

Научная статья  
УДК 343.01:343.13  
<https://doi.org/10.36511/2078-5356-2021-3-91-97>

**Практические аспекты  
статистического моделирования и прогнозирования преступности  
на основе данных рядов динамики**

**Терехов Андрей Михайлович<sup>1</sup>, Кувычков Сергей Иванович<sup>2</sup>, Миронов Никита Андреевич<sup>3</sup>, Смирнов Сергей Александрович<sup>4</sup>**

<sup>1</sup>Российский государственный университет правосудия, Нижний Новгород, Россия

<sup>2</sup>Национальный исследовательский университет «Высшая школа экономики», Нижний Новгород, Россия

<sup>3,4</sup>Нижегородская академия МВД России, Нижний Новгород, Россия

Автор, ответственный за переписку: Сергей Александрович Смирнов, [ser-smir@yandex.ru](mailto:ser-smir@yandex.ru)

**Аннотация.** Цель работы состоит в практической реализации регрессионных моделей преступности, которые покажут комплексное воздействие различных социально-экономических факторов на состояние преступности в России, определяют прогноз на ближайшую перспективу. Результатом данного исследования является построение линейной многофакторной регрессионной модели и прогнозной ARMA-модели с тремя вариантами прогноза. На основе корреляционного анализа определены факторы, имеющие наиболее сильную связь с переменной, характеризующей преступность. Модель множественной регрессии показала взаимосвязанное влияние отдельных факторов на состояние преступности в рамках разработанного уравнения. Качество полученных моделей подтверждено оценочными тестами на статистическую значимость. В качестве исходных данных исследования выступили официальные статистические данные российской статистики за 1992–2020 годы, а также использован массив данных из более 500 наблюдений.

**Ключевые слова:** преступность, прогноз, модель преступности, регрессия, статистический анализ, статистические данные

**Для цитирования:** Терехов А. М., Кувычков С. И., Миронов Н. А., Смирнов С. А. Практические аспекты статистического моделирования и прогнозирования преступности на основе данных рядов динамики // Юридическая наука и практика: Вестник Нижегородской академии МВД России. 2021. № 3 (55). С. 91–97. <https://doi.org/10.36511/2078-5356-2021-3-91-97>.

Original article

**Practical aspects of statistical modeling and forecasting  
of crime based on time series data**

**Andrey M. Terekhov<sup>1</sup>, Sergey I. Kuvychkov<sup>2</sup>, Nikita A. Mironov<sup>3</sup>, Sergey A. Smirnov<sup>4</sup>**

<sup>1</sup>Russian State University of Justice, Nizhny Novgorod, Russian Federation

<sup>2</sup>Research University "Higher School of Economics", Nizhny Novgorod, Russian Federation

<sup>3,4</sup>Nizhny Novgorod academy of the Ministry of internal affairs of Russia, Nizhny Novgorod, Russian Federation

Corresponding author: Sergey A. Smirnov, [ser-smir@yandex.ru](mailto:ser-smir@yandex.ru), [ser-smir@yandex.ru](mailto:ser-smir@yandex.ru)

**Abstract.** The aim of the work is to implement regression models of crime in practice, which will show the complex impact of various socio-economic factors on the state of crime in Russia, and determine the forecast for the near future. The result of this study is the construction of a linear multivariate regression model and a predictive ARMA model with three forecast options. Based on the correlation analysis, the factors that have the strongest relationship with the variable that characterizes crime are determined.

© Терехов А. М., Кувычков С. И., Миронов Н. А., Смирнов С. А., 2021

The multiple regression model showed the interrelated influence of individual factors on the state of crime in the framework of the developed equation. The quality of the obtained models is confirmed by evaluation tests for statistical significance. The initial data of the study were: official statistical data of the Russian statistics for 1992–2020, an array of data from more than 500 observations was used.

