

ПРОГНОЗИРОВАНИЕ СТАТИСТИЧЕСКИХ ДАННЫХ АВТОМОБИЛЬНЫХ ДОРОЖНО-ТРАНСПОРТНЫХ ПРОИСШЕСТВИЙ

Ларионов К.О.

Email: Larionov17163@scientifictext.ru

*Ларионов Константин Олегович – аспирант,
кафедра вычислительной техники и защиты информации,
Оренбургский государственный университет, г. Оренбург*

Аннотация: в статье анализируется статистика автомобильных дорожно-транспортных происшествий. Актуальность разработки и внедрения метода прогнозирования статистики автомобильных дорожно-транспортных происшествий обуславливается тем, что применяемый метод прогнозирования позволяет достаточно точно определить прогнозное значение с максимальной погрешностью меньше 15%. Автоматизированные методы прогнозирования на сегодняшний день помогают быстрому принятию решений в конкретной предметной области за счет быстрой обработки большого количества статистических данных.

Ключевые слова: защита, система, информация, программное обеспечение, дорожно-транспортные происшествия, прогноз, статистика, автомобиль.

FORECASTING STATISTICAL DATA OF ROAD TRANSPORTATION ACCIDENTS

Larionov K.O.

*Larionov Konstantin Olegovich – Postgraduate,
DEPARTMENT OF COMPUTING AND INFORMATION SECURITY,
ORENBURG STATE UNIVERSITY, ORENBURG*

Abstract: The article analyzes the statistics of road traffic accidents. The relevance of the development and implementation of a method for predicting statistics of road traffic accidents is due to the fact that the applied forecasting method allows you to accurately determine the predicted value with a maximum error of less than 15%. Automated forecasting methods today help rapid decision-making in a specific subject area due to the rapid processing of a large amount of statistical data.

Keywords: protection, system, information, software, road traffic accidents, forecast, statistics, car.

УДК 004.421.2

В современном мире все очень быстро меняется и движется. С расширением городов растет и расстояние, которое нужно преодолевать человеку каждый его рабочий день. Таким образом, растет и транспортный поток, люди все реже используют городской транспорт, покупая свои автомобили, мотоциклы из-за их доступности на рынке. Вследствие чего растет количество пробок, аварий, и дорожно-транспортных происшествий с летальным исходом. Автомобильный трафик сегодня в странах порождает одни из самых актуальных мировых проблем, растущее количество пробок на дорогах и дорожно-транспортных происшествий. В данной статье рассмотрена проблема дорожно-транспортных происшествий на территории Российской Федерации. Ниже приведена мировая статистика автомобильных дорожно-транспортных происшествий за 2019 год.



Рис. 1. Статистические данные автомобильные ДТП по странам за 2019 год

Из графика видно, что Россия не занимает лидирующего положения на мировой арене по статистике ДТП на 100 тыс. Однако тенденция роста сохраняется и поэтому вопрос автомобильных ДТП один из важных на сегодняшний день на территории России.

В работе Луценко Е.В. "Адаптивная семантическая информационная модель прогнозирования рисков совершения ДТП" рассматривается подход системно-когнитивного анализа [7]. Суть данного анализа в упрощении модели прогнозирования статистических данных ДТП, а также нахождение зависимостей путем перебора возможных вариантов. В работе не рассматривается вопрос сезонности данных, а также статистических анализ данных ДТП.

В работе Александрова Н.Н. "Прогнозирование количества транспортных средств, вовлеченных в ДТП на двухполосных внегородских автомобильных дорогах" рассматривается подход прогнозирования на основе принципов логики вероятностей на узко-специализированной тематике не рассматривая общее положение дел в статистике мирового ДТП на всех дорогах [6].

В работе Скоробогатченко Д.А. "Нечеткая нейросетевая модель для прогнозирования числа ДТП региона в условиях ограниченной информации" рассматривается современный подход к анализу и прогнозированию данных на базе нейронных сетей и нечеткой логики [8]. Однако нечеткая логика носит характер принятия решения отношения объекта ко множеству на базе функций принадлежности, которые не отражают сезонной компоненты и базовой корреляции временного ряда, а используют подстройку весов обученной нейронной сети.

Целью исследования является снижение риска автомобильных дорожно-транспортных происшествий на основе современных методов анализа и прогнозирования временных рядов.

Для достижения цели были поставлены следующие задачи:

1 Провести сбор статистических данных по автомобильным дорожно-транспортным происшествиям.

Актуальные статистические данные по дорожно-транспортным происшествиям на территории Российской Федерации предоставляет открытый ресурс Госавтоинспекции, режим доступа URL: <http://stat.gibdd.ru/>.

