

Халафян Алексан Альбертович

доктор технических наук,
профессор кафедры прикладной математики
Кубанского государственного университета
dom-hors@mail.ru

Khalafyan Alexan Albertovich

Dr.Sci.,
Professor of the Applied Mathematics Department,
Kuban State University
dom-hors@mail.ru

Ракачев Вадим Николаевич

кандидат исторических наук,
доцент кафедры социологии
Кубанского государственного университета
dom-hors@mail.ru

Rakachev Vadim Nikolayevich

PhD in History,
Assistant Professor of
the Social Science Department,
Kuban State University
dom-hors@mail.ru

**ОПЫТ ДЕМОГРАФИЧЕСКОГО
ПРОГНОЗИРОВАНИЯ
С ПРИМЕНЕНИЕМ МЕТОДОВ
СТАТИСТИЧЕСКОГО
МОДЕЛИРОВАНИЯ (НА ПРИМЕРЕ
КРАСНОДАРСКОГО КРАЯ)**

**DEMOGRAPHIC
FORECASTING
WITH APPLICATION OF
STATISTICAL MODELING
METHODS (CASE STUDY OF
THE KRASNODAR KRAI)**

Аннотация:

В статье приводятся результаты прогнозирования численности и показателей естественного прироста для населения Краснодарского края, полученные с помощью методов статистического моделирования. Рассматриваются особенности применения данных методов при моделировании демографических показателей и процессов.

Ключевые слова:

математические методы, демографический прогноз, методы статистического моделирования, население, Краснодарский край.

Summary:

This paper presents the results of forecasting of the population size and natural increase in the Krasnodar Krai, obtained with application of the statistical modeling methods. The authors consider specifics of introduction of these methods for modeling of the demographic showings and processes.

Keywords:

mathematical methods, demographic forecasting, statistical modeling methods, population, Krasnodar Krai.

Развитие современного общества в значительной степени происходит при участии научного знания, начиная от диагностики и анализа текущих социальных процессов и явлений и заканчивая их прогнозированием и моделированием.

Потребность в социальных прогнозах наиболее актуальна в отношении демографических процессов и явлений. Значимость подобного рода исследований продиктована также тем, что обострение демографической ситуации является следствием серьезных экономических и социальных перемен, которые произошли в обществе за последние десятилетия. Как будет развиваться население в целом, как будет изменяться его структура, какие факторы в первую очередь будут определять эти изменения? Современная наука в состоянии дать ответы на эти вопросы, используя накопленный опыт моделирования и прогнозирования демографических процессов. Такие модели и прогнозы, разработанные на основе синтеза гуманитарных, технических, компьютерных знаний выступают основой при разработке мер социальной политики, социального инжиниринга и проч.

В рамках работы над проектом № 11-06-00346-а «Математическое моделирование демографических процессов в российском регионе», выполняемым при финансовой поддержке Российского фонда фундаментальных исследований, нами была предпринята попытка прогнозирования и создания моделей динамики численности населения

Краснодарского края и его национальной структуры как в ретроспективе, так и в перспективе до 2020 г.

Расчет демографического прогноза и моделей основывался на данных всеобщих переписей населения СССР – Российской Федерации 1926, 1959, 1970, 1979, 1989, 2002, 2010 гг. по Краснодарскому краю. Значительный временной лаг между переписями населения, потребовал дополнительных сведений о численности населения региона в межпереписные периоды. Необходимые данные – материалы текущего статистического учета были выявлены в центральных и региональных архивах. Однако для того, чтобы можно было использовать инструментальный временных рядов при прогнозировании возможной численности населения края на 2020 г., необходимо было заполнить пропущенные данные по численности населения с 1926 по 2010 гг.

Для решения этой задачи было проведено интерполирование недостающих значений численности населения в Краснодарском крае с использованием модуля программы STATISTICA «Множественная нелинейная регрессия». Пропущенные значения рассчитывались посредством аппроксимации (интерполирования) эмпирических данных полиномами третьей или четвертой степени.

Ниже на рисунках отображены графики численности населения (общей и по национальностям) в верхней части которых указаны полиномы для интерполирования пропущенных значений. На втором рисунке отображен график после вычисления пропущенных значений.

В качестве аргументов x использованы номера дат (1, 2, ..., 85), y – расчетная численность населения. При интерполяции возможны погрешности, однако они не будут существенно отличаться от ошибок, которые допускаются при переписи населения.

