

ПРОГНОЗИРОВАНИЕ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ ПОКАЗАТЕЛЕЙ НА ОСНОВЕ ПРИНЦИПА БАЛАНСА ПЕРЕМЕННЫХ

Янковой А.Г.

(Одесса)

Рассматриваются проблемы прогнозирования переменных, связанных балансовыми равенствами. Предлагается принцип отбора уравнений в соответствии с качеством характеристик выравнивания и условием соблюдения балансового равенства.

Для многих развивающихся социально-экономических систем характерно наличие переменных, между которыми наблюдаются балансовые соотношения типа

$$y_{1t} + \dots + y_{m-1t} = y_{mt}. \quad (1)$$

Примером таких переменных могут служить годовые показатели рождаемости, смертности и естественного прироста населения; величины валового общественного продукта, национального дохода и материальных затрат; объемы внешнеторгового оборота, экспорта и импорта.

При решении задач прогнозирования в этой области ведущую роль играют методы экстраполяции, основанные на предпосылке об инерционности развития динамических вероятностных систем. Среди этих методов важное место занимают трендовые модели, описывающие вариацию изучаемых переменных в зависимости от времени t — обобщающего аргумента, аккумулирующего совокупное влияние множества главных факторов. Исходные ряды динамики y_{it} представляются в виде сумм двух ненаблюдаемых слагаемых

$$y_{it} = f_i(t) + \epsilon_i, \quad (2)$$

где $f_i(t)$ — уравнение тренда; ϵ_i — случайная компонента.

Тренды характеризуют основные закономерности движения параметров социально-экономической системы во времени.

Точность и надежность прогнозирования во многом зависят от правильности выбора формы тренда, вида аналитической функции $f_i(t)$. Наиболее точны и надежны результаты прогнозирования, если обнаружен истинный механизм развития процесса во времени. Следовательно, проблема качества прогнозов, полученных по трендовым моделям, сводится к поиску внутренних закономерностей роста, развития, диалектики.

Существует ряд приемов, с помощью которых можно более или менее успешно решать данную проблему в условиях изолированного прогнозирования отдельных переменных. Это прежде всего качественный теоретический анализ изучаемого процесса, позволяющий вскрыть его внутреннюю логику и механизм формирования во взаимосвязи с другими переменными. Однако возможности такого анализа обычно сильно ограничены. В лучшем случае он показывает характер динамики в самых общих чертах, что чаще всего недостаточно для обоснованного выбора формы тренда. Поэтому качественный анализ, являясь содержательной основой прогнозирования, дополняется исследованием эмпирических данных.

Первый путь (простейший) — визуальный — выбор формы тренда в результате визуального изучения графика ряда динамики. Но при этом риск субъективного решения очень велик: разные исследователи, используя один и тот же график, могут при-

вести достаточно убедительные аргументы в пользу различных форм кривой. К тому же на результат выбора сильно влияет масштаб графического изображения.

Второй путь, который обычно рекомендуется в литературе, — метод последовательных разностей. В соответствии с ним исчисляются первые, вторые и т.д. разности уровней ряда. Расчет ведется до тех пор, пока разности не станут примерно равными. Порядок разностей принимается за степень аппроксимирующего полинома. Недостатком такого подхода является возможность подбора кривых, описываемых только многочленами.

Существуют и другие способы определения формы тренда, основанные на анализе эмпирических данных, например метод характеристик прироста [1]. Все они обладают определенными достоинствами и недостатками и часто дают вполне удовлетворительные результаты при анализе и прогнозировании рядов динамики, рассматриваемых изолированно. Однако ситуация принципиально меняется, когда осуществляется одно-временный прогноз нескольких взаимозависимых переменных типа (1). В этом случае приведенные выше методы выбора формы трендов не исключают опасности получить ошибочный, расходящийся прогноз, нарушающий основополагающие количественные соотношения. Как бы глубоко ни были разработаны приемы экстраполяции отдельных изолированных рядов динамики, они не могут застраховать от получения так называемого "веера" прогнозов, отличающегося абсурдностью предсказываемых значений изучаемых переменных с точки зрения их внутренней зависимости.