**Keywords:** crime, forecast, crime model, regression, statistical analysis, statistical data

**For citation:** Terekhov A. M., Kuvychkov S. I., Mironov N. A., Smirnov S. A. Practical aspects of statistical modeling and forecasting of crime based on time series data. *Legal Science and Practice: Journal of Nizhny Novgorod Academy of the Ministry of Internal Affairs of Russia*, 2021, no. 3 (55), pp. 91–97. (In Russ.). <https://doi.org/10.36511/2078-5356-2021-3-91-97>.

*Окончание. Начало в № 2 (54).*

В предыдущей статье «Особенности статистического моделирования и прогнозирования преступности: теоретический аспект» нами был проведен анализ источников литературы, который обозначил результаты исследования отечественных и зарубежных авторов и показал эффективные модели и методики прогнозирования, а также группы факторов, оказывающие высокое влияние на развитие криминальной обстановки. Практическая реализация рассмотренных методик в отдельных странах показала различную достоверность (качество) прогнозных значений.

При этом в мировой практике наибольшую распространенность получили следующие методы прогнозирования: экстраполяция, моделирование, метод экспертных оценок [1]. Высокие показатели получили: метод случайной лесной регрессии (точность прогнозирования до 97 %), метод экспоненциального сглаживания Хольта с ежемесячной сезонностью, логистическая

регрессия и нейросетевая модель, простая линейная экстраполяция, а также некоторые другие методы, показавшие наибольшую эффективность применительно к конкретным пространственно-временным характеристикам и отдельным типам преступности [2, с. 61–70].

В данной работе представлены практические результаты моделирования и прогнозирования преступности в Российской Федерации на основе построения регрессионных моделей по данным динамических рядов.

Для реализации инструментария моделирования и прогнозирования регистрируемой преступности выбраны методы построения линейной модели множественной регрессии и построение авторегрессионной модели [3; 4]. Использование выбранных методик позволит показать величину зависимости преступности от других переменных, благодаря которым можно осуществить планирование мероприятий, направленных на снижение преступности [5].

Таблица 1

**Перечень показателей, использованных для проведения корреляционно-регрессионного анализа**

Наименование группы показателей	Наименование показателя
1	2
1. Показатели, характеризующие состояние правоохранительной системы	1.1. Число зарегистрированных преступлений, тыс. ед. (зависимая переменная Y); 1.2. Количество заключенных, тысяч человек (независимая переменная X1)
2. Показатели, характеризующие общее состояние социальной сферы	2.1. Численность населения страны, тысяч человек (независимая переменная X2); 2.2. Количество безработных, тысяч человек (независимая переменная X3); 2.3. Количество разводов (независимая переменная X4); 2.4. Ожидаемая продолжительность жизни при рождении, лет (независимая переменная X5)
3. Показатели, характеризующие затраты на социальную сферу	3.1. Объемы расходов на культуру из консолидированного бюджета Российской Федерации, млрд руб. (независимая переменная X6); 3.2. Объемы расходов на образование, млрд руб. (независимая переменная X7)
4. Показатели, характеризующие уровень материального благосостояния граждан	4.1. Среднедушевые денежные доходы населения (в месяц), руб. (независимая переменная X8); 4.2. Численность населения с денежными доходами ниже величины прожиточного минимума, млн человек (независимая переменная X9)

*Продолжение таблицы 1*

1	2
5. Показатели, характеризующие здоровье граждан	5.1. Заболеваемость психическими расстройствами и расстройствами поведения (количество взятых под наблюдение пациентов в отчетном году с диагнозом, установленным впервые в жизни), тысяч человек (независимая переменная X10); 5.2. Заболеваемость населения алкоголизмом и алкогольными психозами, тысяч человек (независимая переменная X11)
6. Показатели, характеризующие сферу образования	6.1. Численность студентов, обучающихся по программам подготовки квалифицированных рабочих, служащих, тысяч человек (независимая переменная X12); 6.2. Численность студентов, обучающихся по программам подготовки специалистов среднего звена, тысяч человек (независимая переменная X13); 6.3. Численность студентов, обучающихся по программам бакалавриата, специалитета, магистратуры, тысяч человек (независимая переменная X14)
6. Показатели, характеризующие внешнюю миграцию	6.1. Прибыло в Российской Федерации из стран СНГ и стран дальнего зарубежья, тысяч человек (независимая переменная X15); 6.2. Выбыло из Российской Федерации в страны СНГ и страны дальнего зарубежья, тысяч человек (независимая переменная X16); 6.3. Миграционный прирост, тысяч человек (независимая переменная X17);
7. Показатели, характеризующие уровень развития человека в стране	7.1. Индекс человеческого развития (ИЧР) (независимая переменная X18)
Источник: составлено авторами	

