

Прогноз скорости транспортного потока (на примере г. Хабаровска)



И. Н. Пугачев,
д-р техн. наук, профессор,
заместитель директора
по научной работе
Хабаровского Федераль-
ного исследовательского
центра Дальневосточного
отделения РАН, профессор
кафедры изысканий и про-
ектирования железных
и автомобильных дорог
Дальневосточного государ-
ственного университета
путей сообщения,



Н. Г. Шешера,
д-р техн. наук,
доцент кафедры информа-
ционного и технического
обеспечения ОВД
Дальневосточного
юридического института
имени И. Ф. Шиловой

Современная транспортная инфраструктура представляет собой сложную социально-техническую систему, в которой проектные решения должны не только обеспечивать высокую пропускную способность и экономическую эффективность, но и гарантировать безопасность всех участников дорожного движения.

В условиях роста интенсивности транспортных потоков, увеличения скоростных характеристик транспортных средств и усложнения городской застройки особое значение приобретают фундаментальные параметры проектирования, определяющие как геометрию дороги, так и условия ее эксплуатации. Среди них ключевую роль играет расчетная скорость — величина, лежащая в основе формирования геометрии плана и профиля автомобильных дорог, а также служащая критерием оценки соответствия существующих участков современным требованиям безопасности и пропускной способности [1].

Расчетная скорость — это величина, принимаемая при проектировании автомобильной дороги и определяющая минимально допустимые значения геометрических элементов, таких как радиус кривой в плане и продольном профиле, ширина проезжей части, расстояния видимости, параметры искусственных сооружений и обустройства. Она не всегда совпадает с фактической скоростью транспортного потока после ввода дороги в эксплуатацию.

Важно подчеркнуть, что расчетная скорость — это не просто технический параметр, а компромиссное решение, отражающее совокупность инженерных, экономических, экологических, социальных и правовых факторов, характерных для конкретной территории и категории автомобильной дороги, поэтому при проектировании транспортных сооружений важно добиться высокой точности данного параметра [2]. В Российской Федерации расчетная скорость назначается проектировщиком на основе категории дороги, функционального класса и условий рельефа в соответствии с таблицами из свода правил СП 34.13330.2021¹, являющихся следствием научных изысканий, резуль-

таты которых, как и сам расчет скорости, приведен в ОДМ 218.2.020–2012².

За рубежом подходы к определению и применению расчетной скорости демонстрируют как существенные сходства, так и принципиальные различия. В странах Европейского союза, например, в руководствах PIARC (Международной ассоциации автомобильных дорог и транспорта) и национальных стандартах (как немецкий RAS-L или французский Guide Technique), расчетная скорость остается основой проектирования не как единственная величина для всей дороги, а как диапазон проектных скоростей, зависящий от типа элемента. Например, для кривой малого радиуса может быть назначена одна расчетная скорость, а для прямого участка — другая. Кроме того, в ЕС широкое распространение получила концепция «согласованной скорости» (target speed или desired speed), которую, по замыслу проектировщика, должен фактически выбирать водитель в данных условиях [3, 4]. Эта величина увязывается с геометрией, типом обустройства (наличие разметки, островков безопасности, элементов «успокоения» трафика) и даже с ландшафтным дизайном. Такой подход позволяет управлять поведением водителя «пассивно» — через среду, а не только через знаки и штрафы.

В США расчетная скорость регулируется AASHTO Green Book — ключевым документом Американской ассоциации государственных служащих автомобильных дорог. Она определяется как «скорость, выбранная для проектирования геометрических элементов дороги, при которой водитель может безопасно управлять транспортным средством при благоприятных погодных условиях и при отсутствии помех со стороны других участников движения».

¹ СП 34.13330.2021 «Свод правил. Автомобильные дороги».

² ОДМ 218.2.020–2012 «Отраслевой дорожный документ. Методические рекомендации по оценке пропускной способности автомобильных дорог».

Важной особенностью американской практики является прямая связь проектной скорости с 85-м перцентилем фактической скорости движения, т. е. проектная скорость часто выбирается близкой к скорости, которую не превышают 85% водителей. Это снижает вероятность конфликтов между «быстрыми» и «медленными» участниками и повышает однородность потока. Однако такой подход вызывает дискуссии, поскольку может привести к «уравниванию вверх» — нормализации более высоких, но потенциально опасных скоростей [5].