Во избежание отмеченных недостатков при средне- и долгосрочном прогнозировании переменных, между которыми имеются объективные взаимосвязи, мы предлагаем традиционные методы подбора форм трендов дополнить принципом баланса переменных. Он может быть кратко сформулирован следующим образом: окончательный вывод о приемлемости тех или иных аналитических функций для описания динамики взаимосвязанных показателей (т.е. выбор предиктора) определяется степенью соблюдения для прогнозируемых значений переменных балансового соотношения (1).

Этот принцип как бы реализует невозможное — дает точку опоры в будущем. Действительно, что нам известно о будущем? Только то, что и в 1995 и в 2000 г. балансовое соотношение (1) будет выполняться так же, как оно выполняется сейчас или выполнялось в прошлом. Поэтому вполне логично в качестве предиктора выбрать такое сочетание аналитических функций изучаемых переменных, которое обеспечит наиболее сбалансированное значение прогнозных данных.

Принцип и построенный на нем критерий баланса переменных впервые был разработан в теории эвристической самоорганизации при решении задач технической кибернетики [2, 3]. Он носит универсальный характер и может применяться при изучении систем любой природы. Подходы самоорганизации отличаются эффективным использованием минимального количества априорной информации об исследуемом объекте, когда он рассматривается как "черный ящик". Поэтому принцип баланса переменных в сочетании с традиционными методами выбора формы трендов представляется чрезвычайно плодотворным при прогнозировании параметров сложных вероятностных систем.

Процесс прогнозирования состоит из двух стадий.

1. Определение возможных функций-кандидатов, наиболее точно описывающих изучаемые ряды динамики. Здесь применяются рассмотренные выше традиционные методы выбора формы трендов: качественный теоретический анализ, визуальное исследование графиков, расчет последовательных разностей, характеристик прироста и других параметров, включая критерии математической статистики (F -критерий Фишера, критерий Дарбина — Уотсона). В результате из всего перечня опорных функций, число которых может быть достаточно большим, отбирается несколько главных, наиболее адекватных изучаемым процессам. Здесь же методом наименьших квадратов определяются коэффициенты каждой из отобранных функций, осуществляется экстраполяция и рассчитываются прогнозные значения переменных на заданный период упреждения.

2. Последовательный перебор ограниченного числа главных функций, выделенных на предыдущей стадии. В качестве

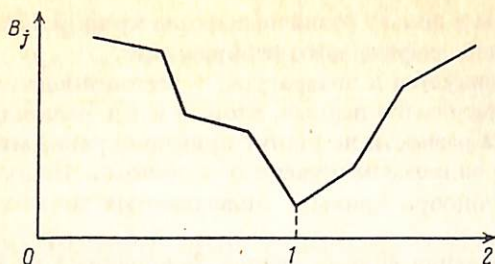


Рис. 1. Изменение значений критерия баланса переменных при последовательном переборе комбинаций функций-кандидатов: 1 — оптимальная комбинация, 2 — комбинация функций-кандидатов

оценки каждой комбинации выступает степень выполнения балансового соотношения (1) для всех точек периода упреждения. Наилучшая комбинация, т.е. предиктор, соответствует минимальному значению критерия баланса переменных

$$B_j = \frac{\sum_{l_1}^{l_2} (y_{mt} - y_{1t} - \dots - y_{m-1t})^2}{\sum_{l_1}^{l_2} y_{mt}^2} \rightarrow \min, \quad (3)$$

где l_1 — l_2 — период упреждения; B_j — суммарная относительная величина разбаланса прогноза по всем точкам периода упреждения. Очевидно, что чем ближе разбаланс прогноза к нулю, тем больше оснований считать именно эту комбинацию функций наилучшим предиктором и наоборот.

Если на первой стадии не пропущена ни одна пригодная функция, достаточно точно описывающая динамику изучаемых взаимосвязанных переменных, то множество значений B_j по мере перебора и усложнения форм трендов обязательно проходит через глобальный минимум, как это показано на рис. 1.

Общее число комбинаций функций-кандидатов j (предикторов), проверяемых на второй стадии, определяется величиной

$$j = \underbrace{a \cdot b \cdot \dots \cdot n}_m, \quad (4)$$

где m — количество взаимосвязанных переменных; $a, b, \dots, n$ — число функций-кандидатов, отобранных на первой стадии для описания каждой переменной.

Если, например, для трех переменных на первой стадии прогнозирования отобрано по 10 первоначальных функций-кандидатов, то $j = 10 \times 10 \times 10 = 1000$.

