

## Выбор метода прогнозирования основных статистических показателей работы ГБУЗ «НИИ СП им. Н.В. Склифосовского Департамента здравоохранения города Москвы»

**Б.Л. Курилин\***, **В.Я. Киселевская-Бабинина**, **Н.А. Карасёв**, **И.В. Киселевская-Бабинина**,  
**Е.В. Кислухина**, **В.А. Васильев**

Лаборатория организации стационарной неотложной помощи  
ГБУЗ «Научно-исследовательский институт скорой помощи им. Н.В. Склифосовского ДЗМ»  
Российская Федерация, 129090, Москва, Б. Сухаревская пл., д. 3

\* Контактная информация: Курилин Борис Леонидович, научный сотрудник лаборатории организации стационарной неотложной помощи ГБУЗ «НИИ СП им. Н.В. Склифосовского ДЗМ». Email: kbl11@yandex.ru

### АКТУАЛЬНОСТЬ

Важнейшей частью социально-экономической политики государства является оптимизация системы здравоохранения, ошибки и сбои в работе которой приводят к потере общественного здоровья и экономическому ущербу. На этом фоне прогнозирование работы лечебных учреждений является основой для успешного развития здравоохранения, учитывая однако, что в настоящее время система здравоохранения, показатели и нормативы медико-социального обеспечения все еще нестабильны, а также не отработана четкая стратегия развития на ближайшие и отдаленные сроки.

### ЦЕЛЬ ИССЛЕДОВАНИЯ

Определение оптимального метода прогнозирования работы медицинского учреждения, основанного на выделении главных тенденций временного ряда при построении модели зависимости параметров или определении поведения данных как стохастического ряда (т.е. моделирование случайных процессов и случайных событий с некоторой случайной ошибкой).

### МАТЕРИАЛ И МЕТОДЫ

Для прогнозирования основных статистических показателей работы НИИ СП им. Н.В. Склифосовского на основании ретроспективного анализа использовались данные, представленные в городское бюро медицинской статистики и внесенные в официальные формы отчетности (форма № 30. Утв. пост. Госкомстата РФ от 10.09.02 № 175): количество госпитализированных больных и показатели летальности по стационарным и реанимационным отделениям.

Для выбора оптимальной методики экспериментальной модели прогноза были использованы статистические данные по больнице за период с 1991 по 2016 г. Показатели 2017 года были приняты за контрольные значения.

### РЕЗУЛЬТАТЫ

В результате сравнения нескольких методов (метод сглаживания по скользящей средней, метод наименьших квадратов, модель Брауна, метод Хольта–Уинтерса, модель автокорреляции, метод Бокса–Дженкинса) применительно к работе ГБУЗ «НИИ СП им. Н.В. Склифосовского ДЗМ» была выбрана модель Хольта–Уинтерса как наиболее отвечающая характеристикам данных.

### ВЫВОДЫ

1. При использовании методов сглаживания по скользящей средней, наименьших квадратов, Бокса–Дженкинса, а также моделей Брауна и автокорреляции на результат прогноза влияют не всегда строго прямолинейные показатели динамического ряда в силу неоднородности динамических рядов и наличия так называемых выскакивающих величин (особенно часто выделяющихся в медицинском учреждении, оказывающем неотложную помощь), которые приводят к значительному снижению достоверности прогнозирования.

2. Применение модели Хольта–Уинтерса, которая учитывает экспоненциальный тренд (тенденция изменения показателей временного ряда) и аддитивную сезонность (периодические колебания, наблюдаемые на временных рядах), является наиболее целесообразным для обработки статистических данных и прогнозирования на долгосрочные, среднесрочные и краткосрочные периоды с учетом специфики стационара, оказывающего неотложную помощь.

3. Выбор оптимального метода прогнозирования работы медицинского учреждения, основанного на выделении главных тенденций временного ряда с учетом большинства особенностей при моделировании случайных процессов и событий, позволил уменьшить относительную ошибку прогноза.

