленной задачей и характером исследуемых процессов. В условиях помесечной сезонности данных для обеспечения высокой точности прогноза выборка должна содержать уже 24–36 помесечных наблюдений КПЭ. В качестве примера в табл. 1 представлен фрагмент цифровой базы для одного из филиалов ОАО «РЖД» для 19 КПЭ энергообеспечения

тяги поездов, грузовых перевозок и обеспечивающих их процессов смежных хозяйств и дирекций за период 46 кварталов (138 месяцев), составленной в табличном процессоре Excel.

Реализация *процедуры 5* методологии по оценке статистических показателей КПЭ процессов энергообеспечения, грузовых перевозок и обеспечивающих их процессов со стороны смежных хозяйств и дирекций основана на анализе ранее сформированной базы данных КПЭ вида табл. 1. [18, 19] В рамках процедуры следует провести анализ законов плотности вероятности выборок КПЭ исследуемых процессов. На рис. 3 представлены примеры гистограмм эмпирической плотности вероятности КПЭ процессов энергообеспечения тяги поездов: *Работа тонно-километровая A_t* , *Расход $W_{тп}$* , *Уд. Расход $W_{туд}$* на одном из направлений тяжеловесного движения за период 24 месяцев. Поскольку представленные на рис. 3, а, б, в эмпирические гистограммы отличаются от плотности вероятности нормального закона (непрерывные кривые) и имеют ассиметричный характер, было предложено использовать показатели: *Мода*, который указывает на наиболее вероятные значения выборки КПЭ; *Коэффициент Ассиметрии K_a* , который оценивает отклонение гистограммы вправо или влево относительно среднего значения; *Коэффициент Эксцесса $K_э$* , который позволяет оценить островершинный или плоский характер плотности вероятности анализируемой выборки показателей.

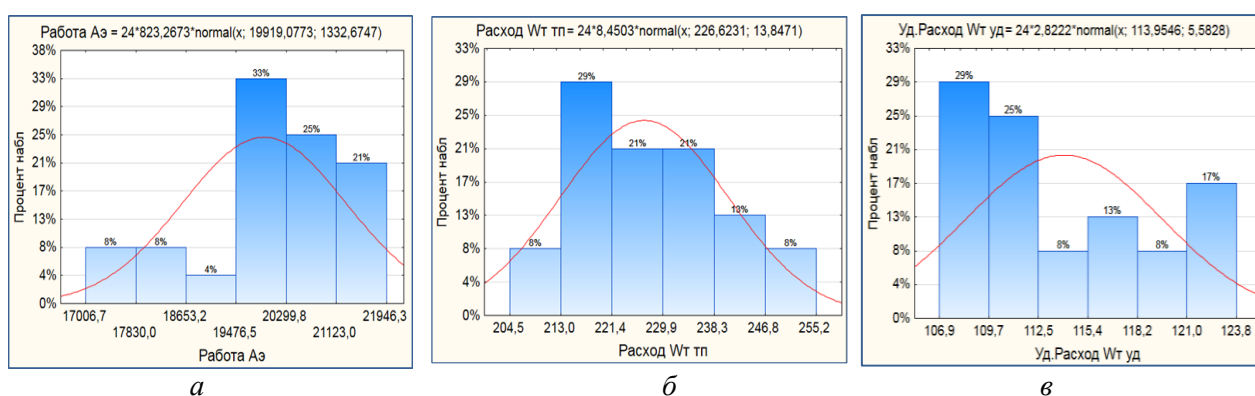


Рис. 3. Пример гистограмм КПЭ процессов:

а – работа тонно-километровая A_m ; б – расход $W_{тп}$; в – удельный расход $W_{туд}$

Реализация *процедуры 6* методологии заключается в идентификации взаимосвязи КПЭ исследуемых для включения в математические модели цифровых двойников процессов энергообеспечения методами корреляционного [4, 5] и кластерного анализов [21], а также выявление при этом «приоритетных» КПЭ, значительно влияющих на энергообеспечение тяги поездов [16–18]. Идентификацию взаимосвязей исследуемых КПЭ предлагается выполнять на основании построения парных коэффициентов корреляции (R) между КПЭ процессов энергообеспечения тяги поездов и КПЭ процессов грузовых перевозок и обеспечивающих их процессов смежных хозяйств и дирекций.

На рис. 4, а представлен пример построения гистограмм коэффициентов корреляции между показателем *Расход $W_{тп}$* – $R(W_{тп})$, а на рис. 4, б – коэффициентов корреляции между показателем *Уд. Расход $W_{туд}$* – $R(W_{туд})$ и КПЭ других хозяйств.

Анализ составленной гистограммы (см. рис. 4, а) показывает, что КПЭ процесса энергообеспечения *Расход $W_{тп}$* на исследуемом полигоне ОАО «РЖД» сильнее всего связан со следующими КПЭ: для *Расход $W_{эпс}$* значение $R(W_{тп}) = 0,92$; для *Работа A_z* значение $R(W_{тп}) = 0,88$; для *Погрузка* значение $R(W_{тп}) = 0,52$; для *Участковая скорость $V_{уч}$* ; *Средний вес поезда*; *Оборот грузового вагона*; *Статическая нагрузка* значение $R(W_{тп}) = -0,31$ и т. д. Исходя из данных гистограммы (см. рис. 4, б) КПЭ *Уд. Расход $W_{туд}$* на исследуемом полигоне ОАО «РЖД» связан с его составляющими следующим образом: для *Расход $W_{тп}$* значение $R(W_{туд})$ составляет не более 0,14; для *Работа A_z* значение $R(W_{туд}) = -0,34$. Это означает, что показатель *Уд. Расход $W_{туд}$* более чем в два раза сильнее зависит от изменения показателя *Работа A_z* , чем от изменения показателя *Расход $W_{тп}$* . Далее КПЭ процесса энергообеспечения *Уд. Расход $W_{туд}$* сильнее всего связан со следующими КПЭ: для *Процент порожнего пробега вагона* значение $R(W_{туд}) = 0,60$; для *Участковая скорость $V_{уч}$* значение $R(W_{туд}) = 0,33$ и т. д.

Для корректного выбора «приоритетных» КПЭ предложено использовать кластерные методы анализа взаимосвязи переменных. На рис. 4, в результаты обработки КПЭ девятнадцати эксплуатационных процессов на исследуемом филиале ОАО «РЖД» представлены в виде дендрограммы.

В дендрограмме степень близости КПЭ изображается в виде уровней объединения одиночных переменных или их групп (кластеров), отстоящих от нулевой оси абсцисс *Расстояние объединений* на

расстояния, пропорциональные разнице между единицей и коэффициентом корреляции R по Пирсону, то есть величине $(1-R \text{ Пирсон})$. При этом если уровень объединений КПЭ или их кластеров находится по оси *Расстояние объединений* на уровне 0, то это соответствует полной положительной взаимосвязи между КПЭ, то есть $R = 1$. Если уровень объединения КПЭ находится на уровне 1, то это соответствует полному отсутствию взаимосвязи между КПЭ, то есть $R = 0$. Если уровень объединения КПЭ равен 2, то это соответствует полной отрицательной взаимосвязи между КПЭ, то есть $R = -1$.