Очевидно, вручную перебрать такое количество комбинаций и рассчитать для каждой величину B_j не представляется возможным. На помощь приходят современные быстродействующие персональные компьютеры, которые позволяют полностью автоматизировать вторую стадию — перебор вариантов и расчет значений критерия баланса переменных. Остается только задать вид функции в числителе (3), вытекающий из сути изучаемых взаимосвязанных переменных. ЭВМ сама выбирает вид предиктора: важно лишь правильно указать среду поиска оптимальной комбинации.

Глубина минимума критерия баланса переменных отражает успешность процесса прогнозирования. Близость к нулю значений B_j указывает на наличие оптимального сочетания прогнозирующих функций, т.е. сигнализирует о нахождении оптимального предиктора. И наоборот, высокие значения разбаланса (порядка 10^{-3} — 10^{-1}) свидетельствуют об ошибках на первой стадии отбора — о пропуске некоторых важных функций-кандидатов, хорошо описывающих динамику изучаемых переменных. Поэтому первая стадия является определяющей во всем процессе прогнозирования. Подобно жерновам мельницы, которые не дают качественной муки при использовании плохого зернового материала, критерий баланса переменных (3) не может определить наилуч-

Таблица 1

Рождаемость, смертность и естественный прирост населения бывшего СССР (в расчете на 1000 чел.)

Год	Число родившихся y_{1t}	Число умерших y_{2t}	Естественный прирост y_{3t}
1955	25,7	8,2	17,5
1956	25,2	7,6	17,6
1957	25,4	7,8	17,6
1958	25,3	7,2	18,1
1959	25,0	7,6	17,4
1960	24,9	7,1	17,8
1961	23,8	7,2	16,6
1962	22,4	7,5	14,9
1963	21,2	7,2	14,0
1964	19,6	6,9	12,7
1965	18,4	7,3	11,1
1966	18,2	7,3	10,9
1967	17,3	7,6	9,7
1968	17,2	7,7	9,5
1969	17,0	8,1	8,9
1970	17,4	8,2	9,2
1971	17,8	8,2	9,6
1972	17,8	8,5	9,3
1973	17,6	8,7	8,9
1974	18,0	8,7	9,3
1975	18,1	9,3	8,8
1976	18,4	9,5	8,9
1977	18,1	9,6	8,5
1978	18,2	9,7	8,5
1979	18,2	10,1	8,1
1980	18,3	10,3	8,0
1981	18,5	10,2	8,3
1982	18,9	10,1	8,7
1983	19,8	10,4	9,4
1984	19,6	10,8	8,8
1985	19,4	10,6	8,8
1986	19,9	9,7	10,2
1987	19,8	9,9	9,9
1988	18,8	10,1	8,7
1989	17,6	10,1	7,5

ший предиктор, если на первой стадии в перебор не включены функции, наиболее адекватные исследуемым показателям. Следовательно, от профессиональных знаний, опыта и навыков ученого во многом зависит выбор оптимального варианта прогнозирующих функций.

В то же время вторая стадия, которая дает возможность оценить каждый вариант прогноза по наиболее целесообразному (с точки зрения поставленных целей) критерию, позволяет устранить неопределенность, часто возникающую на первой стадии. В этом смысле обе стадии играют важную роль в процессе определения оптимальной комбинации форм трендов. Поэтому они должны применяться в неразрывном единстве, как необходимое и достаточное условие успешного прогнозирования взаимосвязанных социально-экономических показателей.

Что касается случайностей при выборе оптимальной комбинации форм трендов, то они практически исключены сообразно закону невозможности свершения маловероятного события. В самом деле, вероятность случайного нулевого разбаланса прогнозных значений переменных даже на одной точке периода упреждения совпадает с

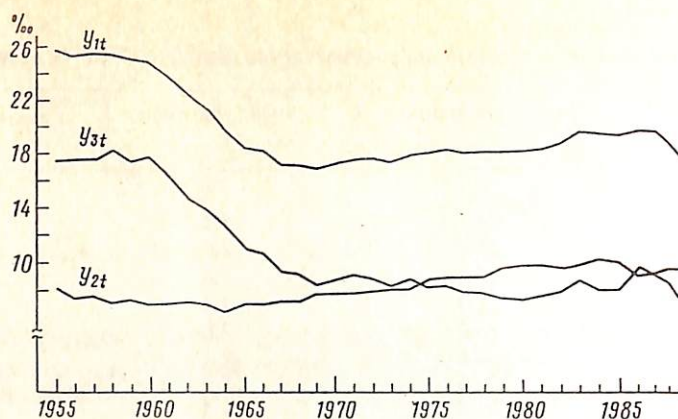


Рис. 2. Динамика рождаемости, смертности и естественного прироста населения бывшего СССР (в расчете на 1000 чел.)