### Ключевые слова:

прогнозирование, выравнивание динамических рядов, оценка достоверности различия показателей

### Для цитирования

Курилин Б.Л., Киселевская-Бабинина В.Я., Карасёв Н.А., Киселевская-Бабинина И.В., Кислухина Е.В., Васильев В.А. Выбор метода прогнозирования основных статистических показателей работы ГБУЗ «НИИ СП им. Н.В. Склифосовского Департамента здравоохранения города Москвы». Журнал им. Н.В. Склифосовского Неотложная медицинская помощь. 2019;8(3):246–256. <https://doi.org/10.23934/2223-9022-2019-8-3-246-256>

### Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов

### Благодарности

Исследование не имеет спонсорской поддержки

AR — авторегрессивная модель

ARIMA — интегрированная модель авторегрессии — скользящего среднего

ARMA — модель авторегрессии — скользящего среднего

MA — модель автокорреляции

## АКТУАЛЬНОСТЬ

Прогнозирование основных статистических показателей работы в медицинском учреждении, оказывающем неотложную помощь — это комплекс аргументированных предположений (выраженных в качественной и количественной формах) относительно структуры оказания медицинской помощи в перспективе, позволяющий правильно и полно предугадывать события на основании факторов и тенденций для принятия соответствующих административных решений, разработки адекватных превентивных мер, которые дадут возможность избежать нежелательных результатов в оказании медицинской помощи больным [1].

Поскольку прогноз работы медицинского учреждения строится на вероятностном развитии событий, то, согласно Федеральным законам (ФЗ) от 20.07.1995 г. N 115-ФЗ «О государственном прогнозировании и программах социально-экономического развития РФ», ФЗ от 21.07.2011 г. N 323-ФЗ «Об основах охраны здоровья граждан в РФ», ФЗ от 29.11.2010 г. N 326-ФЗ «Об обязательном медицинском страховании в Российской Федерации», в зависимости от поставленных задач для построения математических моделей в целях оптимизации процесса оказания медицинской помощи, а также планирования различных методов (в том числе материально-технического обеспечения и учета возможных экономических затрат) необходима исчерпывающая ретроспективная информация на основании статистических показателей работы за значительный (не менее 10–12 лет) период наблюдения [2].

Одно из перспективных направлений развития прогнозирования связано с адаптивными методами. Эти методы позволяют строить самокорректирующиеся модели, способные оперативно реагировать на изменение условий [3]. Адаптивные методы учитывают различную информационную ценность уровней ряда и «старение» информации. Все это делает эффективным их применение для прогнозирования неустойчивых рядов с изменяющейся тенденцией. В адаптивных методах различную ценность уровней в зависимости от их «свежести» можно учесть с помощью системы весов, придаваемых этим уровням.

Многие из базовых методов прогнозирования относятся скорее к отдельным приемам или процедурам, другие являются пакетами методов и отличаются друг от друга количеством частных приемов и/или последовательностью их применения [4].

По степени формализации все методы прогнозирования делятся на интуитивные и формализованные (рис. 1). Интуитивное прогнозирование применяется тогда, когда объект настолько сложен, что аналитически учесть влияние многих факторов практически невозможно. Формализованные же прогнозы строятся при помощи методов вычислительной математики и позволяют получить максимально достоверные данные в более короткий срок.

При выборе методов прогнозирования важным показателем является глубина упреждения прогноза. При этом необходимо не только знать абсолютную величину этого показателя, но и отнести его к длительности эволюционного цикла развития объекта прогнозирования.