Для исследования влияния различных факторов на состояние преступности в Российской Федерации были выбраны данные Росстата по показателям, представленным в таблице 1 [6, с. 35].

Таким образом, было выделено 19 переменных (в том числе зависимая переменная Y). Показатели отражены в динамике за 1992–2018 годы в собственных величинах. С целью недопущения искажения результатов анализа из-за влияния накопленной инфляции стоимостные показатели учитывались в сопоставимых ценах базисного 1992 года. В качестве зависимой переменной, характеризующей состояние преступности, взят показатель «число зарегистрированных преступлений» (тыс. ед.). Далее были выделены корреляционные связи Y с влияющими переменными (табл. 2), рассмотренными в таблице 1.

Интерпретируя полученные результаты, следует обратить внимание на то, что

наиболее тесная связь наблюдается между факторами: X1, X2, X5, X11, X12, X16, X18 и Y. Степень влияния остальных факторов на количество лиц, совершивших преступления, невысокая. При этом следует отметить, что показатели динамических рядов могут быть подвержены влиянию тенденции (наличие тренда), поэтому может наблюдаться явление ложной корреляции.

В программной среде Gretl нами с использованием метода наименьших квадратов была построена линейная модель множественной регрессии. Попытка усовершенствовать модель за счет включения в нее временного тренда показала отсутствие такой необходимости (переменная времени статистически не значима), поэтому в итоговую модель после анализа на мультиколлинеарность методом инфляционных факторов (табл. 3) были включены компоненты, показанные на рисунке 1.

*Таблица 2*

**Коэффициенты парной корреляции зависимой переменной Y и независимых переменных X1 – X18**

гyx1	гyx2	гyx3	гyx4	гyx5	гyx6	гyx7	гyx8	гyx9
0,50	-0,40	0,24	0,12	-0,58	-0,24	-0,31	-0,28	0,21
гyx10	гyx11	гyx12	гyx13	гyx14	гyx15	гyx16	гyx17	гyx18
0,29	0,65	0,52	0,34	0,44	-0,37	-0,44	-0,19	-0,43

Источник: составлено авторами

**Коэффициенты парной корреляции зависимой переменной Y  
и независимых переменных X1 – X18**

Наименование фактора (независимой переменной)	Тестовое значение
X 1	2,675
X 4	1,540
X 9	4,004
X 13	3,436
X 14	7,839
X 15	7,916

Примечание: значения  $> 10,0$  могут указывать на наличие мультиколлинеарности

Источник: составлено авторами в приложении Gretl

Для проверки обоснованности применения линейной модели были выполнены тесты на нелинейность: нелинейность (квадраты), нелинейность (логарифмы). Результат теста на нелинейность (квадраты) показал: р-значение =  $P(\text{Хи-квадрат} > 10,181) = 0,117234$ ;  $0,117234 > 0,05$ . Таким образом, гипотеза о корректности линейной модели не отвергается. Тест на нелинейность (логарифмы) показал: р-значение =  $P(\text{Хи-квадрат} > 7,4114) = 0,284469$ ;  $0,284469 > 0,05$ , что также не отвергает нулевую гипотезу о корректности линейной модели.