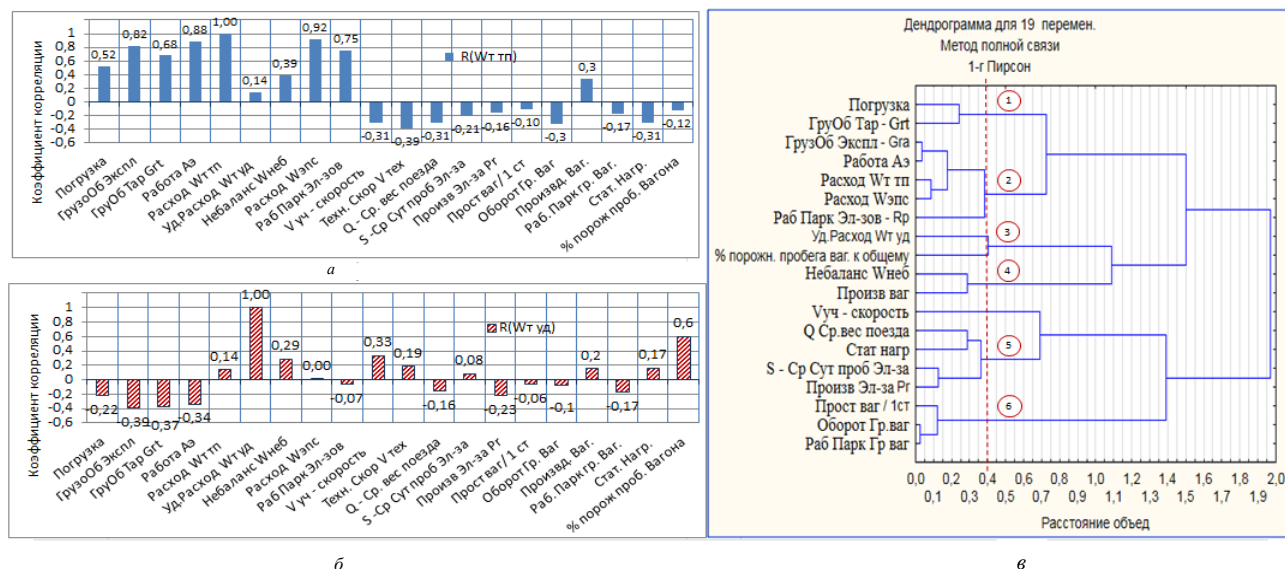


Рис. 4. Результаты идентификации взаимосвязей КПЭ:

а – гистограмма коэффициентов корреляции $R(W_{TP})$;

б – гистограмма коэффициентов корреляции $R(W_{UD})$; в – дендрограмма процессов

Таким образом, чем ближе уровень объединения КПЭ или их кластеров к нулевой отметке, тем выше положительная корреляционная связь. Так, на рис. 4, в на уровне значимых величин коэффициента корреляции $R > 0,4$ по оси *Расстояние объединений* можно различить шесть кластеров, обозначенных цифрами в окружности:

– кластер 1 включает в себя показатели: *Погрузка*; *ГрузОб Тар-Грт*. Уровень взаимосвязи показателей по оси *Расстояние объединений* составляет 2,5 единиц, что соответствует коэффициенту корреляции $R = 0,75$;

– кластер 2 включает в себя показатели: *Расход W_{TP}* ; *Расход $W_{эпс}$* ; *Работа Аз*; *ГрузОб Экспл – Gra*; *Раб Парк Эл-зов – R_p*. Уровень взаимосвязи $R = 0,6$;

– кластер 3 включает в себя показатели: *Уд. Расход $W_{т уд}$* ; *% Порожного пробега вагона к общему*. Уровень взаимосвязи $R = 0,7$;

– кластер 4 включает в себя показатели: *Небаланс $W_{неб}$* ; *Производительность вагона*. Уровень взаимосвязи $R = 0,6$;

– кластер 5 включает в себя показатели: *Q – Ср. вес поезда*; *Стат нагр*; *S – Средне суточный пробег Эл-за*; *Производительность Эл-за*. Уровень взаимосвязи $R = 0,65$;

– кластер 6 включает в себя показатели: *Простой вагона на 1 станции*; *Оборот Грузового Вагона* и *Рабочий Парк Грузового вагона*. Уровень взаимосвязи $R = 0,87$.

Реализация процедуры 7 методологии предусматривает построение комплекса однофакторных экспресс-моделей цифровых двойников процессов энергообеспечения тяги поездов в функции «приоритетных» КПЭ процессов грузовых перевозок и обеспечивающих их хозяйств и дирекций. При построении и верификации данного комплекса моделей используются рекомендации, представленные в методах [23, 24]. Комплекс данных экспресс-моделей может иметь вид

$$\left. \begin{aligned} X_{TЭ1}^* &= F_{TЭ}(X_{ТДУД}) + \xi_{TЭ} \\ X_{TЭ2}^* &= F_{TЭ}(X_{ДТИ}) + \xi_{TЭ} \\ &\dots \\ X_{TЭN}^* &= F_{TЭ}(X_{ТДТ}) + \xi_{TЭ} \end{aligned} \right\}. \quad (2)$$

На основании (2) целесообразно производить построение графиков взаимосвязи КПЭ для последующей их верификации (рис. 5). На рис. 5 представлена матрица графиков гистограмм и корреляционных полей точек взаимосвязи ряда исследуемых КПЭ.

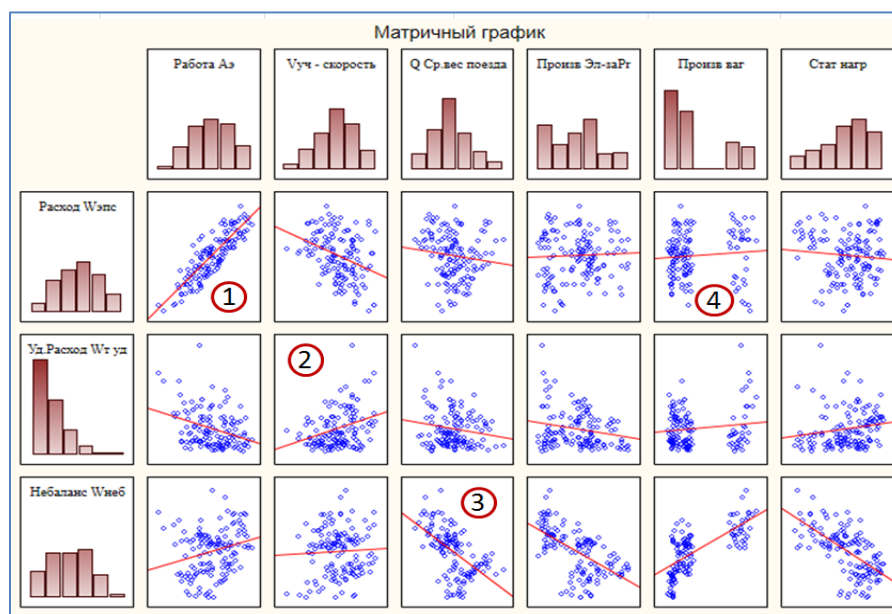


Рис. 5. Пример матричного графика гистограмм и поля точек взаимосвязи КПЭ

На пересечении горизонтальных строк и вертикальных столбцов с видом гистограмм КПЭ представлены корреляционные поля точек их взаимосвязи. Чем меньше степень разброса корреляционного поля точек относительно линий их аппроксимации и чем линия аппроксимации ближе к горизонтальной линии, тем меньше взаимосвязь между исследуемыми КПЭ и тем меньше адекватность построенной экспресс-модели. Так, например, на графике корреляционного поля точек КПЭ *Расход $W_{тпс}$* и КПЭ *Работа A_z* (номер 1, рис. 5) отмечена низкая степень их разброса относительно линии аппроксимации, а также высокий наклон последней. Делается вывод об адекватности данной экспресс-модели и о возможности включения КПЭ *Работа A_z* в математическую модель цифрового двойника энергообеспечения тяги поездов. К таким же адекватным экспресс-моделям для КПЭ *Уд. Расход $W_{т уд}$* и КПЭ *Небаланс $W_{неб}$* по обозначенным выше критериям можно отнести графики под номером 2 и 3.