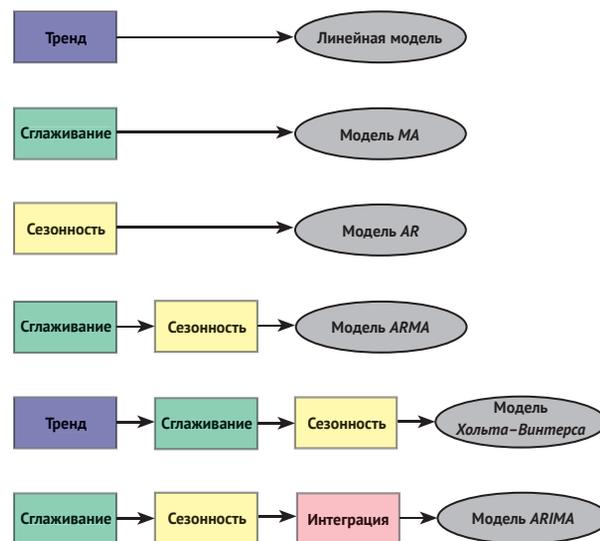


Рис. 1. Схема распределения основных приемов и методов  
Примечания: AR — авторегрессивная модель; ARIMA — интегрированная модель авторегрессии — скользящего среднего; ARMA — модель авторегрессии — скользящего среднего; MA — модель автокорреляции

Fig. 1. The distribution of basic techniques and methods  
Notes: AR — autoregressive model; ARIMA — autoregressive integrated moving average model; ARMA — autoregressive moving average model; MA — moving-average model

## МАТЕРИАЛ И МЕТОДЫ ИССЛЕДОВАНИЯ

Для прогнозирования на основании ретроспективного анализа основных статистических показателей работы НИИ СП им. Н.В. Склифосовского за период с 1991 по 2017 г. были использованы данные, представленные в городское бюро медицинской статистики и внесенные в официальные формы отчетности (форма № 30. Утв. пост. Госкомстата РФ от 10.09.02 № 175): средний день работы койки в году, оборот койки, средний койко-день, показатели летальности по стационарным и реанимационным отделениям. Для выбора оптимальной методики экспериментальной модели прогноза были использованы данные за период с 1991 по 2016 г. Показатели 2017 года в этом случае были приняты за контрольные значения.

Для экспериментальной модели прогноза основных показателей, характеризующих работу ГБУЗ «НИИ СП им. Н.В. Склифосовского ДЗМ», были выбраны несколько методов, основанных на выравнивании динамических рядов:

- метод сглаживания по скользящей средней;
- метод наименьших квадратов;
- модель Брауна;
- метод Хольта–Уинтерса;
- модель автокорреляции;
- метод Бокса–Дженкинса.

Применение этих методов к данным осуществлялось при помощи свободной программной среды R v3.5.1 (*The R Project for Statistical Computing*).

## ОЦЕНКА ДОСТОВЕРНОСТИ МЕТОДОВ

При использовании методов прогнозирования крайне важна оценка статистической значимости полученной модели и ее прогноза. Таковыми могут являться коэффициент детерминации  $R^2$ , ошибка первого рода  $p$ -value и критерий Акаике AIC.

Коэффициент детерминации  $R^2$  рассчитывается по формуле:

$$R^2 = 1 - \frac{\sigma^2(y|x)}{\sigma^2(y)},$$

где  $\sigma^2(y|x)$  — условная дисперсия,  $\sigma^2(y)$  — дисперсия модельных данных. Полученная величина варьирует от 0 до 1 и рассматривается как тот процент реальных данных, который описывает полученная модель. Соответственно, чтобы признать модель статистически значимой, коэффициент детерминации должен быть не ниже 0,5, а для очень хороших моделей он должен превышать 0,8.

Величина  $p$ -value фактически означает вероятность отклонения гипотезы при ее правильности, то есть (применительно к прогнозным моделям) вероятность реальных данных в дальнейшем отклонится от прогноза. Чем меньше величина  $p$ -value, тем более статистически значимы результаты. Традиционно уровень  $p$ -value устанавливается от 5% и меньше. В расчете на более точный прогноз уровень может устанавливаться от 0,5% и меньше.