вероятностью определенного расположения m трендов, часть из которых неверно описывает изучаемые ряды динамики.

Рассмотрим содержание обеих указанных стадий на примере прогнозирования показателей рождаемости, смертности и естественного прироста населения СССР [4]. Их динамика за последние 35 лет представлена в табл. 1*.

Приведенные ряды динамики графически изображены на рис. 2. Визуальный анализ показывает, что для 1950-х и первой половины 1960-х годов характерно довольно резкое снижение рождаемости, что является следствием "демографического эха" войны, при относительно стабильной смертности населения. Это привело к существенному понижению естественного прироста: с 18,1 в 1958 г. до 9,5‰ в 1968 г. В середине 1960-х и в 1970-е годы рождаемость несколько стабилизировалась (ее минимум — 17,0‰ приходится на 1969 г.) и начала постепенно возрастать, однако при этом увеличилась смертность населения. Последнее обусловлено распространением различных эпидемий, в частности гриппа, в марте 1965 г.

В 1975 г. впервые за изучаемый послевоенный период времени смертность превысила естественный прирост, причем вплоть до 1986 г., когда положение изменилось за счет некоторого роста рождаемости и снижения смертности в результате решительных мер, принятых в борьбе с пьянством и алкоголизмом. В последние годы снова наблюдается превышение смертности над естественным приростом, что вызвано главным образом снижением рождаемости: за 1986–1989 гг. она уменьшилась на 2,3‰. Основными причинами такой ситуации являются, на наш взгляд, неблагоприятная экономическая обстановка в стране, нарастание инфляционных процессов, неуверенность в завтрашнем дне, ухудшение половозрастной структуры населения в связи с массовым выездом на постоянное жительство за рубеж.

Первая стадия процесса прогнозирования показателей рождаемости (y_{1t}), смертности (y_{2t}) и естественного прироста (y_{3t}) населения бывшего СССР начинается с констатации очевидного балансового соотношения между изучаемыми переменными

$$y_{1t} - y_{2t} = y_{3t}. \quad (5)$$

Естественный прирост, характеризующий жизненный уровень народа, демографическую ситуацию и социально-экономическое положение трудящихся в стране, определяется как разность между рождаемостью и смертностью. В данном исследовании поставлена задача выбора наилучшего предиктора с целью получения среднесрочного прогноза переменных $y_{1t} - y_{3t}$ на 1990–1994 гг.

* Исходные данные, характеризующие изучаемые переменные, взяты из статистических сборников "Народное хозяйство СССР в цифрах" за соответствующие годы.

Варианты прогнозов изучаемых показателей по исходным рядам динамики за 1955 – 1989 гг. Таблица 2