В то же время, например, корреляционное поле между КПЭ *Расход $W_{т тп}$* и КПЭ *Производительность вагона* (номер 4, рис. 5) носит «раздвоенный» характер вследствие бимодальной гистограммы и нестационарности реализаций на горизонте исследований. При этом включение подобных КПЭ в модель цифрового двойника не будет отвечать требованиям верификации. Кроме того, верификацию экспресс-моделей необходимо проводить по критерию физической и аналитической обоснованности. В качестве примера проверки адекватности по данному критерию могут служить экспресс-модели, представленные на рис. 6, а, б, в. Так, на рис. 6, а зависимости КПЭ *Расход $W_{т тп}$* и КПЭ *Уд. Расход $W_{т уд}$* , построенные в функции КПЭ *Работа A_z* , содержат параметры однофакторных регрессионных уравнений их взаимосвязи и количественную оценку степени адекватности в формате коэффициента корреляции R .

Обе зависимости удовлетворяют требованию верификации по критерию физической и аналитической обоснованности. Данный вывод сделан на основании того, что действительно с ростом выполняемой тонно-километровой работы A_z на участке объем потребленной электроэнергии, расходуемой на тягу поездов – *Расход $W_{т тп}$* , должен также возрастать коэффициент корреляции ($R = 0,88$).

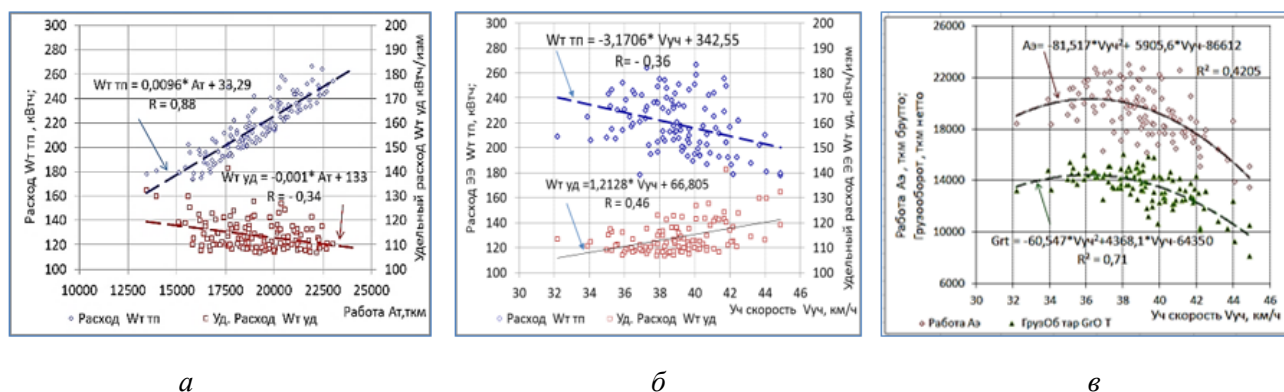


Рис. 6. Пример построения экспресс-моделей КПЭ $W_{т\text{тп}}$, $W_{т\text{уд}}$ в функции:

а – КПЭ Работа A_3 ; б – Участковая скорость – $V_{уч}$;

в – построение экспресс-моделей КПЭ Работа A_3 и Грузооборот – Grt в функции Участковая скорость – $V_{уч}$

Зависимость КПЭ $W_{т\text{уд}}$ в функции КПЭ Работа A_3 также адекватна исходя из структуры формирования показателя. С ростом выполняемой тонно-километровой работы A_3 значение показателя будет уменьшаться – коэффициент корреляции ($R = -0,34$). Однако если обратиться к экспресс-модели КПЭ $W_{т\text{тп}}$, построенной в функции КПЭ Участковая скорость $V_{уч}$ (рис. 6, б), то окажется, что её линия тренда обратно пропорциональна участковой скорости $V_{уч}$. – коэффициент корреляции ($R = -0,36$). Данное обстоятельство не согласуется с физикой процесса перемещения грузовых поездов. Очевидно, что потребность в энергии, расходуемой на перемещение поездов, должна с ростом скорости не убывать, а возрастать пропорционально квадрату скорости перемещения.

Разъяснение данного противоречия состоит в следующем: повышение Участковой скорости $V_{уч}$ является не первопричиной, а следствием снижения грузопотока поездов на полигоне. На рис. 6, в представлены графики зависимости КПЭ Работа A_3 и Грузооборот тарифный (Grt) в функции от КПЭ Участковая скорость $V_{уч}$, ($V_{уч}$ – скорость). Из графиков видно, что максимальная тонно-км работа A_3 на исследуемом полигоне достигалась в объеме 23000 ткм брутто при скорости около 38 км/час. При снижении количества грузовых поездов на участке нелинейно снижаются и КПЭ Работа A_3 и КПЭ Грузооборот – Grt . Как следствие, снижение данных показателей приводит к увеличению межпоездных интервалов, что позволяет грузовым поездам увеличить техническую и участковую скорость. В рассматриваемом примере при снижении КПЭ Работа A_3 от максимально зарегистрированных 20000 до 13000 ед. ткм бр. (то есть на –44 %) участковая скорость $V_{уч}$ возросла до 45 км/час (+18 %). Таким образом, рост участковой скорости $V_{уч}$ с 38 до 45 км/час (рис. 6, б) является следствием уменьшения КПЭ Работа A_3 (рис 6, в), что и приводит к уменьшению КПЭ $W_{т\text{тп}}$ с 225 до 200 ед. кВтч (–11 %), что и отражено на рис. 6, б. Приведенное выше рассуждение показывает, что использование зависимости экспресс-модели КПЭ $W_{т\text{тп}}$ в функции от участковой скорости $V_{уч}$ в отрыве от динамики других КПЭ будет не адекватной для использования в модели цифровых двойников процессов энергообеспечения тяги поездов.

Реализация процедуры 8 методологии предусматривает построение динамических реализаций КПЭ процессов энергообеспечения тяги поездов и КПЭ процессов грузовых перевозок хозяйств, обеспечивающих грузовые перевозки [4]. Пример построения подобных реализаций представлен на рис. 2. Однако также целесообразно проводить построение тех реализаций КПЭ процессов, которые посредством анализа их динамики позволяют определить степень их стационарности. Для построения реализаций возможно воспользоваться данными табл. 1, а для оценки стационарности – статистическими показателями КПЭ. Так, анализ КПЭ Производительность вагона (см. табл. 2) показывает на более чем 200 %-ное различие значений показателя в интервалах № 1 и 2 (14682, 13620) по отношению к значениям интервалов № 45 и 46 (6400, 6762), что связано с изменением методики расчета показателя. Таким образом, данный показатель не отвечает требованиям стационарности и верификации.

Реализация процедуры 9 методологии предусматривает:

– агрегирование данных для переменных для многофакторных моделей цифровых двойников из состава ранее исследуемых «приоритетных» КПЭ;

- идентификацию параметров КПЭ, включенных в состав многофакторных моделей цифровых двойников;
- верификацию полученных многофакторных моделей цифровых двойников посредством устранения мультиколлинеарности «приоритетных» КПЭ процессов грузовых перевозок и обеспечивающих их процессов смежных хозяйств и дирекций;
- построение многофакторных моделей цифровых двойников КПЭ процессов энергообеспечения тяги поездов с определением адекватных показателей.