Критерий Акаике был предложен в качестве альтернативы коэффициенту детерминации.

$$AIC = 2k - 2\ln(L_{max}),$$

где  $k$  — число параметров модели,  $L_{max}$  — максимум функции правдоподобия модели. Как таковой критерий Акаике не может описывать статистическую значимость и не может быть интерпретирован, однако в каждом конкретном случае он вычисляет каждому прогнозу его «относительный вес»; при их сравнении прогноз считается лучшим, если для него значение  $AIC$  оказалось наименьшим.

Для более полного сравнения мы будем вычислять эти три оценки и сравнивать различные методы с их помощью.

#### ЛИНЕЙНЫЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

Временной ряд, для которого делается прогноз, представляет из себя набор значений  $\{y\}_{i=1}^n$  для нескольких точек  $\{x\}_{i=1}^n$ . Простейший метод прогноза заключается в предположении, что  $y$  имеет формульную зависимость от  $x$  с неизвестными параметрами  $\alpha$ :  $y_i = y(\alpha, x_i)$ .

В основе метода наименьших квадратов лежит то, что наилучшим приближением к реальным данным будет минимизация суммы квадратов разностей между фактическим уровнем и теоретическим:

$$\sum_i (y_i - \bar{y}(\alpha, x_i))^2 \rightarrow \min_{\alpha}. \quad (\text{Формула 1})$$

Метод наименьших квадратов — основной метод для определения параметров различных моделей и часто применяется в регрессионном анализе и прогнозировании. В дальнейшем мы будем на него ссылаться при описании других моделей прогнозирования [5–8].

Вид функции  $\bar{y}(\alpha, x)$  выбирается с помощью экспертных оценок — это может быть линейная зависимость, квадратичная, экспоненциальная и т.п. В простейшем случае выбирается линейная зависимость:

$$\bar{y}_i(\alpha, x) = \alpha_0 + \alpha_1 x,$$

где  $\bar{y}$  — выравненные уровни,  $\alpha_0$  — начальный уровень ряда;  $\alpha_1$  — начальная скорость ряда;  $x$  — интервал времени. Таким образом, из временного ряда выделяется его основной тренд — при  $\alpha_1 < 0$  величины уменьшаются с увеличением времени  $x$ , а при  $\alpha_1 > 0$  — увеличиваются.

В общем случае коэффициенты  $\alpha$  вычисляются по формулам:

$$\alpha_0 = \frac{\sum y_i}{n} - \frac{\alpha_1 \sum x_i}{n}, \quad \alpha_1 = \frac{\sum y_i x_i - \frac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}.$$

При условии, что сумма интервалов времени равна нулю ( $\sum x_i = 0$ ), значения  $\alpha_0$  и  $\alpha_1$  вычисляются по формулам:

$$\alpha_0 = \frac{\sum y_i}{n}, \quad \alpha_1 = \frac{\sum y_i x_i}{\sum x_i^2},$$

где  $y$  — уровни ряда,  $n$  — количество уровней ряда.

Средний прогностический показатель определяет при продолжении расчетов  $y_{теор.}$ . Именно  $y_{теор.}$  будущего года является средним прогностическим показателем ( $y_{ср. прогн.}$ ). При расчете по формуле:

$$y_{ср. прогн.} = y_{ср.} + (\alpha_1 x).$$

Величина  $x$  принимает следующее значение с тем же шагом, что и ранее (то есть к последнему значению  $x$  прибавляется 1 или 2).

Для определения пограничных значений среднего прогностического показателя необходимо рассчитать его максимальное и минимальное значения ( $y_{max прогн.}$ ) и ( $y_{min прогн.}$ ). Для расчета  $y_{max прогн.}$  и  $y_{min прогн.}$  использовались значения  $\sum \Delta+$  и  $\sum \Delta-$ . При этом:

$y_{max прогн.}$  где  $\Delta_{ср.}$  — среднее отклонение вверх от линии тенденции, а  $n$  — число лет выше линии тенденции.

$y_{min прогн.}$  где  $\Delta_{ср.}$  — среднее отклонение вниз от линии тенденции, а  $n$  — число лет ниже линии тенденции.