С учетом полученных результатов (рис. 1) модель можно характеризовать следующим образом. Множественный коэффициент корреляции  $R = 0,9582$  показывает, что теснота связи между количеством зарегистрированных преступлений и факторами, включенными в модель, высокая. Множественный коэффициент детерминации (R-квадрат)  $D = 0,9182$ , то есть 91,82 % вариации числа зарегистрированных преступлений объясняется вариацией изучаемых факторов. Проверка значимости коэффициента множественной корреляции путем сравнения фактического и табличного значений F-критерия (тест Фишера) показала адекватность модели ( $F_{\text{факт}} = 37,39 > F_{\text{табл}} = 2,60$ , при вероятности ошибки  $\alpha = 0,05$ ). Таким образом, тест Фишера подтвердил значимость всей модели. Достоверность по уровню значимости критерия Фишера (значимость  $F = 7,62609E-10$ ) значительно меньше 0,05. Таким образом, модель является статистически значимой.

Проверка справедливости выдвинутой гипотезы путем сравнения расчетного критерия

значимости с критическим уровнем значимости ( $p < 0,05$ ) показала значимость гипотезы по всем включенным в модель факторам. Тест Стьюдента ( $t_{\text{факт}} > t_{\text{табл}} = 2,086$ ) также подтвердил статистическую значимость всех факторов. Статистика Дарбина-Уотсона ( $DW \approx 2,0$  при  $d1 = 0,925$  и  $d2 = 1,974$ ) показала отсутствие автокорреляции остатков. Тест Уайта подтвердил нулевую гипотезу об отсутствии гетероскедастичности ( $LM = 7,46554$ ; р-значение =  $P(\text{Хи-квадрат} > 7,46554) = 0,82538 > 0,05$ ). Тестирование модели с использованием методики Бриша-Пэгана также подтвердило гипотезу об отсутствии гетероскедастичности ( $LM = 4,621814$ ; р-значение =  $P(\text{Хи-квадрат} > 4,621814) = 0,593148 > 0,05$ ).

Тест на нормальное распределение ошибок подтвердил гипотезу о том, что остатки распределены по нормальному закону ( $\text{Хи-квадрат} = 1,87978$ ; р-значение =  $0,39067 > 0,05$ ). Результаты теста Рамсея (RESET-тест) на ошибку спецификации указали на то, что в модели не хватает одного или нескольких важных объясняющих факторов ( $F(2, 18) = 9,84742$ ; р-значение =  $P(F(2, 18) > 9,84742) = 0,00129097$ ;  $0,00129097 < 0,05$ ). Здесь следует отметить, что на информационных ресурсах Росстата, МВД, прокуратуры не раскрыта статистическая информация по отдельным важным показателям (факторам), характеризующим, например, эффективность работы правоохранительных органов (количество раскрытых/нераскрытых преступлений), численный состав МВД и так далее, либо такая информация раскрыта за относительно небольшой временной интервал, что не позволило ее учесть при построении регрессионной модели.

Модель 4: МНК, использованы наблюдения 1992–2018 (T = 27)

Зависимая переменная: Y

	Коэффициент	Ст. ошибка	t-статистика	P-значение	
const	-4543,93	821,548	-5,531	2,05e-05	***
X1	2,22232	0,338856	6,558	2,17e-06	***
X4	-0,00254358	0,000485634	-5,238	3,99e-05	***
X9	21,9096	5,74395	3,814	0,0011	***
X13	1,23186	0,241010	5,111	5,33e-05	***
X14	0,512582	0,0490800	10,44	1,52e-09	***
X15	0,00226205	0,000293949	7,695	2,11e-07	***
Среднее зав. перемен	2710,384	Ст. откл. зав. перемен	467,1579		
Сумма кв. остатков	464397,6	Ст. ошибка модели	152,3807		
R-квадрат	0,918156	Испр. R-квадрат	0,893602		
F(6, 20)	37,39434	P-значение (F)	7,63e-10		
Лог. правдоподобие	-169,9722	Крит. Акаике	353,9445		
Крит. Шварца	363,0153	Крит. Хеннана-Куинна	356,6417		
Параметр rho	-0,007890	Стат. Дарбина-Вотсона	1,998796		

Рис 1. Результаты регрессионного анализа (источник: составлено автором в приложении Gretl)

Используя приведенные коэффициенты и учитывая результаты проведенных тестов, составлено уравнение модели:

$$\hat{Y} = -4,543,93 + 2,22 * X1 - 0,00254 * X4 + 21,9 * X9 + 1,23 * X13 + 0,513 * X14 + 0,00226 * X15$$

где  $\hat{Y}$  – число зарегистрированных преступлений, тыс. ед.