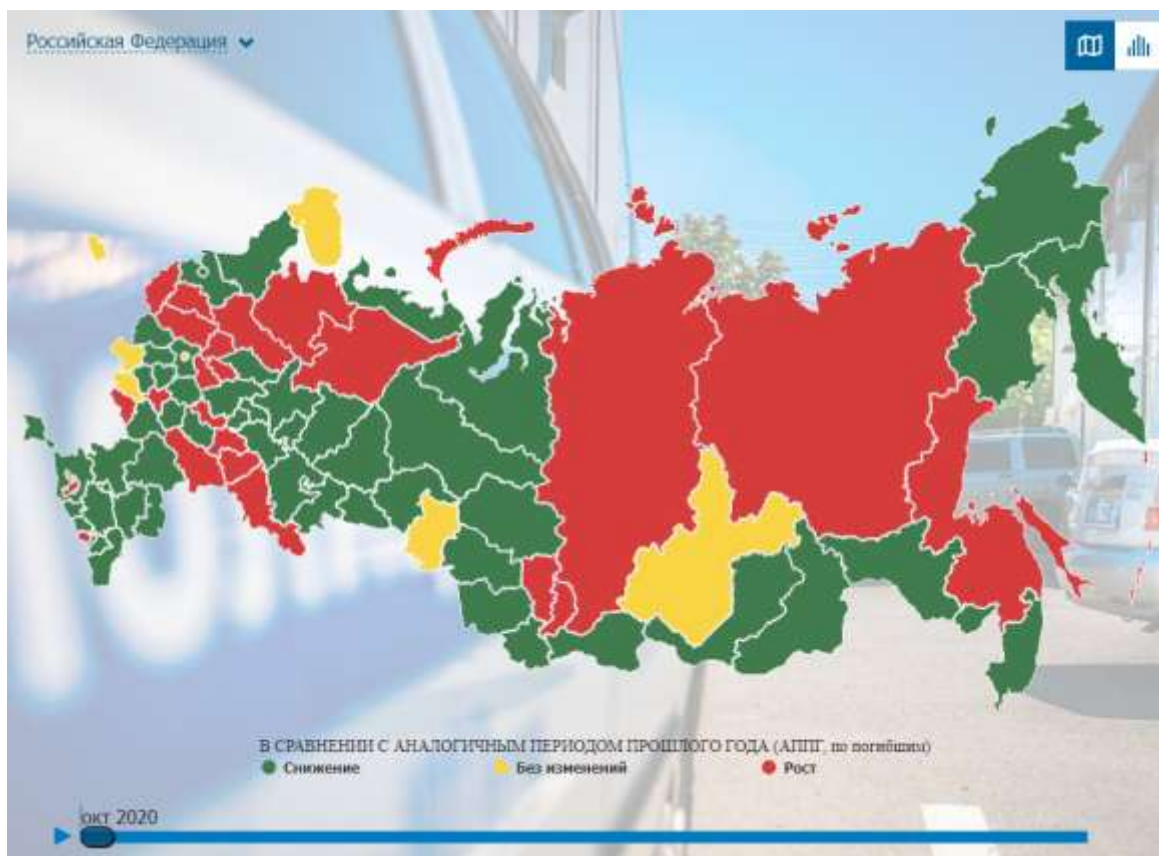


Рис. 2. Сравнительные показатели ДТП с аналогичным периодом прошлого года

На данном ресурсе можно запросить статистику по всей территории России или по конкретной Республике, фильтровать данные по месяцам и по категориям.

2 Провести анализ и прогнозирование статистических данных по автомобильным дорожно-транспортным происшествиям.

На основе созданной выборке данных дорожно-транспортных происшествий с октября 2019 года по октябрь 2020 год был сформирован и проанализирован временной ряд.

На рисунке 3 представлен анализ данных с выделением уровней поддержки и сопротивления графика статистических данных, а также все его статистические показатели, учитывая графики амплитудного и фазового спектра.