Для построения линейной зависимости применительно к нашим данным использовалась функция  $lm$  из стандартного пакета *stats* программы *R*. Линейная модель позволяет четко увидеть линию тренда у данных — на протяжении 15 лет общее число поступлений и число поступлений в реанимацию росло, в то время как число смертей и летальность падали. Однако если в случае с общей летальностью и числом умерших в стационаре реальные данные хорошо отвечают линейному тренду, то в случае с числом поступлений и числом умерших случайные выбросы затрудняют прогнозирование, и достоверность прогноза сильно падает (рис. 2).

К сожалению, ни одно реальное значение за 2017 г. не попало в пределы вычисленных пограничных значений, что обуславливается чересчур резкими скачками. Доверительные интервалы получаются слишком большими, чтобы свидетельствовать о статистической значимой достоверности прогноза. Результаты оценки моделей приведены в табл. 1.

Как видно из табл. 1, наилучший результат достигнут для общей летальности из-за большого сходства поведения данных с прямой. Число поступлений и число смертей в стационаре описываются прямой частично, поскольку до 2000 г. для летальности и после 2010 г. для поступлений наблюдается нестационарное поведение. Самый же недостоверный прогноз у ежегодного числа умерших в реанимации пациентов — слишком высокие скачки и непредсказуемые выбросы — не дают провести прямую, и даже тенденция данных к стойкому снижению или подъему остается под вопросом.

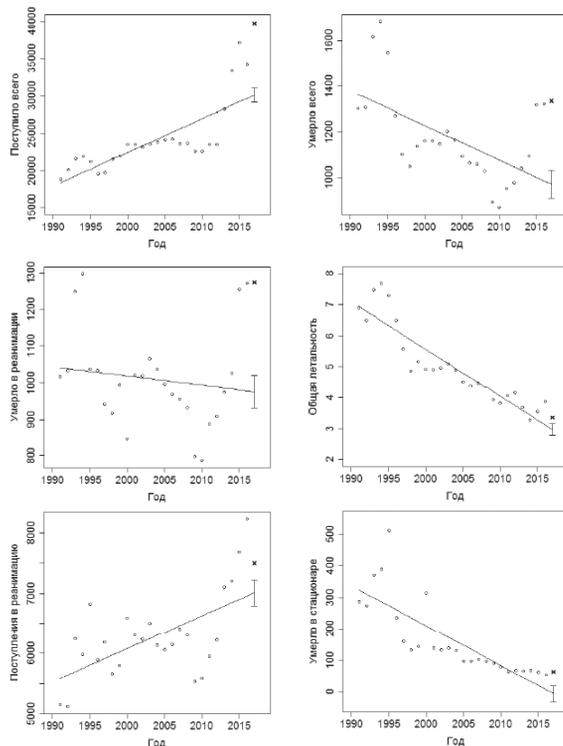


Рис. 2. Линейные модели для данных по обращениям, умершим в стационаре, умершим в реанимации, общей летальности, поступившим в реанимацию, умершим в отделениях. Точками показаны реальные данные, крестиками — значения за 2017 г.

Fig. 2. Linear models for data on admissions, in-hospital deaths, intensive care deaths, general mortality, admission to intensive care, deaths in various departments. Dots show real data, crosses show values for 2017

Таблица 1  
Характеристики линейных моделей для каждого набора данных

Table 1  
Characteristics of linear models for each data set

Прогнозная величина	$R^2$	AIC
Всего поступивших	0,58	491,9
Всего умерших	0,29	345,6
Умершие в реанимации	0,02	332,6
Общая летальность	0,83	44,2
Поступившие в реанимацию	0,32	408,5
Умершие в стационаре	0,63	301,0

Таким образом, плюсы метода наименьших квадратов линейной модели применительно к работе медицинского учреждения заключаются в возможности определения тенденции развития к росту или снижению с определением ее выраженности, но основанный на нем прогноз имеет недостаточную точность.