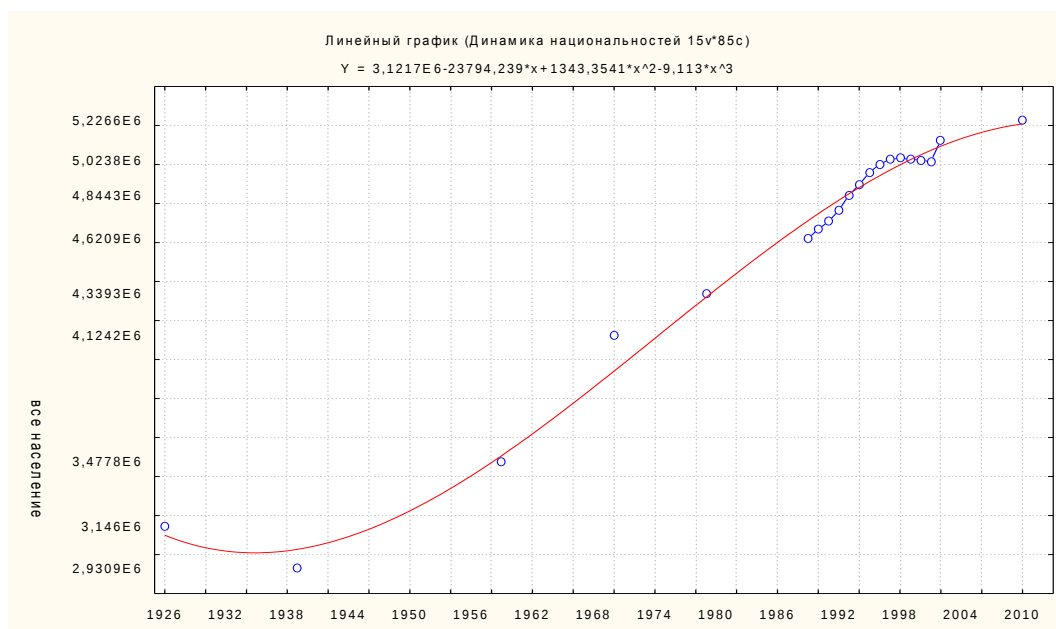


Рисунок 1 – Численность населения Краснодарского края с 1926 по 2010 гг. до интерполирования

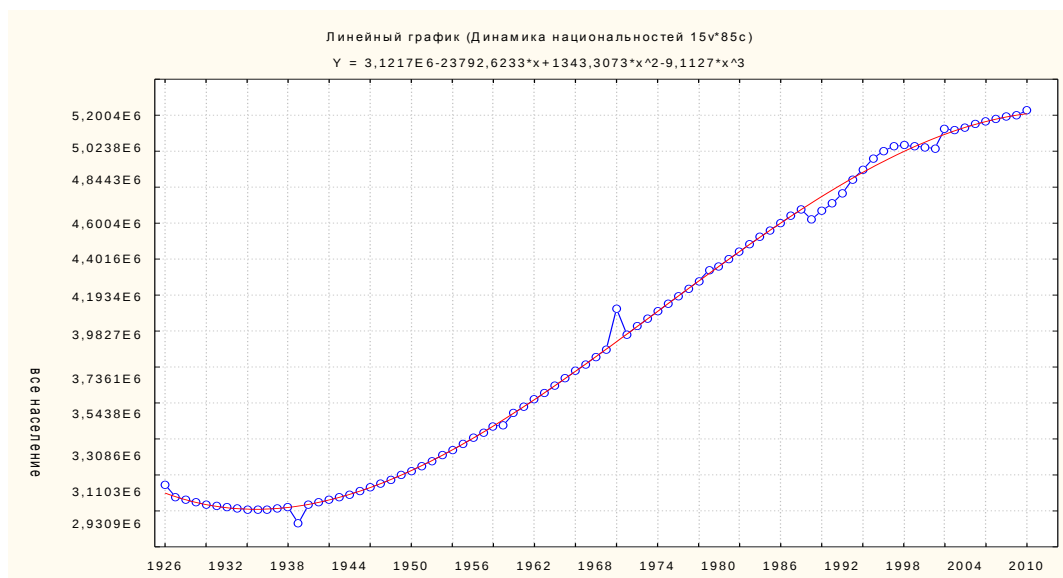


Рисунок 2 – Численность населения Краснодарского края с 1926 по 2010 гг. после интерполирования (выбросы соответствуют данным переписей)

На основе полученного ряда данных было осуществлено прогнозирование численности населения Краснодарского края в целом и отдельных этнических групп в его составе в частности с использованием метода экспоненциального сглаживания.

При данном методе наиболее простая и ясная модель временного ряда имеет следующий вид: $X_t = \beta + E_t$, где β – константа; E_t – случайная ошибка. Константа β относительно стабильна на каждом временном интервале, но может также медленно изменяться со временем. Один из интуитивно ясных способов выделения β состоит в том, чтобы использовать сглаживание скользящим средним, в котором последним наблюдениям приписывается больший вес, чем предпоследним, предпоследним – еще больший вес, чем предпредпоследним и т. д. Простое экспоненциальное сглаживание именно так и устроено. Здесь более старым наблюдениям приписываются экспоненциально убывающие веса, при этом, в отличие от скользящего среднего, учитываются все предшествующие наблюдения ряда, а не те, что попали в определенное окно. Точная формула простого экспоненциального сглаживания имеет вид

$$S_t = \alpha X_t + (1 - \alpha) S_{t-1}.$$

Когда эта формула применяется рекурсивно, каждое новое сглаженное значение (которое является также прогнозом) вычисляется как взвешенное среднее текущего наблюдения и сглаженного ряда. Очевидно, результат сглаживания зависит от параметра α (альфа). Если α равно 1, то предыдущие наблюдения полностью игнорируются. Если α равно 0, то игнорируются текущие наблюдения. Значения α между 0 и 1 дают промежуточные результаты. Эмпирические исследования показали, что весьма часто простое экспоненциальное сглаживание дает достаточно точный прогноз.