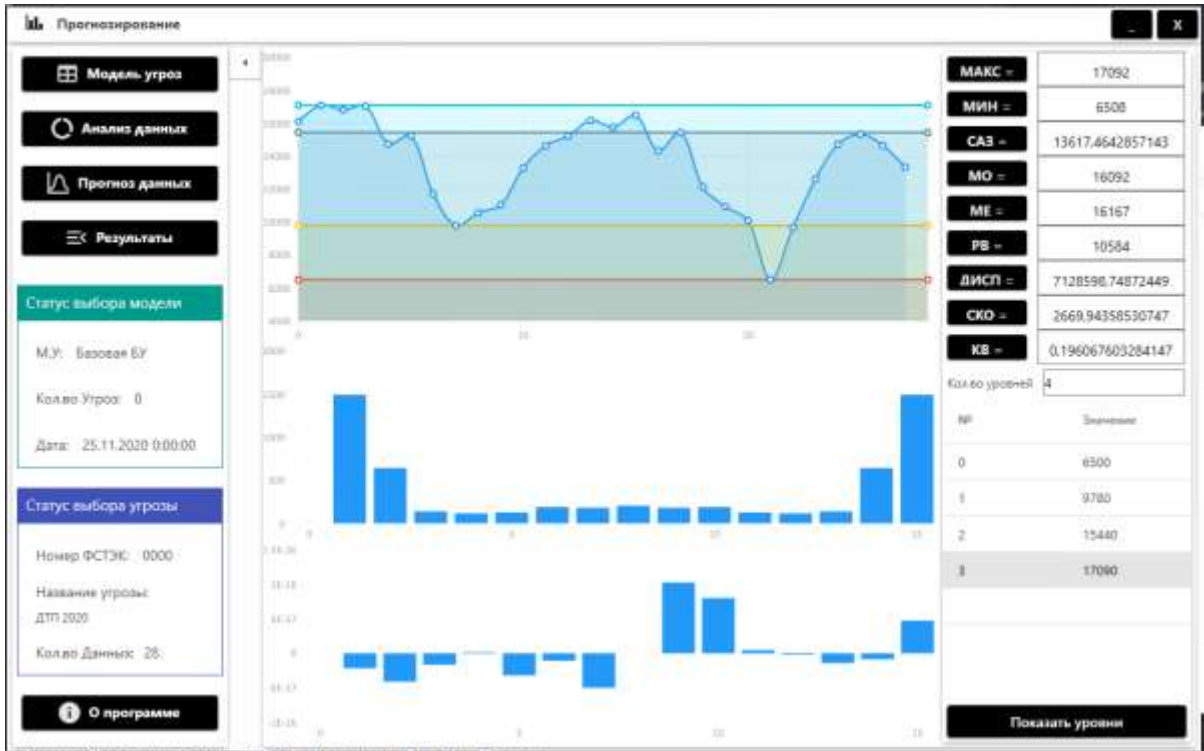


Рис. 3. Анализ статистических данных дорожно-транспортных происшествий с октября 2019г. по октябрь 2020 г.

Проведем разложение временного ряда на его базовые составляющие: тренд, сезонная компонента, случайная компонента. Математическая модель полигармонического полинома для разложения на базовые характеристики, прогнозирования и анализа данных представлена на формулах:

$$a_0 = \frac{\sum_{t=1}^N y_i}{N}, \quad (1)$$

где:

N – число элементов исходного ряда;

y – значение основного исходного ряда;

t – порядковый номер элементов исходного ряда, $t = 1, 2, \dots$.

$$a_k = \frac{2 \cdot \sum_{t=1}^N y_i \cdot \cos\left(2 \cdot \pi \cdot k \cdot \frac{t}{N}\right)}{N}, \quad (2)$$

$$b_k = \frac{2 \cdot \sum_{t=1}^N y_i \cdot \sin\left(2 \cdot \pi \cdot k \cdot \frac{t}{N}\right)}{N}, \quad (3)$$

где:

N – число элементов исходного ряда;

y – значение основного исходного ряда;

k – номер гармоники;

t – порядковый номер элементов исходного ряда, $t = 1, 2, \dots$.

$$X(t) = a_0 + \sum_{i=1}^n [a_i \cdot \cos(2 \cdot \pi \cdot K_i \cdot t / N) + b_i \cdot \sin(2 \cdot \pi \cdot K_i \cdot t / N)] + \varepsilon(t) + d_0 + d_1 \cdot t, \quad (4)$$

где:

N – число элементов исходного ряда;
 n – число гармоник полигармонического полинома;
 K_i – коэффициенты, определяющие номер гармонии;
 $\hat{a}(t)$ – прогнозная оценка случайной компоненты;
 t – порядковый номер элементов исходного ряда, $t = 1, 2, \dots$

Критерий Дарбина—Уотсона (или DW-критерий) — статистический критерий, используемый для тестирования автокорреляции первого порядка элементов исследуемой последовательности. Наиболее часто применяется при анализе временных рядов и остатков регрессионных моделей [6].