#### МОДЕЛЬ СКОЛЬЗЯЩЕГО СРЕДНЕГО (МА)

Этот метод хорошо работает, если данные представляют собой сумму  $q$  белых шумов:

$$y_t = \sum_{j=0}^q b_j \varepsilon_{t-j}, \quad (\text{Формула 2})$$

где  $x$  — члены ряда,  $b$  — параметры модели,  $\varepsilon$  — белый шум. То есть последующие значения временно-го ряда накапливают ошибки предыдущих.

Уравнение первого порядка можно переписать по-другому:

$$y_t = \varepsilon_t - \sum_{j=1}^q b_j y_{t-j}.$$

То есть каждое последующее значение опирается на предыдущие, причем все с весами, стремящимися к нулю.

Сущность метода состоит в том, что фактические показатели динамического ряда заменяют на теоретически ожидаемые путем осреднения двух или трех рядом стоящих фактических показателей. Таким образом, из данных выделяется явный тренд. Чем больше окно, тем больше значимость получаемого тренда, однако мелкие изменения стираются в процессе.

Простая скользящая средняя — расчет по  $2n$  точкам с помощью среднего арифметического:

$$\bar{y}_t = \frac{1}{2n} \sum_{i=t-n}^{t+n} y_i.$$

Скользящая средняя взвешенная — расчет по  $n$  точкам с учетом большей или меньшей значимости отдельных значений на основе арифметической прогрессии:

$$\bar{y}_t = \frac{2}{n(n+1)} \sum_{i=0}^{n-1} (n-i) y_{t-i}.$$

При расчете скользящей средней для всех периодов, прогноз строится на один период по формуле:

$$y_{t+1} = m_{t-1} + \frac{1}{n} (y_t - y_{t-1}),$$

где  $t+1$  — прогнозный период;  $t$  — период, предшествующий прогнозному периоду (год, месяц и т.д.);  $y_{t+1}$  — прогнозируемый показатель;  $m_{t-1}$  — скользящая средняя за два периода до прогнозного;  $n$  — число уровней, входящих в интервал сглаживания;  $y_t$  — фактическое значение исследуемого явления за предшествующий период;  $y_{t-1}$  — фактическое значение исследуемого явления за два периода, предшествующих прогнозному.

Для построения скользящей средней и ее прогнозных значений применяется функция  $ma$  из пакета *forecast* программы *R*. На рис. 3 приведены результаты применения метода скользящего среднего для статистических данных стационара. На графиках видно, что этот метод дает хорошие прогностические оценки данных по общей летальности и умершим в отделениях. Однако число обращений и умерших в стационаре прогнозируются хуже. Из табл. 2 видно, что по сравнению с линейным методом наименьших квадратов, метод скользящего среднего более статистически значим, так как его коэффициент детерминации выше у всех прогнозных величин.

Поступления в реанимацию и число умерших в реанимациях не могут быть прогнозированы с хорошей точностью из-за получившейся разницы в реальных и прогнозных значениях.

#### МОДЕЛЬ БРАУНА

В модели Брауна, также известной как метод экспоненциального сглаживания, динамический ряд сглаживается с помощью взвешенной скользящей средней, в которой значения подчиняются экспоненциальному закону.

Идея метода заключается в том, что прогнозное значение  $\bar{y}_{t+1}$  определяется через предыдущее спрогнозированное значение  $\bar{y}_t$ , но скорректированное с