Подготовка данных

Скоростной режим автомобилей определяется не только геометрией дороги (радиусы кривых, ширина проезжей части и др.) и временными факторами (часы суток, сезон), но и условиями внешней среды. Как показывают исследования, даже при благоприятных геометрических характеристиках участка и отсутствии формальных препятствий совокупное влияние компонентов системы ВАДС¹ может создавать скрытые риски, приводя к росту аварийности. В таких случаях водители, интуитивно оценивая угрозу личной безопасности, сознательно снижают скорость — часто ниже нормативной или ожидаемой по проекту.

С целью количественной оценки этих эффектов в течение двух лет на улично-дорожной сети г. Хабаровска проводился сбор эмпирических данных, охватывающих динамику скорости транспортного потока, геометрические параметры проезжей части (число полос, ширина, наличие выделенных полос для маневров), временные режимы (время суток, день недели, месяц), метеорологические условия (температура, влажность, ветер, осадки и др.). Замеры скорости выполнялись с помощью GPS/ГЛОНАСС-трекера, установленного на контрольном автомобиле, который двигался в составе реального транспортного потока без искусственного ускорения или торможения.

Целевой переменной в анализе выступала максимальная зафиксированная скорость (*max_speed*) на данном участке в конкретный момент — показатель, отражающий не предельную техническую возможность транспортных средств, а максимум, реализуемый водителями в сложившихся условиях. Данные о по-

¹ ВАДС (водитель, автомобиль, дорога, среда) — сложная целостная динамическая система, где каждый из элементов находится в определенных отношениях и связях друг с другом.

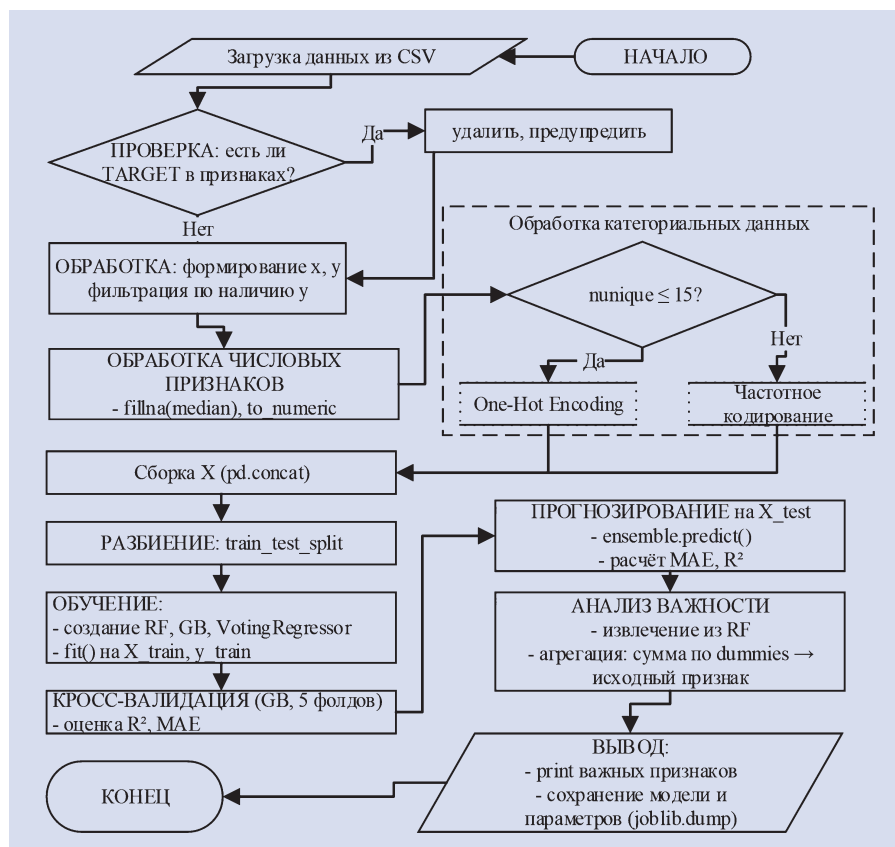


Рис. 1. Алгоритм программы обучения и прогноза

годе привязывались по времени замера с использованием открытых геоинформационных сервисов (с разрешением 1 ч).