Пользователь может задавать начальное значение параметров сглаживания, начальное значение тренда и (если требуется) сезонные факторы. Для тренда и сезонной составляющей могут быть заданы независимые параметры сглаживания.

Для экспоненциального сглаживания была использована демпфированная модель без сезонной составляющей. Параметры модели определялись автоматическим поиском.

Для оценки адекватности модели используются графики, на которых вместе с исходным рядом в подходящем масштабе по оси Y изображаются его сглаженный вариант, прогноз и ряд остатков.

Ниже на рис. 3 показаны исходные значения ряда (синим цветом) и сглаженные значения (красным цветом). С 2010 г. пунктиром красного цвета изображены прогнозные значения ряда до 2020 г.

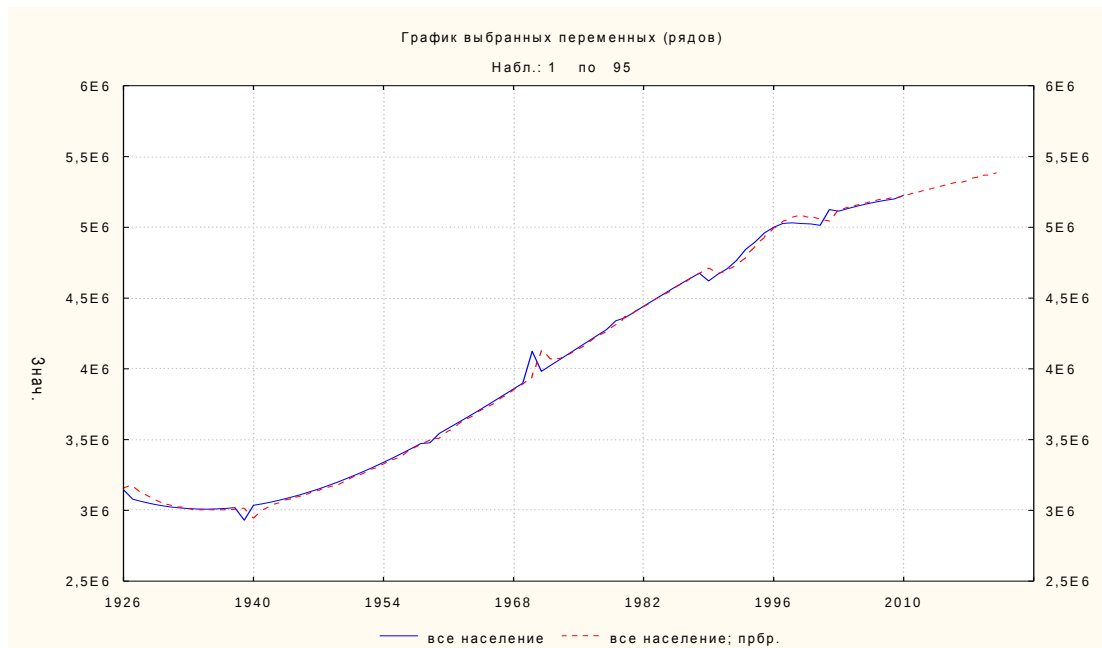


Рисунок 3 – Фактическая и прогнозная численность населения Краснодарского края в период с 1926 по 2020 гг.

Прогнозные значения, полученные в результате расчетов, составили на 2013 г. – 5,274 млн чел., на 2014 г. – 5,291; на 2015 г. – 5,307; на 2016 г. – 5,323; на 2017 г. – 5,339; на 2018 г. – 5,354; на 2019 г. – 5,369; на 2020 г. – 5,385 млн чел. Как видим, динамика, обозначенная в прогнозе, носит положительный характер, однако темпы невелики.

С использованием фактических данных за период с 1959 по 2012 гг. рассчитан прогноз общей численности населения края с применением аппроксимации эмпирических значений квадратичной или кубической функций.

Были получены прогнозные значения общей численности населения Краснодарского края с использованием полиномиальных функций (с применением полинома второй и третьей степени).

На рис. 4 представлены графики исходной численности и аппроксимирующий полином второй степени. Уравнение полинома выписано в верхней части рисунка. Очевидно, что данное уравнение можно использовать для прогнозирования значений численности на период до 2020 г.