Показатель	Вид функции	Уравнение тренда	Средняя ошибка аппроксимации, % $\bar{E}$	Коэффициент детерминации, R^2	Прогноз, годы				
					1990	1991	1992	1993	1994
Рождаемость, y_{1t}	1. Парабола 2-й степени	$\hat{y}_{1t} = 27,51611 - 0,89099t + 0,0198t^2$	4,91	0,830	21,095	21,649	22,242	22,876	23,548
	2. Экспонента	$\hat{y}_{1t} = \exp(3,14109 - 0,00958t)$	3,09	0,421	16,741	16,586	16,419	16,234	16,107
	3. Степенная	$\hat{y}_{1t} = 27,27164t - 0,12294$	2,29	0,627	17,554	17,495	17,438	17,382	17,328
	4. 1-я функция Торквиста	$\hat{y}_{1t} = \frac{18,55995t}{t - 0,43456}$	8,23	0,442	18,787	18,781	18,775	18,769	18,764
	5. Авторегрессия логистического типа	$\hat{y}_{1t} = \frac{y_{1t-1}}{0,92607 + 0,00429y_{1t-1}}$	2,30	0,940	17,572	17,546	17,523	17,501	17,480
Смертность, y_{2t}	1. Линейная	$\hat{y}_{2t} = 6,72401 + 0,11073t$	5,43	0,818	10,710	10,821	10,932	11,042	11,153
	2. Парабола 2-й степени	$\hat{y}_{2t} = 7,08943 + 0,05145t + 0,00165t^2$	5,08	0,833	11,075	11,247	11,422	11,600	11,781
	3. Линейно-гиперболическая	$\hat{y}_{2t} = 5,98217 + 0,13657t + \frac{2,32707}{t}$	4,03	0,887	10,964	11,099	11,234	11,369	11,504
	4. Экспонента	$\hat{y}_{2t} = \exp(1,92688 + 0,01268t)$	2,47	0,813	10,843	10,981	11,122	11,264	11,407
Естественный прирост, y_{3t}	1. Парабола 2-й степени	$\hat{y}_{3t} = 20,42897 - 0,94275t + 0,01815t^2$	6,94	0,915	10,012	10,394	10,813	11,268	11,759
	2. Экспонента	$\hat{y}_{3t} = \exp(2,79776 - 0,02364t)$	5,77	0,713	7,006	6,842	6,682	6,526	6,373
	3. Линейно-логарифмическая	$\hat{y}_{3t} = 20,9967 - 3,72478 \ln t$	10,81	0,815	7,649	7,547	7,448	7,351	7,256
	4. Степенная	$\hat{y}_{3t} = 23,50998t - 0,29827$	4,28	0,805	8,073	8,008	7,944	7,883	7,824
	5. Авторегрессия логистического типа	$\hat{y}_{3t} = \frac{y_{3t-1}}{0,95506 + 0,00654y_{3t-1}}$	5,13	0,924	7,469	7,440	7,413	7,386	7,362

Рассматривая на рис. 2 общую тенденцию изменения изучаемых показателей, следует отметить у переменных y_{1t} и y_{3t} хотя и не совсем четко выраженные минимальные значения. Это позволило выдвинуть гипотезу о криволинейности изменения уровней рождаемости и естественного прироста населения страны за исследуемый период времени. Поэтому в качестве первоначальных опорных функций для трендовых моделей указанных показателей использовались следующие: параболы второй и третьей степени, экспоненциальная и степенная функции, логистическая кривая, функция Гомперца и Торнквиста, авторегрессионные модели различных типов, колебательная функция и некоторые их комбинации.

Что касается смертности, то здесь тенденция вполне очевидна — небольшой, но довольно стабильный прирост переменной y_{2t} , поэтому для построения трендовых моделей применялись следующие функции: линейная, парабола второй степени, гипербола, экспонента и их комбинации.

В результате расчетов на персональных компьютерах, проведенных с помощью пакета прикладных программ АРМ СТАТИСТИКА, были получены три группы моделей, включающих 43 уравнения. Статистический анализ функций на базе F -, t -критериев позволил выделить 14 трендов, наиболее адекватно описывающих изучаемые признаки. В табл. 2 дан перечень отобранных функций, показатели точности аппроксимации и прогнозные значения переменных на 1990–1994 гг.

Данные граф 4 и 5 табл. 2 свидетельствуют о том, что все отобранные на первой стадии уравнения достаточно точно отражают динамику исследуемых процессов: средняя ошибка аппроксимации ($\bar{E}$) для трендов рождаемости колеблется от 2,30 до 8,23%, для моделей смертности — от 2,47 до 5,43, для моделей естественного прироста населения — от 4,28 до 10,81%.

В условиях изолированного прогнозирования социально-экономических показателей напрашивающимся и вполне логичным является традиционный подход, согласно которому в качестве моделей прогноза следует взять наиболее точные тренды. Для рождаемости это авторегрессия логистического типа ($\bar{E} = 2,30\%$, $R^2 = 0,940$), для смертности — экспонента ($\bar{E} = 2,47\%$, $R^2 = 0,813$), для естественного прироста — степенная функция ($\bar{E} = 4,28\%$, $R^2 = 0,805$). Иными словами, наилучшим на первый взгляд представляется сочетание трендов 5–4–4 табл. 2 ("автоэкспостепенной" предиктор).