Построение программы обучения и прогноза

После создания банка статистических данных с различных информационных ресурсов разработана программа, которая включает функции предварительной обработки данных, построения и оценки ансамблевой регрессионной модели для прогнозирования максимальной скорости транспортного средства (*max_speed*) на основе множества числовых и категориальных признаков, связанных с дорожно-транспортными происшествиями (рис. 1) [7].

Программа реализована на языке Python с использованием библиотек *pandas*, *numpy*, *scikit-learn* и *joblib* [8]. В качестве источника данных выступает единый CSV-файл, содержащий структурированную информацию о характеристиках, описывающих скорость транспортного потока. На этапе загрузки проводится валидация структуры, проверяется наличие целевой переменной и исключается ее возможное дублирование в списке предикторов, что предотвращает утечку информации [9].

Пропущенные значения в числовых признаках заменяются медианой, что

обеспечивает устойчивость к выбросам. Для категориальных признаков применяется комбинированный подход, признаки с числом уникальных значений до 15 преобразуются методом *one-hot encoding*, а признаки высокой кардинальности кодируются на основе частотного распределения их значений в обучающей выборке. Такой подход позволяет сохранить информативность без избыточного роста размерности [10].

Обучение модели выполняется на основе ансамбля двух регрессоров — случайного леса и градиентного бустинга, объединенных с помощью усреднения предсказаний (*voting regressor*) [11]. Такая композиция повышает устойчивость и обобщающую способность модели по сравнению с использованием одного алгоритма. Для оценки качества используются метрики средней абсолютной ошибки (*MAE*) и коэффициента детерминации (*R²*), а также кросс-валидация (пять фолдов), что позволяет объективно оценить стабильность модели на разных подвыборках данных.

Важность признаков агрегируется на уровне исходных переменных (с суммированием вкладов всех *dummys* переменных для категориальных признаков), что позволяет выявить ключевые факторы, оказывающие наибольшее влияние на скорость.

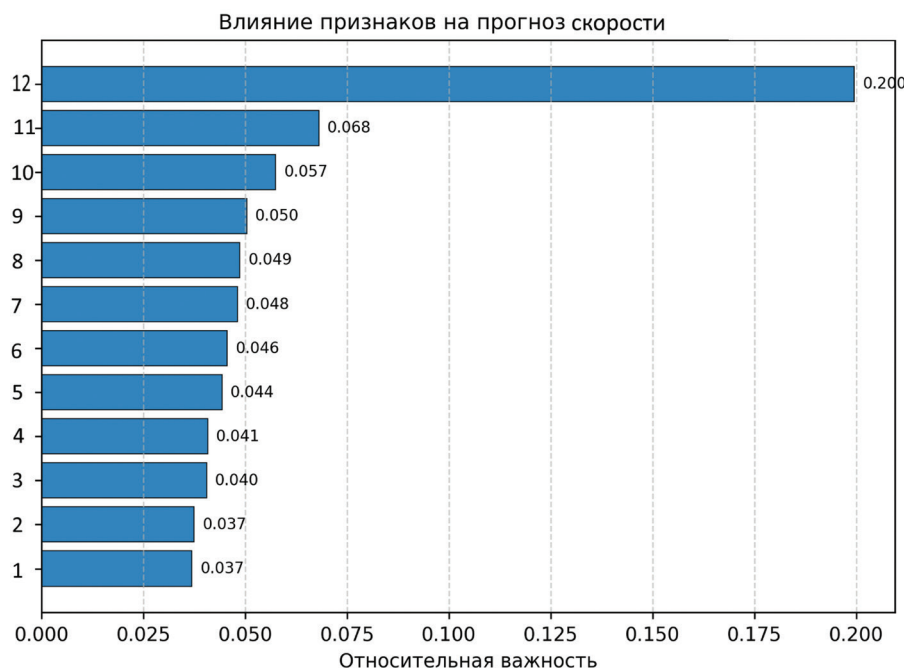


Рис. 2. Определение важности признаков методом перестановки

1 — ширина дороги; 2 — месяц; 3 — направление ветра; 4 — дефицит насыщения воздуха;
 5 — температура почвы; 6 — температура воздуха; 7 — вид парковки; 8 — относительная влажность воздуха;
 9 — день недели; 10 — время; 11 — атмосферное давление; 12 — вид полосы для левого поворота