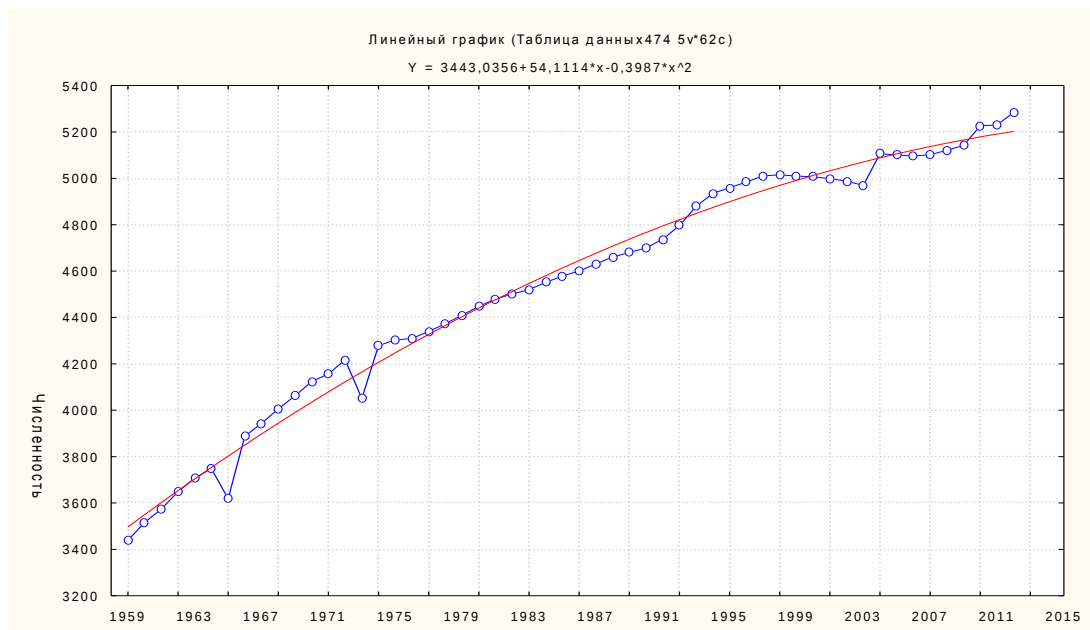


Рисунок 4 – Аппроксимация квадратичным полиномом численности населения Краснодарского края

На рис. 5 представлены графики исходной численности и аппроксимирующий полином третьей степени. Уравнение полинома выписано в верхней части рисунка. Данное уравнение также можно использовать для прогнозирования значений численности на период до 2020 г.

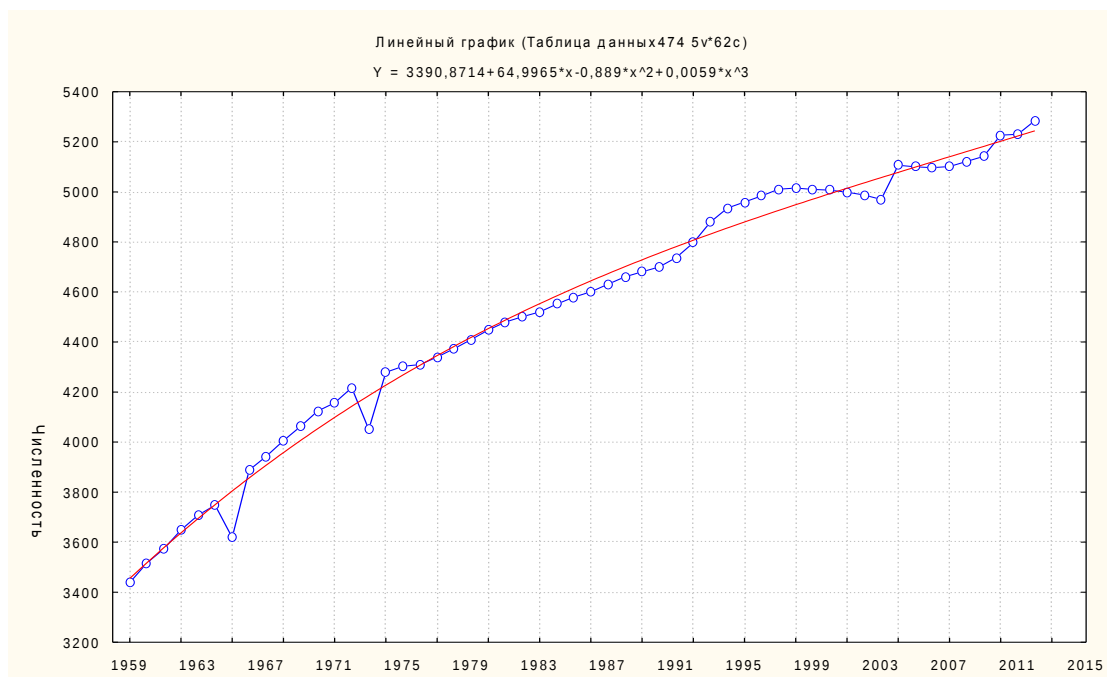


Рисунок 5 – Аппроксимация кубическим полиномом численности населения Краснодарского края

Сравнивая оба графика, отметим, что кубический полином отражает тенденцию более быстрого роста численности населения с 2007 г., чем квадратичный полином.