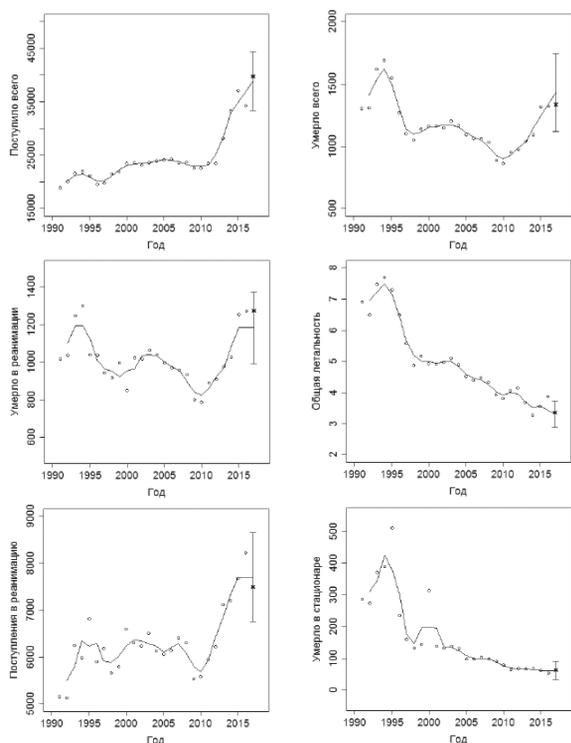


Рис. 3. Сглаживание по скользящей средней для данных по обращениям, умершим в стационаре, умершим в реанимации, общей летальности, поступившим в реанимацию, умершим в отделениях. Точками показаны реальные данные, крестиками — значения за 2017 г.  
 Fig. 3. Smoothing by the moving average for data on admissions, in-hospital deaths, intensive care deaths, general mortality, admission to intensive care, deaths in various departments. Dots show real data, crosses show values for 2017

Таблица 2  
**Характеристики моделей скользящего среднего для каждого набора данных**  
 Table 2  
**Characteristics of the moving average models for each data set**

Прогнозная величина	R <sup>2</sup>	AIC
Всего поступивших	0,96	402,4
Всего умерших	0,94	278,1
Умершие в реанимации	0,94	274,7
Общая летальность	0,96	9,5
Поступившие в реанимацию	0,95	349,6
Умершие в стационаре	0,83	233,5

некоторым коэффициентом на величину отклонения факта  $y_t$  от прогноза:  $\hat{y}_{t+1} = \hat{y}_t + \alpha(y_t - \hat{y}_t)$ . Достаточно часто эту модель представляют в другом виде:  $\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)\hat{y}_t$ , однако смысл модели от этого не меняется: она в той или иной степени (в зависимости от значения коэффициента  $\alpha$ ) адаптируется к новой поступающей информации.

Метод экспоненциального сглаживания аналогичен методу скользящего среднего. У них есть общий главный принцип — каждая точка зависит от значений соседних с некоторыми весами. Главные отличия метода Брауна — начальная точка не затрагивается и остается неизменной, по мере удаления от начальных данных веса экспоненциально стремятся к нулю, кроме того, данные сглаживаются не в центре, а к ближайшему предыдущему значению.

Вообще же модель Брауна может применяться в двух случаях:

— когда нужно сгладить имеющийся ряд данных для выявления тенденции (обычно в случае со стационарными процессами с небольшой погрешностью каждого измерения), задается значение  $\alpha$  в пределах от 0 до 1;

— когда нужно сделать краткосрочный прогноз. В таком случае наилучший результат прогноза получается при задании  $\alpha$  в пределах от 0 до 2.

Стандартизированного способа нахождения параметра  $\alpha$  для всех случаев не существует. В зависимости от длины интервала сглаживания  $n$  параметр вычисляется как  $\alpha = 2/(n+1)$ . Оптимальный параметр может быть получен при минимизации ошибок прогноза относительно  $\alpha$  (формула 1).

На рис. 4 показаны результаты прогнозирования статистических данных, полученных в результате применения модели Брауна. В целом можно сказать, что модель хуже прогнозирует данные, чем простой метод скользящего окна. Это может быть объяснено резкими изменениями рассматриваемых показателей, на которые система не смогла отреагировать.

Из табл. 3 видно, что коэффициент детерминации для всех прогнозных величин модели Брауна ниже, чем при применении метода скользящего среднего.

Поскольку в прогнозируемых данных изменения проходят скачкообразно, модель Брауна дает статистически незначительные результаты.