В завершение модель и все параметры предобработки (медианы, частотные карты, список ожидаемых столбцов) сохраняются в сериализованном виде, что обеспечивает воспроизводимость и возможность последующего применения в автоматизированных системах прогнозирования без повторного обучения.

Для обеспечения надежности последующего моделирования и интерпретируемости результатов разработан и внедрен специализированный диагностический модуль `validators.py`, выполняющий комплексную верификацию входного массива данных и соответствия его структуры принятой методологической схеме. Модуль реализует многоэтапную процедуру контроля, охватывающую проверку корректности загрузки, отсутствие утечки целевой переменной, оценку информативности признаков и визуализацию важности на основе обученной модели.

После фильтрации проводится статистический анализ числовых признаков, сопровождаемый оценкой доли пропущенных значений и их последующей импутацией на этапе основной обработки (в другом модуле). Особое внимание уделяется корреляционной структуре данных: рассчитываются коэффициенты Пирсона и Спирмена между всеми числовыми признаками и целевой переменной скорости (`max_speed`).

Такой двойной подход позволяет

выявить как линейные, так и монотонные, но нелинейные зависимости, что особенно важно при работе с транспортными характеристиками, часто имеющими пороговые или насыщающиеся эффекты (например, влияние радиуса кривизны на допустимую скорость). Признаки с низкой абсолютной корреляцией ($|r| < 0,15$) фиксируются для последующего анализа — их включение в модель может увеличивать дисперсию оценок без значимого прироста качества.

Для категориальных признаков выполняется верификация полноты охвата, проверяется наличие каждого из заявленных признаков в исходной таблице, а также оценивается разнообразие значений (число уникальных категорий) и доминирование отдельных уровней (частота модального значения). Это позволяет своевременно выявить признаки с низкой вариативностью (например, «тип дороги = шоссе» в 98 % случаев), чья информативность может быть ограничена.

Завершающий этап модуля `validators.py` — верификация обученной модели с точки зрения интерпретируемости. При наличии сериализованной модели (`model_max_speed.pkl`) и параметров предобработки (`preprocessing_params.pkl`) модуль автоматически извлекает вектор важности признаков, поддерживаемый ансамблевыми методами на основе деревьев решений.

Поскольку в предобработке применяется частотное кодирование категориальных переменных, важность агрегируется по исходным признакам, для числовых — напрямую, для категориальных — по сумме вкладов всех порожденных бинарных или частотных компонент.

Определение значимых признаков

На основе агрегированного вектора строится горизонтальная столбчатая диаграмма наиболее влиятельных факторов (рис. 2). Такая визуализация обеспечивает прозрачность модели и служит основой для содержательного анализа, например, для выявления доминирующей роли геометрических параметров трассы или характеристик транспортного средства, что имеет прямое прикладное значение для совершенствования норм проектирования.

С учетом экспериментального исключения значений определено 12 переменных, которые в совокупности лучше всего описывают изменения скоростного режима. Точность модели составила: $R^2 = 0,774$, MAE (ошибка скорости) = 3,73 км/ч.

Проверка модели на тестовых данных

Для проверки разработанной модели выполнен тестовый прогноз на 30 данных, которые не участвовали в обучении (см. таблицу).

С учетом полученных данных можно сделать вывод о том, что расхождения фактических значений с прогнозом не являются существенными для проектирования транспортных сооружений. В редких случаях они превышают 10 км/ч.

Заключение

Проведенное исследование продемонстрировало зависимость изменения скорости от совокупности таких характеристик, как погодные условия, геометрические элементы дорог и временные параметры (месяц, день недели, время суток). Но это не значит, что перечень независимых признаков исчерпывающий. Дальнейшие исследования будут выявлять новые закономерности и зависимости.