Однако учитывая поставленную задачу — комплексное прогнозирование показателей рождаемости, смертности и естественного прироста населения, не будем торопиться с выводами. Ведь существует всего 100 ($5 \cdot 4 \cdot 5$) предикторов и соответственно 100 различных вариантов оценки будущих значений взаимосвязанных демографических показателей, различающихся по крайней мере величиной одного признака. Не исключено, что среди этих вариантов окажутся сочетания, точность которых по критерию баланса переменных будет выше, чем у тривиального решения (5–4–4).

Перейдем ко второй стадии прогнозирования — непосредственному перебору всех 100 комбинаций изучаемых функций по B_j . Перебор вариантов и расчет значений критерия баланса переменных по пятилетним прогнозным данным (гр. 6–10 табл. 2) осуществлялись на персональном компьютере с помощью программы, написанной на языке "Бейсик". Укажем здесь лишь основные результаты расчетов, представляющие практический интерес. Десять наименьших расчетных значений критерия баланса переменных приведены в табл. 3.

Из нее видно, что наиболее сбалансированный прогноз дает предиктор 1–2–1 ($B_{1-2-1} = 1 \cdot 10^{-6}$), когда ряды динамики всех трех изучаемых показателей описываются параболой второй степени. Далее идет предиктор 4–3–3 (1-я функция Торнквиста, линейно-гиперболическая и линейно-логарифмическая функции), но степень его разбалансированности в 16 раз выше, чем у варианта параболического прогноза. Отметим, что в десятку лучших попал также и тривиальный предиктор 5–4–4, но его точность по критерию (3) в 398 раз ниже, чем у предиктора 1–2–1.

Возникает вопрос: не является ли высокая согласованность прогнозов, полученных по параболическому предиктору, следствием того, что для экстраполяции всех пере-

менных использовалась одна и та же форма трендов? Для ответа на него рассмотрим другие "одноименные" предикторы и их разбаланс.

Вариант 2-4-2 (экспоненциальный предиктор) был рассчитан первоначально, но не попал в десятку лучших. Его разбаланс $5250 \cdot 10^{-6}$. Линейный и авторегрессионный предикторы были построены специально, и их разбаланс составил соответственно $1258 \cdot 10^{-6}$ и $547 \cdot 10^{-6}$. Поэтому на поставленный вопрос следует дать отрицательный ответ. По крайней мере данный пример не подтверждает выдвинутую гипотезу. По-видимому, дело не в простом совпадении форм трендов, а в таком их сочетании, которое отражало бы внутренний механизм формирования изучаемых социально-экономических показателей.

Анализируя коэффициенты уравнений, нельзя не отметить между ними очевидную взаимосвязь

$$\begin{aligned} \hat{y}_{1t} &= 27,51611 - 0,89099t + 0,01980t^2 \\ \hat{y}_{2t} &= 7,08943 + 0,05145t + 0,00165t^2 \\ \hat{y}_{3t} &= 20,42897 - 0,94275t + 0,01815t^2, \end{aligned} \quad (6)$$

т.е. коэффициенты тренда, описывающего динамику естественного прироста населения, приближенно равны разности соответствующих коэффициентов трендов, описывающих изменение рождаемости и смертности. В самом деле, расхождения составляют: $\Delta a_0 = 0,00299$; $\Delta a_1 = 0,00031$; $\Delta a_2 = 0$. Отсюда такая высокая сбалансированность выравненных значений переменных и для точек периода упреждения.

По нашему мнению, глубокий минимум критерия (3) и выполнение балансового соотношения между самими коэффициентами уравнений системы (6) свидетельствуют о проявлении сложившейся в последние годы параболической закономерности развития рождаемости, смертности и естественного прироста населения СССР. Эта закономерность за 1955–1989 гг. характеризуется среднегодовым ускорением роста рождаемости, равным $0,0396\%$, смертности – $0,0033$, естественного прироста – $0,0363\%$.

Для проверки устойчивости выявленной закономерности и выбора наилучшего предиктора, способного успешно решить поставленную задачу прогнозирования, был использован итеративный метод. Сущность его – в образовании новых усеченных рядов динамики изучаемых показателей путем исключения из исходного ряда k , $k = 1, \dots, n/2$, первых точек. По каждому усеченному ряду строятся трендовые модели, указанные в табл. 2, рассчитываются прогнозные значения всех трех показателей и определяются значения критерия баланса переменных для десяти предикторов табл. 3. Иными словами, исследуется поведение показателя разбаланса десяти наилучших предикторов при постепенном сокращении длины периода предыстории.