В качестве способа построения моделей используется множественная регрессия [25], пакеты прикладных программ Statistica [26].

Агрегирование данных для переменных для математических многофакторных моделей цифровых двойников, структура которых определена выражением (1), целесообразно производить из состава «приоритетных» КПЭ, ранее исследуемых в экспресс-моделях вида (2). В качестве формы математических многофакторных моделей цифровых двойников предлагается использовать форму линейных многофакторных уравнений регрессии [25]. Опираясь на состав переменных, входящих в выражение (1), выражение, например, для КПЭ *Расход* $W_{т\text{тп}}$, *Уд.* *Расход* $W_{т\text{уд}}$ и % *Небаланс* $W_{неб}$, можно представить в виде:

$$W_{т\text{тп}}^* \Delta t = a_0 + (a_1 \cdot X_{тдуд} \Delta t + a_2 \cdot X_{тдт} \Delta t + \dots + a_m \cdot X_{дти} \Delta t) + \xi_{т\text{тп}}, \quad (3)$$

$$W_{туд}^* \Delta t = b_0 + (b_1 \cdot X_{тдуд} \Delta t + b_2 \cdot X_{тдт} \Delta t + \dots + b_m \cdot X_{дти} \Delta t) + \xi_{туд}, \quad (4)$$

$$W_{неб}^* \Delta t = c_0 + (c_1 \cdot X_{тдуд} \Delta t + c_2 \cdot X_{тдт} \Delta t + \dots + c_m \cdot X_{дти} \Delta t) + \xi_{неб}, \quad (5)$$

где $W_{т\text{тп}}^* \Delta t$, $W_{туд}^* \Delta t$, $W_{неб}^* \Delta t$ – расчетные дискретные (цифровые) оценки фактических дискретных значений процесса показателей $W_{т\text{тп}} \Delta t$, $W_{туд} \Delta t$, $W_{неб} \Delta t$ во временном интервале Δt ; $X_{тдуд} \Delta t$, $X_{тдт} \Delta t$, $\dots$, $X_{дти} \Delta t$ – фактические дискретные наблюдаемые значения «приоритетных» КПЭ процессов территориальных дирекций ТДУД, ТДТ, ТДИ во временном интервале Δt ; a_0 , b_0 , c_0 – свободные члены линейных многофакторных уравнений регрессии (3)–(5); a_1 , a_2 , $\dots$, a_m ; b_1 , b_2 , $\dots$, b_m ; c_1 , c_2 , $\dots$, c_m – параметры моделей (коэффициенты линейных многофакторных уравнений регрессии (3)–(5) при КПЭ процессов, поиск которых осуществляется методами идентификации [6, 7];

$\xi_{т\text{тп}}$, $\xi_{туд}$, $\xi_{неб}$ – стохастические составляющие цифровых моделей уравнений регрессии (3)–(5) из-за неучтенных факторов.

Для корректного агрегирования переменных в уравнениях вида (3)–(5) должен быть устранен эффект мультиколлинеарности переменных, который основан на подборе и использовании в многофакторном уравнении регрессии переменных с уровнем их взаимной корреляционной связи не более $R = 0,6$. В предлагаемой методологии устранение эффекта мультиколлинеарности переменных производится посредством использования дендрограммы переменных, выполненной методом кластерного анализа (см. рис. 4, в). При этом из дендрограммы выбираются те переменные, которые: принадлежат к разным группам кластеров; наиболее подходят к независимой переменной по физическому смыслу из состава переменных, включенных в кластер. Так, для зависимого КПЭ *Расход* $W_{т\text{тп}}$ при подборе независимых переменных в уравнении вида (3) следует рассматривать кластеры из дендрограммы рис. 4, в, уровень взаимосвязи которых будет $R \leq 0,6$. Так, из кластера 1 может быть выбран только один из сильно коррелированных КПЭ ($R = 0,75$), например, «Погрузка». Из кластера 2 целесообразно выбрать независимый КПЭ *Работа А*, и т. д. Таким образом, в состав переменных предварительного уравнения для выражения (3) были включены 15 КПЭ, характеризующих энергообеспечение тяги поездов и КПЭ из состава кластеров рис. 4, в. Результаты идентификации параметров данного уравнения представлены в табл. 2.

Таблица 2

**Результаты идентификации параметров пятнадцатифакторного уравнения регрессии
для цифрового двойника КПЭ *Расход* $W_{т.тп}$**

№	КПЭ	Скорректир. коэффициент регрессии Beta	Коэффициенты регрессионного уравнения $a_0, a_1, a_2, \dots, a_m$	Достигнутый p-level при идентификации (допустимый $p = 0,05$)
0	<i>Свободный член уравнения a_0</i>		-413,387	0,000002
1	Погрузка	-0,018	-0,001	0,758316
2	<i>Грузооборот Тарифный – Grt</i>	0,322	0,005	0,015433
3	Грузооборот Эксплуатационный – Gra	-0,420	-0,007	0,094373
4	<i>Работа тонно-км А Эл-зов</i>	0,903	0,010	0,000001
5	<i>Небаланс Wнеб</i>	0,271	175,593	0,000001
6	Рабочий Парк Эл-зов – Rp	0,228	0,101	0,212433
7	$V_{уч}$ – скорость участковая	0,118	1,000	0,111155
8	Q – Средний вес поезда	-0,102	-0,021	0,120915
9	S – Срeнесуточн пробег Эл-за	0,103	0,041	0,354482
10	Производительность Эл-за – Pr	-0,072	-0,010	0,682518
11	Простой вагона на 1-ой станции	-0,048	-0,439	0,658401
12	Оборот Грузового вагона	0,545	5,762	0,101299
13	<i>Производительность вагона</i>	0,845	0,006	0,000000
14	Рабочий Парк Грузового вагона	0,298	0,000	0,482537
15	<i>Статическая нагрузка на ось</i>	0,267	5,757	0,000106

Идентификация параметров КПЭ в уравнении вида (3) производилась в пакете программ Statistica [26]. В табл. 2 представлены следующие показатели: **Beta** – скорректированный коэффициент регрессии, характеризующий степень влияния КПЭ; коэффициенты регрессионного уравнения (3) для моделирования КПЭ *Расход* $W_{т.тп}$; **p-level** – достигнутый в результате идентификации уровень значимости и сравниваемый с допустимым уровнем $p = 0,05$.

Точность 15-факторной модели показателя *Расход* $W_{т.тп}$ характеризуется при этом следующими показателями: множественный коэффициент корреляции (*Multiple R*) = 0,954; множественный коэффициент детерминации (*Multiple R²*) = 0,911; скорректированный коэффициент детерминации (*Adjusted R²*) = 0,899; критерий Фишера $F(15,122) = 82,9$, что намного больше нормативного $F_{табл} = 1,7$ при уровне значимости $p < 0,05$; стандартная ошибка оценки (*Std. Err. of Estimate*) = 7,032; средняя абсолютная ошибка прогноза (*MAPE*) по всей реализации = 2,26 %. Однако, несмотря на полученную высокую точность согласно методике, изложенной для пакета программ Statistica, пригодными для дальнейшего использования могут быть только те КПЭ, у которых достигнутый p-level при идентификации будет меньше принятого при расчетах, т. е. $p < 0,05$. В табл. 2 таких переменных только пять (выделены курсивом и красным цветом).