В качестве третьей альтернативной модели было использовано экспоненциальное сглаживание посредством демпфированной модели. Из рис. 6 видно, что прогнозная кривая модели принимает значения значительно меньшие, чем значения исходной численности, также прогнозные значения на период с 2012 по 2020 гг. не отражают тенденции роста численности населения, что свидетельствует о неадекватности построенной модели.

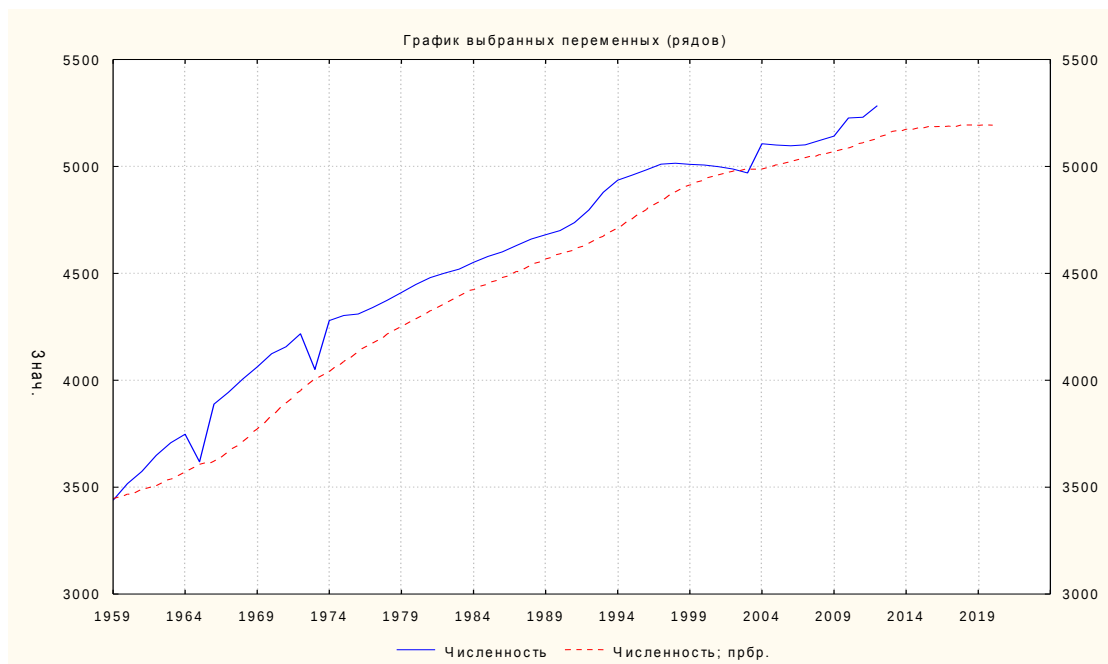


Рисунок 6 – Демпфированная модель экспоненциального сглаживания численности населения Краснодарского края

В качестве альтернативы демпфированной модели экспоненциального сглаживания была построена модель экспоненциального сглаживания с линейным трендом. В результате удалось построить более адекватную модель, но также обладающую определенным недостатком – прогнозная часть графика с 2012 по 2020 гг. имеет линейный тренд роста, что вряд ли будет соответствовать действительности (рис. 7).

В соответствии с предложенными моделями – квадратичная и кубическая аппроксимация, экспоненциальное сглаживание с линейным трендом – были вычислены прогнозные значения численности населения с 1959 по 2020 гг.

Таким образом, посредством различных математических методов при различных уровнях достоверности и степени приближения, получено четыре предсказанных значения общей численности населения Кубани на 2020 г.: 5265,3; 5409,5; 5493,3; 5389,4 (тыс. чел.).