Расчеты свидетельствуют о достаточно устойчивой сбалансированности прогнозов, полученных по параболическим трендам, на фоне общего ухудшения качества (в смысле критерия (3)) всех остальных предикторов. На отдельных усеченных рядах значение разбаланса для параболического предиктора оказалось ниже, чем на исходном ряду динамики. При $k = 14$ наблюдалось наименьшее значение критерия B_j (табл. 4).

Сравнение данных табл. 3 и 4 показывает, что разбаланс всех вариантов, кроме параболического, увеличился более чем на порядок. Это свидетельствует о их непригодности для решения поставленной задачи. В то же время параболический предиктор, освободившись от груза старой информации начальных точек ряда динамики, адаптировался к изменениям последних лет в показателях рождаемости, смертности и естественного прироста населения. В результате сбалансированность прогнозов улучшилась и есть все основания полагать, что новый предиктор обеспечит более точное предсказание будущего по сравнению с исходным.

Таким образом, предлагаемый подход открывает возможность выбирать наилучший предиктор не только в пространстве опорных функций, но и на множестве точек периода предыстории. Наименьшее значение критерия B_j , найденное по рядам дина-

Расчетные значения критерия баланса переменных

Вариант	Предиктор									
	1-2-1	3-1-2	3-4-2	4-1-4	4-2-2	4-3-3	4-4-3	4-4-4	5-3-4	5-4-4
$B_j \cdot 10^6$	1	27	262	46	142	16	174	273	221	398

Таблица 4

Расчетные значения критерия баланса переменных по усеченным рядам динамики за 1969–1989 гг.

Вариант	Предиктор									
	1-2-1	3-1-2	3-4-2	4-1-4	4-2-2	4-3-3	4-4-3	4-4-4	5-3-4	5-4-4
$B_j \cdot 10^5$	0,027	54	872	301	567	212	1526	3214	2517	2903

мики различной длины, указывает на оптимальный вариант решения поставленной задачи.

Трендовые модели, построенные по данным 1969–1989 гг., имеют вид

$$\begin{aligned} \hat{y}_{1t} &= 16,75212 + 0,25168t - 0,00683t^2 \\ \hat{y}_{2t} &= 7,39856 + 0,34781t - 0,01051t^2 \\ \hat{y}_{3t} &= 9,36382 - 0,09930t + 0,00381t^2 \end{aligned} \quad (7)$$

Так же как и в исходной системе (6), между коэффициентами уравнений (7) наблюдается приближенный баланс. Расхождения составляют: $\Delta a_0 = 0,01026$; $\Delta a_1 = 0,00317$; $\Delta a_2 = 0,00013$.

В пользу предиктора (7) говорит еще один важный факт. Монотонно возрастающие прогнозные значения всех трех показателей, полученных по системе (6) (см. табл. 2), плохо согласуются с последними фактическими уровнями рождаемости и естественного прироста населения, которые, как уже отмечалось, сейчас снижаются. Предиктор (7) в результате адаптации к последним тенденциям развития отличается от предиктора (6) знаками коэффициентов a_2 , что позволяет получить не противоречащий здравому смыслу прогноз изучаемых показателей (табл. 5).

Данные табл. 5 могут служить полезным ориентиром при проведении плановых расчетов в области демографии, экономики, социологии. Так, к 1994 г. рождаемость в нашей стране следует ожидать в пределах 17,0–19,3, смертность – 8,5–10,1, естественный прирост – 7,4–10,3 ‰.

Согласно уравнениям (7), за 1969–1989 гг. динамика рождаемости и смертности в бывшем СССР характеризуется среднегодовым замедлением роста на 0,0137 и 0,0210 ‰ и ускорением естественного прироста населения на 0,0076 ‰. По мере возникновения свежей информации, проявления новых тенденций в изменении изучаемых показателей с целью уточнения указанных параметров, а также прогнозных значений переменных необходимо пересчитывать коэффициенты параболических трендов, проверяя соблюдение балансовых соотношений (3) и (7).

Определенный интерес представляет сопоставление статистических и балансовых характеристик предикторов (6) и (7), представленных в табл. 6, в которой наряду со средней ошибкой аппроксимации E и коэффициентом детерминации R^2 приводятся также остаточная дисперсия $S_{ост}^2$ и расчетное значение F -критерия.