Окончательно качество модели проверяется на величине статистики Дарбина-Уотсона d (Формула 14) [5].

$$d = \frac{\sum_{t=2}^N (e_t - e_{t-1})^2}{\sum_{t=2}^N e_t^2}, \quad (5)$$

где:

N – число элементов исходного ряда;

e – значение автокорреляции полученное из разницы значений точек основного ряда от значений точек анализируемого исходного ряда;

t – порядковый номер элементов исходного ряда, $t = 1, 2, \dots$

Значение статистики Дарбина-Уотсона изменяется в диапазоне от 0 до 4. При этом $dd = 2$ указывает на отсутствие автокорреляции элементов временного ряда. Если dd меньше двух, то имеет место положительная автокорреляции, а больше двух — отрицательная.



Рис. 4. Разложение временного ряда на базовые компоненты

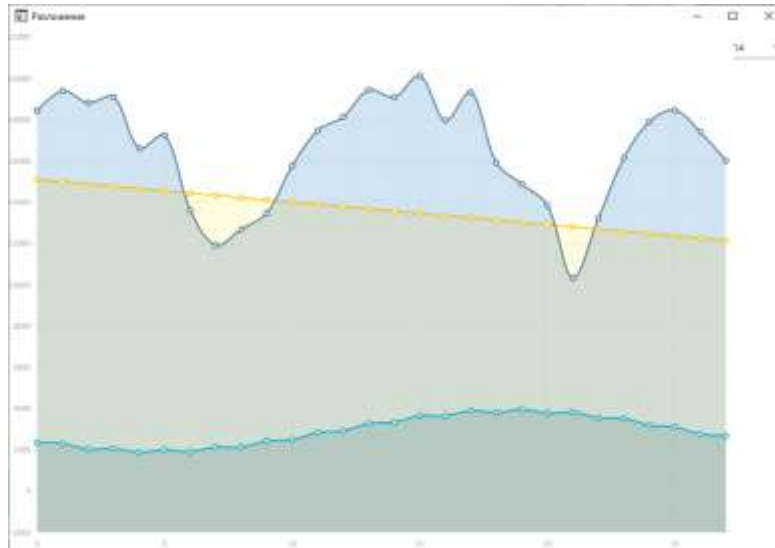


Рис. 5. Разложение временного ряда на базовые компоненты

Прогнозирование временного ряда показано на рисунке 6. Из него видно, что прогноз был построен на базе математической модели полигармонического полинома с использованием критерия Дарбина-Уотсона. Для формирования линии обучения было использовано 3 гармоник, коэффициенты тренда что дают понижающую линию и вычитание случайно компоненты для остановки критерия Дарбина-Уотсона. Длина линии прогноза выбиралась с учетом наложения на основную линию статистических данных для сравнения значений.

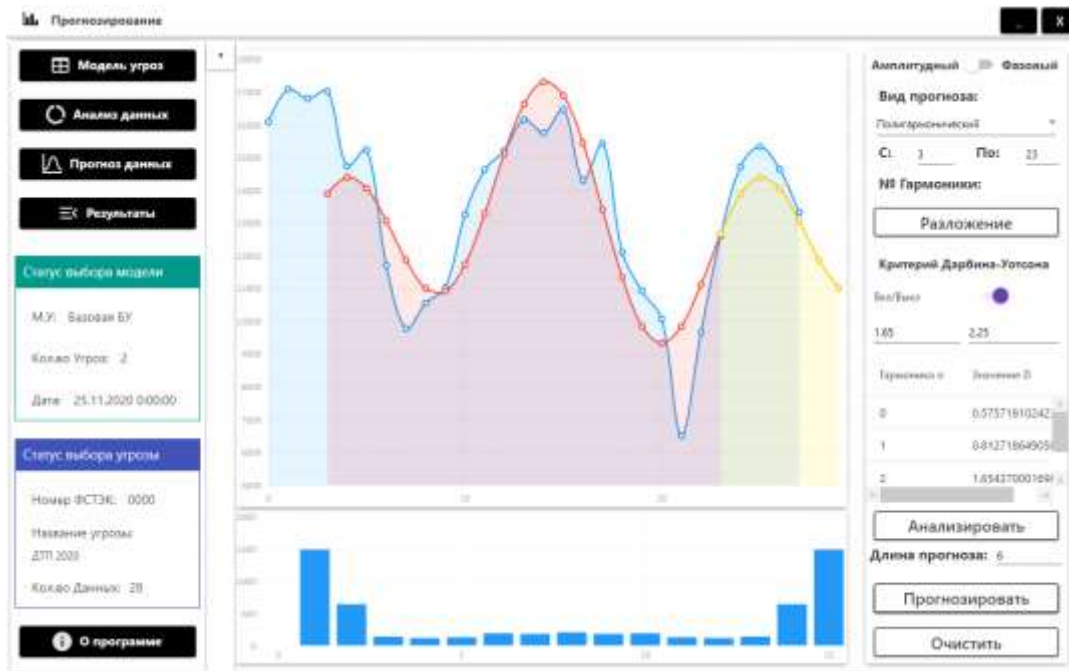


Рис. 6. Прогнозирование статистических данных дорожно-транспортных происшествий.