В рамках этапа методологии была произведена оптимизация составленного 15-факторного регрессионного уравнения (3), которое заключалось в последовательном исключении из уравнения переменных КПЭ с показателем уровня значимости меньше допустимого, т. е. $p\text{-level} \leq 0,05$. Кроме того, оптимизация КПЭ уравнения должна производиться с учетом ранее описанных методов выбора переменных при использовании дендрограммы (см. рис. 4, в) и результатов ранее описанного анализа корректности экспресс-моделей. Подбор параметров с устранением эффекта мультиколлинеарности допустимо производить также в автоматизированном режиме в пакете программ Statistica из состава КПЭ, предварительно включенных в состав многофакторных моделей вида (3).

Параметры одного из возможных вариантов полученного четырехфакторного регрессионного уравнения после процедуры оптимизации представлены в табл. 3.

Таблица 3

**Результаты параметров четырехфакторного уравнения регрессии
для цифрового двойника КПЭ $Расход W_{т\text{тп}}$**

№	КПЭ	Скорректир. коэффициент регрессии Beta	Коэффициенты регрессионного уравнения (3) $a_0, a_1, a_2, \dots, a_m$	Достигнутый p-level – при идентификации (допустимый $p = 0,05$)
0	<i>Свободный член уравнения a_0</i>		107,371	0,002
1	<i>Работа тонно-км А Эл-зов</i>	0,964	0,010	0,000
2	<i>$V_{уч}$ – скорость участковая</i>	0,280	2,364	0,000
3	<i>Q – Средний вес поезда</i>	–0,243	–0,051	0,000
4	<i>Простой вагона на 1-й станции</i>	0,150	1,366	0,003

Точность модели показателя $Расход W_{т\text{тп}}$ при данной конфигурации регрессионного уравнения несколько хуже, чем у исходной 15-факторной, но такая модель более устойчива к вариации данных.

Характеристики адекватности 4-факторного уравнения регрессии для цифрового двойника КПЭ процесса $Расход W_{т\text{тп}}$: множественный коэффициент корреляции составляет ($Multiple R$) = 0,908; множественный коэффициент детерминации ($Multiple R^2$) = 0,824; скорректированный коэффициент детерминации ($Adjusted R^2$) = 0,819; критерий Фишера $F(4,133) = 156,13$, что намного больше нормативного $F_{табл} = 2,44$ при уровне значимости $p < 0,05$; стандартная ошибка оценки ($Std. Err. of Estimate$) = 9,438.

Для оценки параметров КПЭ $Уд. Расход W_{т\text{уд}}$ в уравнении вида (4) могут быть использованы другие переменные территориальных дирекций ТДУД, ТДТ, ТДИ, представленные на схеме рис. 1.

Для определения параметров уравнения (4) и последующей верификации зависимостей КПЭ $Уд. Расход W_{т\text{уд}}$ было составлено предварительное уравнение из 12 КПЭ, представленных в дендрограмме на рис. 4, в. Предполагается, что из состава переменных предварительного уравнения должны быть исключены КПЭ, характеризующие энергообеспечение тяги поездов. Результаты идентификации параметров предварительного уравнения для моделирования КПЭ $Уд. Расход W_{т\text{уд}}$ представлены в табл. 4.

Таблица 4

**Результаты идентификации параметров 12-факторного уравнения регрессии
для цифрового двойника КПЭ $Уд. Расход W_{т\text{уд}}$**

№	КПЭ	Скорректир. коэффициент регрессии Beta	Коэффициенты регрессионного уравнения (4) $b_0, b_1, b_2, \dots, b_m$	Достигнутый p-level – при идентификации (допустимый $p = 0,05$)
0	Свободный член уравнения a_0		–76,4509	0,070459
1	Работа тонно-км А электровозов	–0,359404	–0,0011	0,077635
2	Рабочий Парк электровозов – Рр	0,501556	0,0605	0,106053
3	$V_{уч}$ – скорость участковая	0,124752	0,2881	0,478320
4	Q – Средний вес поезда	–0,072825	–0,0041	0,653385
5	S – Среднесуточн пробег электровоза	0,382222	0,0416	0,096793
6	Производительность электровоза – Рг	–0,160959	–0,0061	0,604625
7	Простой вагона на 1-й станции	0,347630	0,8642	0,111061
8	<i>Статическая нагрузка на ось</i>	0,381123	2,2403	0,004190
9	<i>% порожн. пробега ваг. к общему</i>	0,300355	0,9190	0,009918
10	Рабочий Парк Грузового вагона	–0,611552	–0,0002	0,502525
11	Производительность вагона	0,511288	0,0009	0,146255
12	Оборот Грузового вагона	0,473817	1,3698	0,487317

Точность представленной в табл. 4 модели показателя $Уд. Расход W_{т\text{уд}}$ характеризуется следующими показателями: множественный коэффициент корреляции составляет ($Multiple R$) = 0,756; множе-

ственный коэффициент детерминации ($Multiple R^2$) = 0,572; скорректированный коэффициент детерминации ($Adjusted R^2$) = 0,531; критерий Фишера $F(12,125) = 13,9$, что намного больше нормативного $F_{табл} = 1,8$ при уровне значимости $p < 0,05$; стандартная ошибка оценки ($Std. Err. of Estimate$) = 4,155.

Оптимизация данного 12-факторного регрессионного уравнения производилась методом, предусмотренным для пакета программ Statistica, и с учетом ранее описанных методов выбора переменных при использовании дендрограммы (рис. 4, в). Параметры одного из возможных вариантов полученного четырехфакторного регрессионного уравнения после процедуры оптимизации представлены в табл. 5.

Таблица 5

**Результаты параметров 4-факторного уравнения регрессии
для цифрового двойника КПЭ Уд. Расход $W_{туд}$**

№	КПЭ	Скорректир. коэффициент регрессии Beta	Коэффициенты регрессионного уравнения $b_0, b_1, b_2, \dots b_m$	Достигнутый p-level – при идентификации (допустимый $p = 0,05$)
0	<i>Свободный член уравнения a_0</i>		31,73057	0,013708
1	<i>Работа тонно-км А Эл-зов</i>	-0,367802	-0,00109	0,000454
2	<i>Рабочий Парк электровозов – Rp</i>	0,450876	0,05434	0,000013
4	<i>$V_{уч}$ – скорость участковая</i>	0,259226	0,59868	0,000525
3	<i>% порожн. пробега ваг. к общему</i>	0,494295	1,51247	0,000000

Характеристики адекватности 4-факторного уравнения регрессии для цифрового двойника КПЭ Расход $W_{туд}$: множественный коэффициент корреляции составляет ($Multiple R$) = 0,69; множественный коэффициент детерминации ($Multiple R^2$) = 0,48; скорректированный коэффициент детерминации ($Adjusted R^2$) = 0,46; критерий Фишера $F(4,133) = 30,13$, что намного больше нормативного $F_{табл} = 2,44$ при уровне значимости $p < 0,05$; стандартная ошибка оценки ($Std. Err. of Estimate$) = 4,462.

По аналогии с приведенным подходом может быть построена модель цифрового двойника для КПЭ Небаланс электроэнергетики, % (Небаланс $W_{неб}$).

Реализация процедуры 10 методологии предусматривает по результатам идентификации проведение валидации построенных моделей цифровых двойников, например, для КПЭ Расход $W_{тп}$ и для КПЭ Уд. Расход $W_{туд}$ (табл. 6 и рис. 7).