Таблица 5

Точечный и интервальный прогнозы показателей рождаемости, смертности и естественного прироста населения бывшего СССР, ‰ (1 - α = 0,95)

Показатель	Годы				
	1990	1991	1992	1993	1994
Рождаемость	18,981 ± 1,208	18,925 ± 1,295	18,855 ± 1,402	18,772 ± 1,528	18,675 ± 1,674
Смертность	9,962 ± 0,556	9,837 ± 0,596	9,690 ± 0,645	9,523 ± 0,704	9,335 ± 0,770
Естественный прирост	9,022 ± 1,382	9,094 ± 1,481	9,173 ± 1,602	9,261 ± 1,747	9,355 ± 1,914

Таблица 6

Статистические и балансовые характеристики параболических предикторов

Годы	Показатель	Критерий				
		$\bar{E}$	R^2	$S_{\text{ост}}^2$	$F_{\text{расч}}$	$B_j \cdot 10^6$
1955–1989	Рождаемость	4,91	0,830	1,326	78,4	1,0
	Смертность	5,08	0,833	0,256	79,6	
	Естественный прирост	6,94	0,915	1,047	172,3	
1969–1989	Рождаемость	2,11	0,609	0,274	14,0	0,27
	Смертность	2,00	0,914	0,058	95,5	
	Естественный прирост	4,82	0,871	0,358	6,1	

Анализ табл. 6 показывает, что критерии $\bar{E}$, $S_{\text{ост}}^2$ и B_j согласованно отразили общее улучшение статистических характеристик параболических трендов при сокращении длины периода предыстории. Что касается R^2 и F , то они отметили повышение точности и надежности аппроксимации только для уравнения смертности. В отношении трендов рождаемости и естественного прироста зафиксировано ухудшение качества параболических моделей. Причем расчетные значения F -критерия для уравнений этих показателей резко снизились, приблизившись к табличному значению ($F_{0,05; 2; 19} = 3,52$).

Дальнейшее сокращение периода предыстории ($k = 15, 16, 17$) привело к построению предикторов, в которых тренды рождаемости и смертности оказались статистически незначимыми по F -критерию. Разбаланс прогнозов при этом возрос незначительно, по-прежнему указывая на наличие параболической закономерности в изменении рождаемости, смертности и естественного прироста населения нашей страны.

Разное поведение во времени статистических и балансовых показателей исследуемых предикторов свидетельствует о более высокой устойчивости критерия B_j по сравнению с остальными. Особенно она проявляется на коротких рядах динамики, когда мощность статистических критериев резко падает и они часто допускают ошибку второго рода. Являясь свободным от распределения критерием, B_j слабее реагирует на сокращение исходной выборки, сохраняет способность четко различать испытываемые гипотезы относительно форм трендов.

Указанное свойство открывает перед критерием баланса переменных важную область применения, в которой он обладает определенными преимуществами по сравнению с традиционным подходом. Такой областью является, на наш взгляд, изучение

динамики и прогнозирование взаимосвязанных показателей по коротким периодам предыстории. Последние, как известно, преобладают в социально-экономических исследованиях. Типична ситуация, когда экспериментатор располагает сравнительно однородными рядами динамики длиной всего 10–15 точек. В этом случае принцип баланса переменных и построенный на нем критерий (3) могут оказаться практически незаменимыми при решении задач прогнозирования социально-экономических показателей. Что касается средних и длинных рядов динамики, то здесь преимущества критерия баланса переменных по сравнению со статистическими критериями не столь очевидны и требуют дополнительных сравнительных исследований на разнообразном эмпирическом материале.

ЛИТЕРАТУРА

1. *Четыркин Е.М.* Статистические методы прогнозирования. М.: Статистика, 1977.
2. *Ивахненко А.Г., Мюллер Й.А.* Самоорганизация прогнозирующих моделей. Киев: Техніка, 1985.
3. *Ивахненко А.Г., Дукарский О.М., Браверман В.Я.* Модели и методы управления энергетическим строительством. Киев: Будівельник, 1985.
4. *Колузанов К.В., Янковой А.Г.* Принцип баланса переменных в задачах долгосрочного прогнозирования взаимосвязанных социально-экономических показателей // Модели и методы оптимального планирования в сельском хозяйстве. Одесса: ОСХИ. 1982.

Поступила в редакцию
31 III 1989