Комплексное влияние на объемы преступности в рамках данной модели оказывают факторы: X1, X4, X9, X13, X14, X15.

На основе данной модели трудно прогнозировать состояние преступности, поскольку невозможно определить поведение всех влияющих факторов, включенных в модель на перспективу. Для прогнозирования в среде Gretl нами была построена ARMA-модель (2, 0) со следующими характеристиками: уровень надежности – 95 %, R2 = 0,79. Значения информационных критериев следующие: AIC = 404,42, HQC = 406,13, BIC = 409,89. Использован точный метод максимального правдоподобия. Динамический ряд (1992–2020) предварительно был протестирован на стационарность (ADF-тест), ряд оказался стационарным в уровнях ряда. Модель успешно прошла все основные тесты на статистическую значимость: тест на нормальное распределение

ошибок, тест на отсутствие ARCH-процессов, LM-тест на отсутствие автокорреляции. Уравнение модели имеет следующий вид:

$$y_t = 2626,42 + 1,18243 * y_{t-1} - 0,371075 * y_{t-2} + e_t$$

где  $y_t$  – число зарегистрированных преступлений, тыс. ед.

В рамках данной модели нами рассчитаны варианты прогноза на среднесрочную перспективу 2021–2023 годов (рис. 2):

- расчетный прогноз (предсказание): 2021 год – 2 161,418 тыс. ед., 2022 год – 2 292,636 тыс. ед., 2023 год – 2 404,295 тыс. ед.;
- позитивный сценарий: 2021 год – 1 732,938 тыс. ед., 2022 год – 1 629,096 тыс. ед., 2023 год – 1 608,084 тыс. ед.;
- негативный сценарий: 2021 год – 2 589,899 тыс. ед., 2022 год – 2 956,177 тыс. ед., 2023 год – 3 200,506 тыс. ед.

Таким образом, расчетный прогноз показывает незначительный рост регистрируемой преступности на среднесрочную перспективу, остальные варианты свидетельствуют, соответственно, снижение и значительный рост показателя. При этом реализация конкретного сценария будет зависеть от эффективности проводимой государством социально-экономической политики.



**Рис 2. Результаты прогнозирования количества зарегистрированных преступлений в Российской Федерации за период 2021–2023 годов (источник: составлено авторами)**

Следует отметить, что данная модель не может прогнозировать реальные объемы преступности, поскольку значительная их часть латентна и не регистрируется. При этом разработка методик прогнозирования общего объема преступности актуальна и требует выработки новых подходов к количественной оценке не учитываемых (из-за их латентности) показателей.

Подводя итог проведенному исследованию, отметим следующие основные результаты:

- выделены статистически значимые факторы, оказывающие взаимосвязанное влияние на состояние преступности в рамках полученной многофакторной модели: количество заключенных, количество разводов, численность населения с денежными доходами ниже величины прожиточного минимума, численность студентов, обучающихся по программам подготовки специалистов среднего звена, численность студентов, обучающихся по программам бакалавриата, прибыло в Российской Федерации из стран СНГ и стран дальнего зарубежья;

- разработанная прогнозная ARMA-модель показала три варианта прогноза регистрируемой преступности. Реализация того или иного варианта будет зависеть от эффективности реализуемой государством социально-экономической политики.

Использование прогностического подхода в практике деятельности аппарата государственного управления способствует повышению качества принятия управленческих решений.