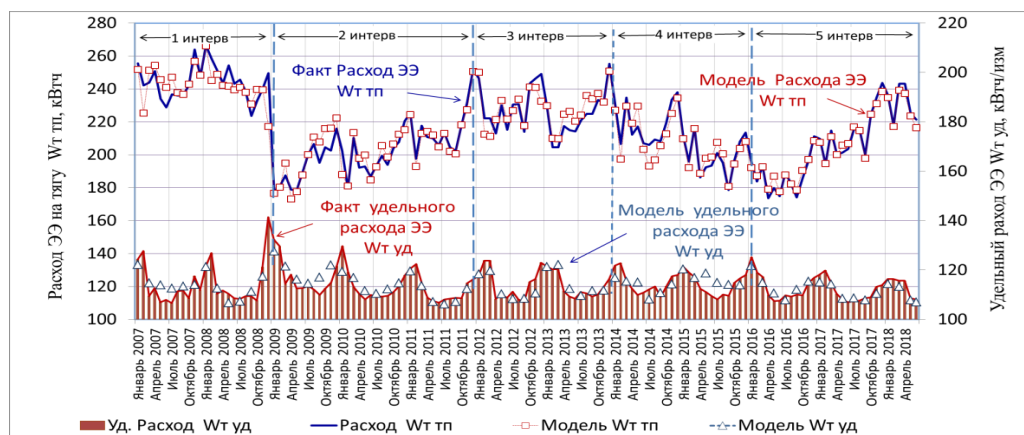
Таблица 6

**Результаты фактических, моделируемых значений реализации КПЭ Расход $W_{тп}$
и оценки адекватности моделирования**

Дата	Факт Расход ЭЭ $W_{тп}$	Модель Расход $W_{тп}$ (4 фак- тора)	Остаток (Факт – Модель) $W_{тп}$	Относит. ошибка модели $W_{тп}$	Факт Удельный Расход ЭЭ $W_{туд}$	Модель Удельный Расход $W_{туд}$ (4 фактора)	Остаток (Факт – Модель) $W_{туд}$	Относит. ошибка модели $W_{туд}$
1	255,4	240,3	15,1	5,9 %	124,3	118,8	5,5	4,4 %
2	242,2	217,8	24,4	10,6 %	127,7	115,3	12,4	9,7 %
...								
45	236,3	231,7	4,6	1,6 %	116,3	114,1	0,4	1,9 %
46	243,3	236,2	7,1	1,2 %	115,7	113,4	1,7	2,0 %
Минимум	173,7	169,5	-14,1	-8,5 %	106,7	104,0	-8,9	-8,4 %
Максимум	266,5	256,4	35,2	12,7 %	141,4	126,9	17,0	12,0 %
MAPE-оценка				3,6 %				3,1 %

Валидация согласно рекомендациям метода кросс-валидации [28, 29] предусматривает оценку точности моделирования показателей в сопоставлении с фактическими данными за весь период идентификации – 138 месяцев. О результатах оценки адекватности модельных реализаций КПЭ Расход $W_{тп}$ и КПЭ Уд. Расход $W_{туд}$ можно судить по данным, представленным в табл. 6 и на рис. 7.

Построенные модельные реализации процесса энергообеспечения тяги поездов КПЭ *Расход электроэнергии на тягу поездов по показаниям на ТП, кВтч (Расход $W_{т\text{ тп}}$)* были использованы в работе [30] для оценки ресурса нагрузочной способности СТЭ на одном из территориальных участков дирекции Трансэнерго ОАО «РЖД».



**Рис. 7. Реализации фактических и модельных значений
КПЭ *Расход $W_{т\text{ тп}}$ и Уд. *Расход $W_{т\text{ уд}}$****

Выводы

1 Использование на железнодорожном транспорте цифровых технологий потребовало разработать методологию построения моделей цифровых двойников процессов энергообеспечения тяги грузовых поездов для территориальных филиалов компании ОАО «РЖД», включающую свыше пятнадцати различных методов, способов и программных продуктов. Методология построена на основании матричной структуры взаимодействия процесса грузовых перевозок с обеспечивающими его процессами территориальных дирекций, в том числе и с процессом энергообеспечения тяги грузовых поездов.

2 В качестве переменных в математических моделях цифровых двойников процесса энергообеспечения тяги поездов целесообразно использовать «приоритетные» показатели процессов грузовых перевозок и показатели хозяйств, обеспечивающих грузовые перевозки, которые в наибольшей степени оказывают влияние на динамику энергообеспечения тяги поездов.

3 Для корректного построения моделей цифровых двойников энергообеспечения тяги поездов, предназначенной для использования в дирекции Трансэнерго, целесообразно производить процедуру верификации закономерностей как на основании физически обоснованного и аналитического анализа результатов моделирования, проведенного как на базе комплекса однофакторных экспресс-моделей, так и на базе многофакторных моделей цифровых двойников при окончательном их построении.

4 Принятие решений о составе параметров моделей цифровых двойников процессов энергообеспечения тяги поездов и их идентификация позволяют производить моделирование динамических реализаций данных процессов после валидации моделей. На исследуемых полигонах территориальных филиалов ОАО «РЖД» достигнута высокая точность моделирования процессов энергообеспечения поездов. Погрешность моделирования не превышает уровня 3,1–3,7 %.

Список литературы

- 1 Распоряжение правительства от 19 марта 2019 года № 466-р «Об утверждении долгосрочной Программы развития ОАО «Российские железные дороги» до 2025 года» // Digital Russia : [сайт]. – URL: <http://d-russia.ru/utverzhdyon-plan-perehoda-rzhd-na-tsifrovuyu-zheleznuyu-dorogu.html> (дата обращения: 20.05.2023).
- 2 ГОСТ Р 57700.37–2021. Компьютерные модели и моделирование. Цифровые двойники изделий. Общие положения. – Москва : Российский институт стандартизации, 2021. – 10 с.

References

- 1 Government Decree dated March 19, 2019 No. 466-r “On Approval of the Long-Term Development Program of JSC Russian Railways until 2025” // Digital Russia : [website]. – URL: <http://d-russia.ru/utverzhdyon-plan-perehoda-rzhd-na-tsifrovuyu-zheleznuyu-dorogu.html> (date of access: 05/20/2023).
- 2 GOST 57700.37–2021. Computer Models and Modeling. Digital twins of products. General Provisions. – Moscow : Russian Institute for Standardization, 2021. – 10 p.

3 **ГОСТР 57188–2016.** Численное моделирование физических процессов. Термины и определения. – Москва : Стандартинформ, 2016. – 8 с.

4 **Вентцель, Е. С.** Теория случайных процессов и её инженерные приложения / Е. С. Вентцель, Л. А. Овчаров. – Москва : Наука, 1991. – 384 с. – ISBN 5-02-014125-9.

5 **Айвазян, С. А.** Прикладная статистика и основы эконометрики / С. А. Айвазян, В. С. Мхитарян. – Москва : Юнити, 1998. – 1022 с.

6 **Эйкхофф, П.** Основы идентификации систем управления. Оценивание параметров и состояния / П. Эйкхофф, В. А. Лотоцкий, А. С. Мандель ; под редакцией Н. С. Райбмана. – Москва : Мир, 1975. – 686 с.

7 Дисперсионная идентификация / Н. С. Райбман, В. В. Капитonenko, Ф. А. Овсепян, П. М. Варлаки ; под редакцией Н. С. Райбмана. – Москва : Наука. Гл. ред. физико-математической литературы, 1981. – 336 с.

8 **Митрофанов, А. Н.** Моделирование процессов прогнозирования и управления электропотреблением тяги поездов : монография / А. Н. Митрофанов. – Самара : СамГАПС, 2005. – 174 с. – ISBN 5-98941-020-4.