Предложенный алгоритм является готовым инструментом для исследования и по мере дальнейшего накопления статистических данных точность модели будет только увеличиваться. На данный

Результаты тестового прогноза

Фактические данные, км/ч	Прогноз, км/ч	Ошибка, км/ч
43,692	44,947	1,255
53,961	53,008	0,954
53,125	44,794	8,331
42,731	43,890	1,159
38,166	39,328	1,162
42,226	42,776	0,551
37,726	38,172	0,446
28,017	31,090	3,073
48,709	45,437	3,272
41,060	42,059	0,999
37,044	35,603	1,441
21,516	36,099	7,582
42,671	43,421	0,750
53,431	37,303	9,129
44,467	44,174	0,293
46,835	48,485	1,650
53,724	53,172	0,552
35,189	35,445	0,256
37,395	49,678	9,282
39,295	39,237	0,058
34,198	34,315	0,117
59,932	59,385	0,547
33,367	34,172	0,805
33,826	33,903	0,077
19,557	22,550	2,994
50,305	49,539	0,766
45,231	45,127	0,104
21,693	25,069	3,376
58,417	57,877	0,540
65,431	64,723	0,708

момент рекомендуется скорректировать СП 34.13330.2021 в части назначения скоростей с учетом предложенных в работе решений.

Литература

1. Пугачев И. Н., Скрипко П. Б., Шешера Н. Г. Программный подход к комплексному сбору и подготовки данных об интенсивности движения транспортных средств, погодных условий и есте-

ственной освещенности в часовых интервалах // Т-Comm: Телекоммуникации и транспорт. 2023. Т. 17. № 10. С. 43–51.

2. Пугачев И. Н., Шешера Н. Г., Григоров Д. Е. Использование инструментов современных библиотек Python по работе с географическими координатами для решения задач безопасности дорожного движения // Транспорт Азиатско-Тихоокеанского региона. 2024. № 3(40). С. 60–66.

3. Mahmoud N., Abdel-Aty M., Cai Q., Abuzwidah M. Analyzing the difference between operating speed and target speed using mixed-effect ordered logit model // Transportation Research Record. 2022. Vol. 2676. № 9. P. 596–607.

4. Sutheerakul Ch., Kronprasert N., Satiennam W., Zaw M. S. C. Classification of roadway context and target speed for multilane highways in thailand using fuzzy expert system // Sustainability. 2024. Vol. 16. № 9. P. 3865.

5. Chen L., Zhou Q., Cai Y. et al. Cae-gan: a hybrid model for vehicle trajectory prediction // IET Intelligent Transport Systems. 2022. Vol. 16. № 12. P. 1682–1696.

6. Пугачев И. Н., Шешера Н. Г., Григоров Д. Е. Определение эффективных широт интервалов влияния температурных режимов на интенсивность транспортного потока с использованием ранговой корреляции Спирмена // Т-Comm: Телекоммуникации и транспорт. 2024. Т. 18. № 3. С. 34–40.

7. Фурман С. И., Чантиева М. Э. Ансамблевые методы, алгоритмы семантической обработки видеоизображений, архитектуры машинного обучения для режима реального времени // Инновации. Наука. Образование. 2020. № 22. С. 389–393.

8. Зацерковный А. В., Нурминский Е. А. Нейросетевой анализ транспортных потоков городских агломераций на основе данных публичных камер видеонаблюдения // Компьютерные исследования и моделирование. 2021. Т. 13. № 2. С. 305–318.

9. Mazloumi E., Rose G., Currie G., Moridpour S. Prediction intervals to account for uncertainties in neural network predictions: methodology and application in bus travel time prediction // Eng. Appl. Artif. Intell. 2011. № 24(3). P. 534–542.

10. Донченко Д. С., Садовникова Н. П., Парыгин Д. С. Прогнозирование степени тяжести последствий ДТП с использованием методов машинного обучения // Вестник Воронежского института высоких технологий. 2019. № 13(4).

11. Habib M., Bashir B., Als Salman A., Bachir H. Evaluating the accuracy and effectiveness of machine learning methods for rapidly determining the safety factor of road embankments // Multidiscipline Modeling in Materials and Structures. 2023. Vol. 19. № 5. P. 966–983.