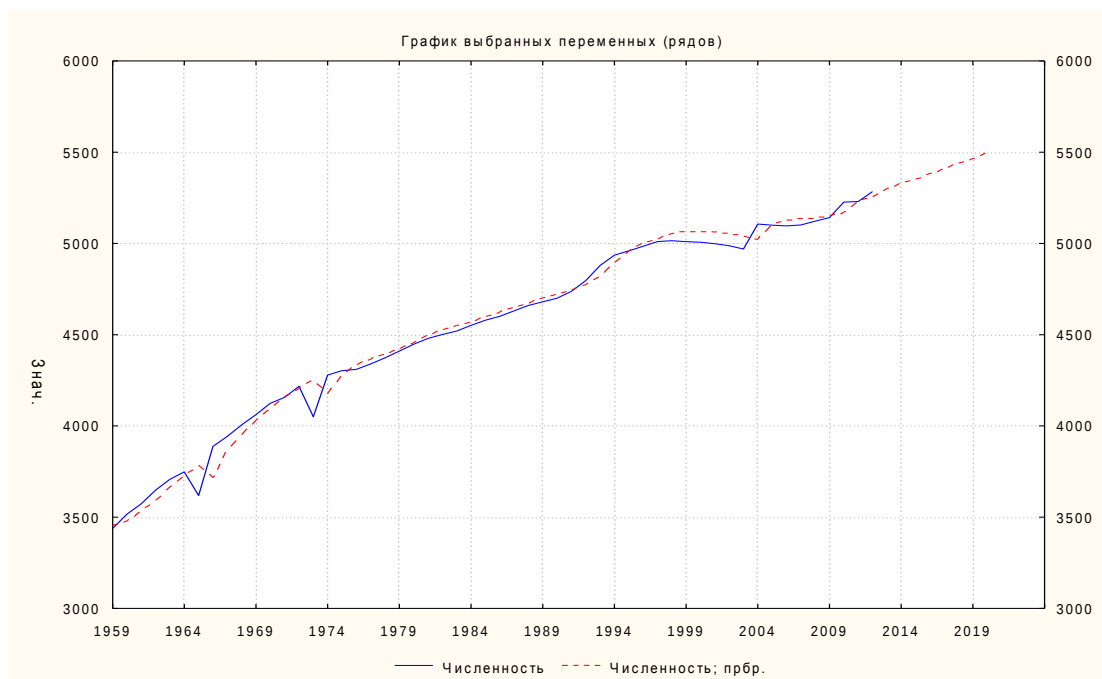


Рисунок 7 – Модель экспоненциального сглаживания с линейным трендом численности населения Краснодарского края

Так как каждая из рассмотренных моделей достаточно адекватна и отличается тенденцией изменения численности на интересующий нас момент времени, имеет смысл в качестве окончательного прогноза рассмотреть их усредненные значения.

В результате в качестве окончательного прогнозного значения численности населения на 2020 г. можно рассматривать цифру в 5389,4 тыс. чел. Таким образом, достоверным является утверждение, что если сохранятся тенденции изменения численности населения Краснодарского края, т. е. «будущее будет похоже на прошлое», численности населения достигнет к 2020 г. величины в 5500 тыс. человек.

Кроме моделирования общей численности населения Краснодарского края было проведено моделирование естественного прироста за период с 1960 по 2020 гг.

Исходные данные, использованные для расчетов, представлены с 1960 по 2011 гг. Известные значения численности населения Краснодарского края даны во временном интервале 1960–1989 гг. через каждые 5 лет: 1960, 1965, 1970, 1975, 1980, 1985, 1990 гг. Начиная с 1989 г. прирост населения представлен за каждый год. Динамика численности на основе исходных данных имеет следующий характер. С 1960 по 1990 гг. естественный прирост снижался и в 1990 г. стал отрицательным. Дальнейшее уменьшение естественного прироста происходит до 2000 г., затем проявляется тенденция увеличения прироста, но сохраняется его отрицательное значение.

Для возможности прогноза на 2020 г. посредством инструментария временных рядов необходимо было заполнить пустые ячейки таблицы с 1960 по 1989 гг. Для этого использовался модуль программы STATISTICA «Множественная нелинейная регрессия». Пропущенные значения рассчитывались посредством аппроксимации (интерполирования) эмпирических данных полиномами второй, третьей и четвертой степени.

На рисунках 8, 9 и 10 отображены графики прироста населения. В верхней части графиков указаны соответствующие полиномиальные функции. В качестве аргументов x использованы номера дат (1, 2, ..., 85), y – расчетный прирост населения.