9 **Митрофанов, А. Н.** Прогнозирование показателей энергообеспеченности перевозочного процесса на базе многофакторных моделей / А. Н. Митрофанов, Г. М. Третьяков, С. В. Копейкин // Электротехника. – № 3. – Москва : СамГУПС, 2017. – С. 12–17. – EDN XWVTVN.

10 **Жарков, Ю. И.** Автоматизация диагностирования систем релейной защиты и автоматики электроустановок : монография / Ю. И. Жарков, В. Г. Лысенко, Е. А. Стороженко. – Москва : Маршрут, 2005. – 175 с. – ISBN 5-89035-253-9.

11 Комплексное решение прикладных технологических задач автоматизации тягового электрообеспечения / Ю. И. Жарков, О. В. Кубкина, В. Г. Лысенко, Н. А. Попова // Вестник Ростовского государственного университета путей сообщения. – № 1 (81). – 2021. – С. 130–139. – DOI 10.46973/0201-727X_2021_1_130.

12 **Бодров, П. А.** Предиктивный анализ работы технических средств электрифицированных железных дорог в целях повышения их надежности / П. А. Бодров, Ю. И. Жарков, А. Л. Ганашек // Труды Ростовского государственного университета путей сообщения. – № 3 (56). – 2021. – С. 4–6. – EDN PEXXUJ.

13 **Шеер, А.-В.** ARIS – моделирование бизнес-процессов / А.-В. Шеер. – Вильямс, 2000. – 175 с. – ISBN 978-5-8459-1449-1.

3 **GOST 57188–2016.** Numerical modeling of physical processes. Terms and Definitions. – Moscow : Standartinform, 2016. – 8 p.

4 **Ventsel, E. S.** Theory of random processes and its engineering applications / E. S. Ventsel, L. A. Ovcharov. – Moscow : Nauka Publ., 1991. – 384 p. – ISBN 5-02-014125-9.

5 **Ayvazyan, S. A.** Applied Statistics and Fundamentals of Econometrics / S. A. Ayvazyan, V. S. Mkhitaryan. – Moscow : Unity Publ., 1998. – 1022 p.

6 **Eickhoff, P.** Fundamentals of Identification of Control Systems. Estimation of parameters and state / P. Eickhoff, V. A. Lototsky, A. S. Mandel ; edited by N. S. Raibman. – Moscow : Mir, 1975. – 686 p.

7 Dispersion identification / N. S. Raibman, V. V. Kapitonenko, F. A. Ovsepyan, P. M. Varlaki ; edited by N. S. Raibman. – Moscow : Science. Ch. ed. Physics and Mathematics Literature, 1981. – 336 p.

8 **Mitrofanov, A. N.** Modeling of processes of forecasting and control of electric power consumption of train traction : monograph / A. N. Mitrofanov. – Samara : SamGAPS, 2005. – 174 p. – ISBN 5-98941-020-4.

9 **Mitrofanov, A. N.** Forecasting indicators of the energy supply of the transportation process on the basis of multifactorial models / A. N. Mitrofanov, G. M. Tretyakov, S. V. Kopeykin // Electrical engineering. – No. 3. – Moscow : SamGUPS, 2017. – P. 12–17. – EDN XWVTVN.

10 **Zharkov, Yu. I.** Automation of diagnosing systems of relay protection and automation of electrical installations: monograph / Yu. I. Zharkov, V. G. Lysenko, E. A. Storozhenko. – Moscow : Route, 2005. – 175 p. – ISBN 5-89035-253-9.

11 Complex solution of applied technological problems of automation of traction power supply / Yu. I. Zharkov, O. V. Kubkina, V. G. Lysenko, N. A. Popova // Vestnik Rostovskogo Gosudarstvennogo Universiteta Putey Soobshcheniya. – No. 1 (81). – 2021. – P. 130–139. – DOI 10.46973/0201-727X_2021_1_130.

12 **Bodrov, P. A.** Predictive analysis of the work of technical means of electrified railways in order to improve their reliability / P. A. Bodrov, Yu. I. Zharkov, A. L. Ganashek // Trudy Rostovskogo Gosudarstvennogo Universiteta Putey Soobshcheniya. – No. 3 (56). – 2021. – P. 4–6. – EDN PEXXUJ.

13 **Scheer, A.-W.** ARIS – Business Process Modeling / A.-W. Scheer. – Williams, 2000. – 175 p. – ISBN 978-5-8459-1449-1.

14 **Никонов, В. В.** КОМПАС-3D : создание моделей и 3D-печать / В. В. Никонов. – Санкт-Петербург : Питер, 2020. – 208 с. – ISBN 978-5-4461-1456-6.

15 **Парментер, Д.** Ключевые показатели эффективности : разработка, внедрение и использование успешных KPI / Д. Парментер. – Нью-Джерси, США : John Wiley & Sons, inc., 2007. – 233 с. – ISBN 0-470-09588-1.

16 **Митрофанов, А. Н.** Подход к построению моделей цифровых двойников процессов энергообеспечения тяги поездов / А. Н. Митрофанов, Ю. И. Жарков, С. А. Окладов // Вестник Ростовского государственного университета путей сообщения. – 2023. – № 3. – С. 92–102. – DOI 10.46973/0201-727X_2023_3_92.

17 Методологическое обоснование построения цифровых двойников процессов энергообеспечения тяги поездов / А. Н. Митрофанов, С. А. Окладов // Вестник транспорта Поволжья. – 2023. – № 3 (99). – С. 24–32. – EDN HFMEMR.

18 **Окладов, С. А.** Формирование моделей и баз данных бизнес-процессов энергообеспечения для дирекции Трансэнерго ОАО «РЖД» / С. А. Окладов // Наука и образование транспорту. – 2023. – № 1. – С. 286–290. – EDN CRWITS.

19 **Дейт, К. Дж.** Введение в системы баз данных / К. Дж. Дейт. – 8-е изд. – Москва : Вильямс, 2005. – 1328 с. – ISBN 5-8459-0788-8.

20 **Майер-Шенбергер, В.** Большие данные. Революция, которая изменит то, как мы живём, работаем и мыслим / В. Майер-Шенбергер, К. Кукиер ; пер. с англ. Инны Гайдюк. – Москва : Манн, Иванов, Фербер, 2014. – 240 с. – ISBN 987-5-91657-936-9.

21 **Жамбю, М.** Иерархический кластер-анализ и соответствия / М. Жамбю. – Москва : Финансы и статистика, 1988. – 345 с.

22 **Григорьев, Ю. Д.** Методы оптимального планирования эксперимента: линейные модели : учебное пособие / Ю. Д. Григорьев. – Санкт-Петербург : Лань, 2015. – 320 с.

23 **Брюс, П.** Разведочный анализ данных / П. Брюс, Э. Брюс // Практическая статистика для специалистов Data Science. – Санкт-Петербург : БХВ-Петербург, 2018. – С. 19–58. – 304 с.

24 **Миронов, А. М.** Методы верификации программ / А. М. Миронов. – Москва : ДМК Пресс, 2023. – 332 с.

25 **Дрейпер, Н.** Прикладной регрессионный анализ. Множественная регрессия / Н. Дрейпер, Г. Смит. – 3-е изд. – Москва : Диалектика, 2007. – 912 с. – ISBN 0-471-17082-8.

14 **Nikonov, V. V.** COMPASS-3D : Model creation and 3D printing // Saint Petersburg : Piter, 2020. – 208 p. – ISBN 978-5-4461-1456-6.

15 **Parmenter, D.** Key Performance Indicators : Developing, Implementing and Using Winning KPI's. – New Jersey, USA : John Wiley & Sons, inc., 2007. – 233 p. – ISBN 0-470-09588-1.