3 Оценить эффективность автоматизированного метода прогнозирования и анализа данных на статистических данных по автомобильным дорожно-транспортным происшествиям.

Прогноз временного ряда был выполнен с погрешностью в 12% что приемлемо для данной статистической выборки, так как полностью отражает фактическое положение дел по вопросу дорожно-транспортных происшествий. Прогнозируемые значения показывают сезонный спад ДТП в России в ноябре и декабре 2020 года и приближение к уровню поддержки в 9780 единиц ДТП. Прогнозирование статистических данных

временного ряда ДТП в России необходимо для дальнейшего принятия решений ГИБДД и каждого участника дорожного движения, в частности.

Достоинства данной работы заключаются в реализации математической модели полигармонического полинома на практике прогнозирования статистических данных ДТП в России, рассмотрении общей картины ДТП в мире на сегодняшний день, а также к достоинствам работы можно отнести актуальность статистических данных. К недостаткам работы можно отнести выбор одной модели прогнозирования и отсутствие сравнения с другими моделями прогнозирования временных рядов.

В заключение можно отметить, что автоматизированные методы прогнозирования на сегодняшний день помогают быстрому принятию решений в конкретной предметной области за счет быстрой обработки большого количества статистических данных. Что касается статистических данных ДТП в России, то выборка была сформирована локальная за один год и она не может отражать полной картины ДТП в России за все время, соответственно и прогнозирование статистических данных на базе такой выборки может происходить в локальном промежутке не на длительный срок.

Список литературы / References

1. Анализ рисков информационной безопасности для малого и среднего бизнеса [Электронный ресурс]. Режим доступа: <https://publications.hse.ru/mirror/pubs/share/folder/gjzfg00ohn/direct/157681455.pdf/> (дата обращения: 16.16.2020).
2. *Аралбаев Т.З.* Построение адаптивных систем мониторинга и диагностирования сложных промышленных объектов на основе принципов самоорганизации; Т.З. Аралбаев; Уфа: Гилем, 2003. 247 с.: ил.
3. *Аралбаев Т.З.* Проектирование вычислительных систем: лабораторные работы и методические указания / Т.З. Аралбаев, Р.Р. Галимов, Р.И. Хасанов; Оренбургский гос. ун-т. Оренбург: ОГУ, 2012. 64 с.
4. *Атаманова М.В.* Прогнозирование показателей производственной безопасности / М.В. Атаманова. [Электронный ресурс]. Режим доступа: <https://cyberleninka.ru/article/n/prognozirovanie-v-korotkih-vremennyh-ryadah-metodologicheskie-i-metodicheskie-aspekty/> (дата обращения: 09.04.2020).
5. *Барбашова Е.В.* Прогнозирование в коротких временных рядах: методологические и методические аспекты / Е.В. Барбашова, И.В. Гайдамакина, Н.В. Польшакова. [Электронный ресурс]. Режим доступа: <https://cyberleninka.ru/article/n/prognozirovanie-v-korotkih-vremennyh-ryadah-metodologicheskie-i-metodicheskie-aspekty/> (дата обращения: 09.04.2020).
6. *Александров Н.Н.* Прогнозирование количества транспортных средств, вовлеченных в ДТП на двухполосных внегородских автомобильных дорогах / Александров Н.Н. Владимир: Издательство: ФАУ РОСДОРНИИ (Москва), 2012. С. 173-189.
7. *Луценко Е.В.* Адаптивная семантическая информационная модель прогнозирования рисков совершения ДТП / Луценко Е.В., Коржаков В.Е. Майкоп: Издательство: Адыгейский государственный университет (Майкоп), 2008. С. 55-59.
8. *Скоробогатченко Д.А.* Нечеткая нейросетевая модель для прогнозирования числа ДТП региона в условиях ограниченной информации / Скоробогатченко Д.А., Ерохин А.В. Волгоград: Издательство: Волгоградский государственный технический университет (Волгоград), 2014. С. 174-181.