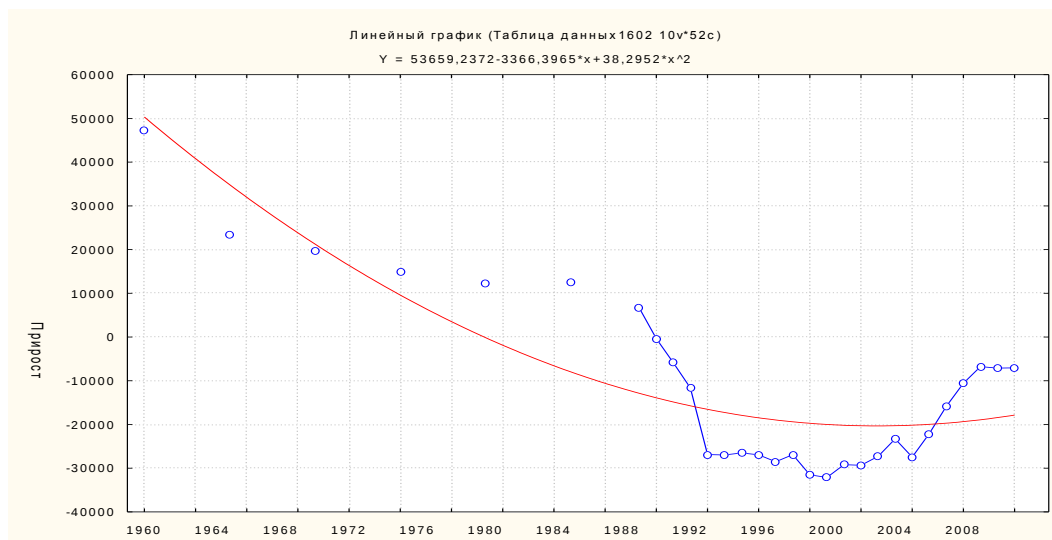


Рисунок 8 – Интерполирование численности населения Краснодарского края с использованием полиномиальной функции 2-й степени

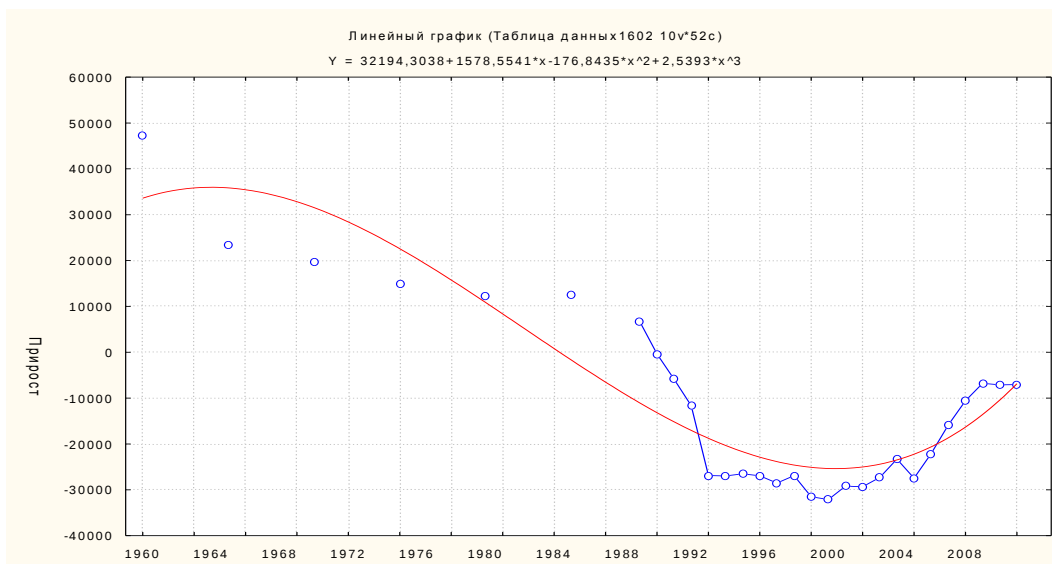


Рисунок 9 – Интерполирование численности населения Краснодарского края с использованием полиномиальной функции 3-й степени

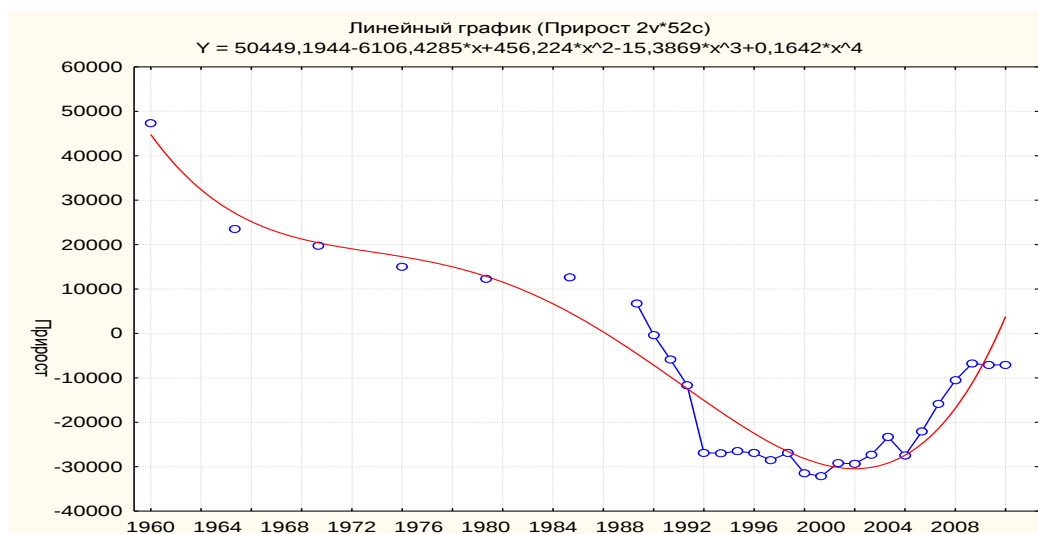


Рисунок 10 – Интерполирование численности населения Краснодарского края с использованием полиномиальной функции 4-й степени

Из графиков видно, что наилучшее приближение к фактическому тренду дает полином 4-й степени. В таблице 1 приведены итоги регрессионного анализа, в столбце *В* отображены коэффициенты регрессионного уравнения – полинома четвертой степени.