16 **Mitrofanov, A. N.** Approach to building digital twin models of business processes of energy supply for train traction / A. N. Mitrofanov, Yu. I. Zharkov, S. A. Okladov // Vestnik Rostovskogo Gosudarstvennogo Universiteta Putey Soobshcheniya. – 2023. – No. 3. – P. 92–102. – DOI 10.46973/0201-727X_2023_3_92.

17 Methodological justifications for the construction of digital twins of business processes for energy supply of traction of trains / A. N. Mitrofanov, S. A. Okladov // Bulletin of Transport of the Volga Region. – 2023. – No. 3 (99). – P. 24–32. – EDN HFMEMR.

18 **Okladov, S. A.** Formation of models and databases of business processes of energy supply for the Directorate of Transenergo of JSC “Russian Railways” / S. A. Okladov // Science and Education for Transport. – 2023. – No. 1. – P. 286–290. – EDN CRWITS.

19 **Date, C. J.** Introduction to Database Systems / C. J. Date. – 8th ed. – Moscow : Williams Publ., 2005. – 1328 p. – ISBN 5-8459-0788-8.

20 **Mayer-Schönberger, V.** Big Data. A Revolution That Will Transform How We Live, Work, and Think / V. Mayer-Schönberger, K. Cuquier. – Moscow : Mann, Ivanov, Ferber, 2014. – 240 p. – ISBN 987-5-91657-936-9.

21 **Zhambyu, M.** Hierarchical Cluster-Analysis and Correspondences / M. Zhambyu. – Moscow : Finance and Statistics Publ., 1988. – 345 p.

22 **Grigor'ev, Y. D.** Methods of optimal experimental planning: linear models : textbook / Y. D. Grigor'ev. – Saint Petersburg : Lan', 2015. – 320 p.

23 **Bruce, P.** Exploratory Data Analysis / P. Bruce, E. Bruce // Practical Statistics for Data Science Specialists. – Saint Petersburg : BHV-Petersburg Publ., 2018. – P. 19–58. – 304 p.

24 **Mironov, A. M.** Methods of program verification / A. M. Mironov. – Moscow : DMK Press, 2023. – 332 p.

25 **Draper, N.** Applied regression analysis / N. Draper, G. Smith. – 3rd ed. – Moscow : Dialectics Publ., 2007. – 912 p. – ISBN 0-471-17082-8.

26 **Боровиков, В. П.** Statistica : искусство анализа данных на компьютере / В. П. Боровиков. – 2-е изд. – Санкт-Петербург : Питер, 2003. – 688 с. – ISBN 5-272-00078-1.

27 **Карлберг, К.** Регрессионный анализ в Microsoft Excel / К. Карлберг. – Москва : Вильямс, 2017. – 400 с. – ISBN 978-5-9908462-7-2.

28 **Арло, С.** Обзор процедур перекрестной проверки для выбора модели // С. Арло, А. Селисс // Статист. – 2010. – Вып. 4. – С. 40–79. – DOI 10.1214/09-SS054.

29 Верификация и валидация компьютерных моделей / А. В. Сальников, М. С. Французов, К. А. Виноградов [и др.] // Известия высших учебных заведений. Машиностроение. – 2022. – № 9. – С. 100–115. – DOI 10.18698/0536-1044-2022-9-100-115.

30 **Митрофанов, А. Н.** Построение моделей цифровых двойников электротехнической инфраструктуры железных дорог для оценки ресурса её нагрузочной способности / А. Н. Митрофанов, М. А. Гаранин, С. А. Окладов // Электротехника. – 2023. – № 10. – С. 2–7. – DOI 10.53891/00135860_2023_10_2.

26 **Borovikov, V. P.** Statistica : The Art of Data Analysis on a Computer / V. P. Borovikov. – 2nd ed. – Saint Petersburg : Piter, 2003. – 688 p. – ISBN 5-272-00078-1.

27 **Karlberg, K.** Regression Analysis in Microsoft Excel / K. Karlberg. – Moscow : Williams, 2017. – 400 p. – ISBN 978-5-9908462-7-2.

28 **Arlot, S.** A survey of cross-validation procedures for model selection / S. Arlot, A. Celisse / Statistics Surveys. – 2010. – Iss. 4. – DOI 10.1214/09-SS054.

29 Verification and validation of computer models / A. V. Salnikov, M. S. Frantsuzov, K. A. Vinogradov [et al.] // BMSTU Journal of Mechanical Engineering. – 2022. – No. 9. – P. 100–115. – DOI 10.18698/0536-1044-2022-9-100-115.

30 **Mitrofanov, A. N.** Construction of models of digital twins of the elektrotechnical infrastructure of railways for assessing the resource of its loading capacity / A. N. Mitrofanov, M. A. Garanin, S. A. Okladov // Elektrotehnika. – 2023. – No. 10. – P. 2–7. – DOI 10.53891/00135860_2023_10_2.

A. N. Mitrofanov, Yu. I. Zharkov, S. A. Okladov

METHODOLOGY FOR DESIGNING MODELS OF DIGITAL TWINS OF TRAIN TRACTION POWER SUPPLY PROCESSES

Abstract. The paper presents a methodology for designing models of digital twins of train traction energy supply processes, which provides the correct selection of factors and identification of parameters of a multifactor model for the territorial branches of JSC Russian Railways.

The following processes of power supply were considered: electricity consumption for traction, specific electricity consumption, imbalance of traction power supply. As independent variables of the model, the processes of freight transportation and the processes to ensure them were studied: cargo loading, cargo turnover, train speed, indicators of the use of electric locomotives, wagons, etc. According to the results of the validation procedure, it was established that it was possible to simulate the dynamic implementation of the processes of power supply of train traction with an error of no more than 3.1–3.7% of the actual values.

Keywords: digital twins, express models of processes, energy supply, train traction, correlation, cluster analysis, verification of a multivariate model of a digital twin.

For citation: Mitrofanov, A. N. Methodology for designing models of digital twins of train traction power supply processes / A. N. Mitrofanov, Yu. I. Zharkov, S. A. Okladov // Vestnik Rostovskogo Gosudarstvennogo Universiteta Putey Soobshcheniya. – 2024. – No. 2. – P. 82–99. – DOI 10.46973/0201-727X_2024_2_82.

Сведения об авторах

Митрофанов Александр Николаевич

Самарский государственный университет путей сообщения (СамГУПС),
кафедра «Электроснабжение железнодорожного транспорта»,
доктор технических наук, профессор,
e-mail: almit77@mail.ru

Information about the authors

Mitrofanov Alexander Nikolayevich

Samara State Transport University (SSTU),
Chair «Power Supply of Railway Transport»,
Doctor of Engineering Sciences, Professor,
e-mail: almit77@mail.ru

Жарков Юрий Иванович

Ростовский государственный университет путей
сообщения (РГУПС),
кафедра «Автоматизированные системы
электрообеспечения»,
доктор технических наук, профессор,
e-mail: zharkov43@gmail.com

Окладов Сергей Анатольевич

Самарский государственный университет путей
сообщения (СамГУПС),
кафедра «Электрообеспечение железнодорожного
транспорта»,
старший преподаватель,
e-mail: okladov@samgups.ru

Zharkov Yury Ivanovich

Rostov State Transport University (RSTU),
Chair «Automated Power Supply Systems»,
Doctor of Engineering Sciences, Professor,
e-mail: zharkov43@gmail.com

Okladov Sergey Anatolievich

Samara State Transport University (SSTU),
Chair «Power Supply of Railway Transport»,
Senior Lecturer,
e-mail: okladov@samgups.ru