Проведение дальнейших исследований, направленных на совершенствование моделей и инструментов моделирования, будет способствовать формулировке более точных выводов относительно взаимосвязи между различными параметрами состояния преступности.

#### Список источников

1. Пилюгина Т. В., Пичкурено Е. А., Вендина А. А. Математические методы прогнозирования преступности // *Aspectus*. 2015. № 1. С. 98–103.
2. Пилюгина Т. В., Натура Д. А. Использование метода математического моделирования при прогнозировании региональной преступности в вопросах ее предупреждения // *Всероссийский криминологический журнал*. 2017. Т. 11. № 1. С. 61–70.
3. Богданова М. В., Паршинцева Л. С., Квачко В. Ю. Методика моделирования и прогнозирования преступности в Российской Федерации // *Правовая информатика*. 2019. № 4. С. 50–62.
4. Осипов В. А. Некоторые аспекты прогнозирования развития преступности экстремистской направленности в Российской Федерации // *Вестник Белгородского юридического института МВД России имени И. Д. Путилина*. 2019. № 4. С. 17–22.
5. Маркина С. А. Оценка и прогнозирование экономической преступности в системе обеспечения экономической безопасности // *Известия Юго-Западного государственного университета. Серия: Экономика. Социология. Менеджмент*. 2019. Т. 9. № 6 (35). С. 147–162.
6. Буньковский Д. В., Капустюк П. А. Экономические факторы, влияющие на развитие организованной преступности // *Деятельность правоохранительных органов*. 2019. № 1. С. 10–15.

тельных органов в современных условиях: сборник материалов XXII международной научно-практической конференции: в 2 т. Иркутск, 2017. С. 35–39.

### References

1. Pilyugina T. V., Pichkurenko E. A., Vendina A. A. Mathematical methods of crime forecasting. *Aspectus*, 2015, no. 1, pp. 98–103. (In Russ.)

2. Pilyugina T. V., Natura D. A. Using the method of mathematical modeling in predicting regional crime in the issues of its prevention. *All-Russian criminological journal*, 2017, vol. 11, no. 1, pp. 61–70. (In Russ.)

3. Bogdanova M. V., Parshintseva L. S., Kvachko V. Yu. Methods of modeling and fore-casting crime in the Russian Federation. *Legal Informatics*, 2019, no. 4, pp. 50–62. (In Russ.)

4. Osipov V. A. Some aspects of forecasting the development of extremist crime in the Russian Federation. *Bulletin of the Belgorod law Institute of the Ministry of internal Affairs of Russia named after I. D. Putilin*, 2019, no. 4, pp. 17–22. (In Russ.)

5. Markina S. A. Assessment and forecasting of economic crime in the system of ensuring economic security. *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta. Series: Economics. Sociology. Management*, 2019, vol. 9, no. 6 (35), pp. 147–162. (In Russ.)

6. Bunkovsky D. V., Kapustyuk P. A. Economic factors influencing the development of organized crime. In the book: The activities of law enforcement agencies in modern conditions: collection of materials of the XXII International scientific and practical conference: in 2 vol. Irkutsk, 2017. Pp. 35–39. (In Russ.)

### Информация об авторах

**А. М. Терехов** – кандидат экономических наук, доцент;

**С. И. Кувычков** – кандидат юридических наук, доцент;

**Н. А. Миронов** – кандидат физико-математических наук;

**С. А. Смирнов** – без ученой степени.

### Information about the authors

**A. M. Terekhov** – Candidate of Sciences (Economy), Associate Professor;

**S. I. Kuvychkov** – Candidate of Sciences (Law), Associate Professor;

**N. A. Mironov** – Candidate of sciences (Physics and Mathematics);

**S. A. Smirnov** – no academic degree.

Статья поступила в редакцию 02.04.2021; одобрена после рецензирования 28.06.2021; принята к публикации 31.08.2021.

The article was submitted 02.04.2021; approved after reviewing 28.06.2021; accepted for publication 31.08.2021.