Таблица 1 – Коэффициенты регрессионного уравнения – полинома четвертой степени

N = 29	Итоги регрессии для зависимой переменной: Прирост $R=,96246658$ $R^2=,92634193$ Скорректир. $R^2=,91406558$ $F(4,24)=75,457$ $p<,00000$ станд. ошибка оценки: 6002,5				
	БЕТА	Стд. ош. БЕТА	В	Ст. ош. В	t (24)
Св. член			50449,19	6957,394	7,25116
Номер года	-4,0043	1,172781	-6106,43	1788,458	-3,41435
V1**2	17,5957	5,079264	456,22	131,696	3,46422
V1**3	-31,7314	7,420275	-15,39	3,598	-4,27631
V1**4	17,6399	3,513738	0,16	0,033	5,02027

В таблице в столбце *Прогноз пол. 4-й ст.* указаны значения прироста, посчитанные в соответствии с построенной моделью. В столбце *Прирост комб.* объединены значения столбцов *Прирост* и *Прогноз пол. 4-й ст.* добавлением в столбец *Прирост* недостающих значений из столбца *Прогноз пол. 4-й ст.*

На рис. 11 представлен линейный график комбинированного прироста с аппроксимирующим полиномом четвертой степени.

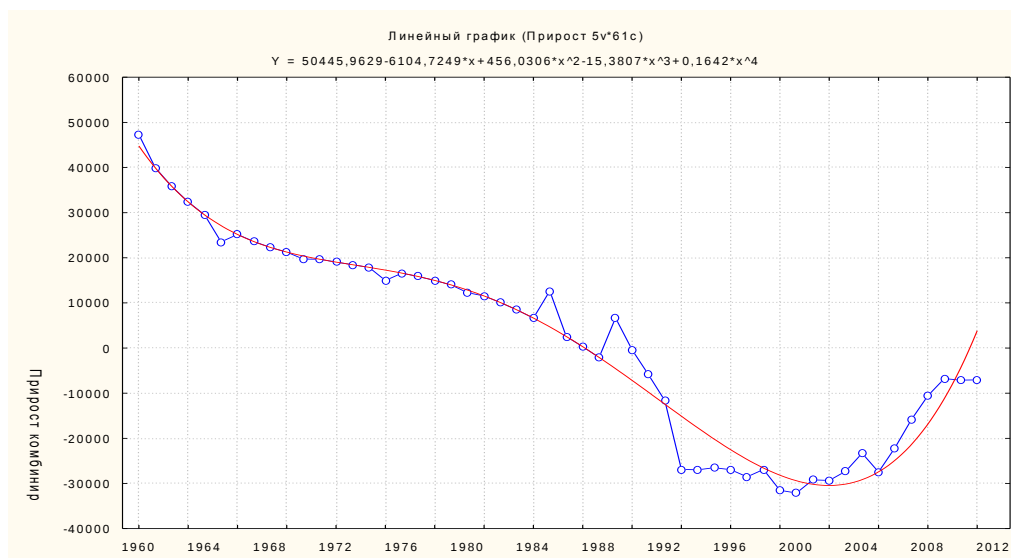


Рисунок 11 – Комбинированный прирост численности населения Краснодарского края с аппроксимирующим полиномом четвертой степени

Исходя из имеющихся данных, возможно спрогнозировать несколько сценариев развития демографической ситуации в Краснодарском крае. По первому сценарию, естественный прирост продолжит уменьшаться. По второму сценарию, естественный прирост может сохранять стабильное значение на отметке около -7000. По третьему сценарию, возможен рост показателей естественного прироста. Менее вероятен первый сценарий, так как с 2000 г. наметился тренд на увеличение прироста населения, маловероятен также и второй сценарий.

Согласно прогнозу до 2020 г., построенному методом экспоненциального сглаживания с демпфированным трендом и соответствующему второму сценарию, уровень прироста стабилизировался на отметке -6500.

Согласно прогнозу до 2020 г., построенному методом экспоненциального сглаживания с линейным трендом и соответствующему третьему сценарию, уровень прироста к 2020 г. достиг значения 6000.

Таким образом, сопоставляя полученные в результате моделирования данные с известными фактическими данными, можно сделать вывод о том, что в целом динамика численности населения региона во второй половине XX в. не подвергалась воздействию экстремальных факторов, что подтверждается близкими значениями реальной и прогнозной численности.

Согласно полученным прогнозам, численность населения Краснодарского края в перспективе до 2020 г. продолжит расти, однако темпы прироста будут невелики. Для показателей естественного прироста также наиболее вероятен вариант их дальнейшего роста, но они по-прежнему будут в минусе.