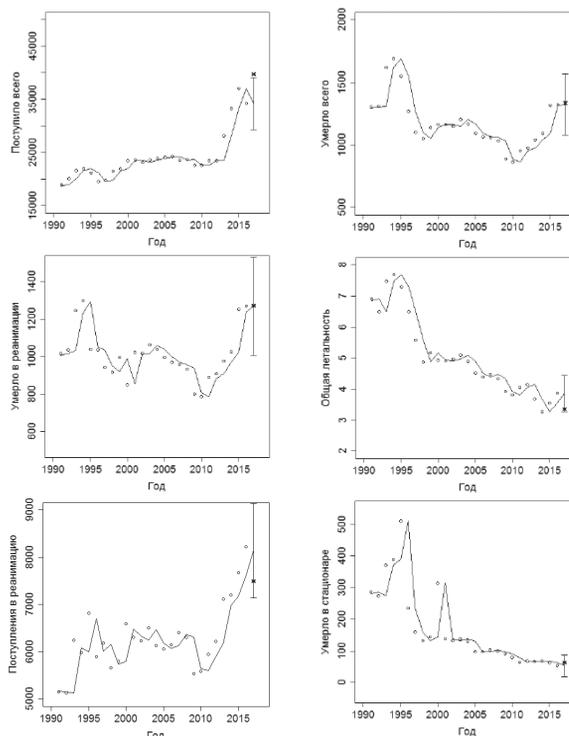


Рис. 4. Экспоненциальное сглаживание для данных по обращениям, умершим в стационаре, умершим в реанимации, общей летальности, поступившим в реанимацию, умершим в отделениях. Точками показаны реальные данные, крестиками — значения за 2017 г.  
 Fig. 4. : Exponential smoothing for data on admissions, in-hospital deaths, intensive care deaths, general mortality, admission to intensive care, deaths in various departments. Dots show real data, crosses show values for 2017

Таблица 3

## Характеристики моделей Брауна для каждого набора данных

Table 3

## The Brown model characteristics for each data set

Прогнозная величина	$R^2$	AIC
Всего поступивших	0,93	474,9
Всего умерших	0,91	332,6
Умершие в реанимации	0,89	330,3
Общая летальность	0,92	39,8
Поступившие в реанимацию	0,49	412,6
Умершие в стационаре	0,67	288,3

## МЕТОД ХОЛЬТА–УИНТЕРСА

Метод Хольта–Уинтерса является модификацией метода экспоненциального сглаживания для сезонных рядов. Этот метод также учитывает экспоненциальный тренд (тенденция изменения показателей временного ряда) и аддитивную сезонность (периодические колебания, наблюдаемые на временных рядах) [9, 10].

Модель может быть в двух видах, в зависимости от сезонной компоненты либо аддитивной, либо мультипликативной.

Пусть задан временной ряд  $y_0, \dots, y_t, y_i \in R$ . Необходимо решить задачу прогнозирования временного ряда:

$$\begin{cases} \hat{y}_{t+d} = (a_t + kr)_t \Theta_{t+k-s}, \\ a_t = \alpha \frac{y_t}{\Theta_{t-s}} + (1 - \alpha)(a_{t-1} + r_{t-1}), \\ r_t = \gamma(a_t - a_{t-1}) + (1 - \gamma)r_{t-1}, \\ \Theta_t = \beta \frac{y_t}{a_t} + (1 - \beta) \Theta_{t-s}; \end{cases}$$

$$\begin{cases} \hat{y}_{t+d} = a_t + kr_t \Theta_{t+k-s}, \\ a_t = \alpha(y_t - \Theta_{t-s}) + (1 - \alpha)(a_{t-1} + r_{t-1}), \\ r_t = \gamma(a_t - a_{t-1}) + (1 - \gamma)r_{t-1}, \\ \Theta_t = \beta \frac{y_t}{a_t} + (1 - \beta) \Theta_{t-s}, \end{cases}$$

где  $s$  — период сезонности,  $\Theta_t, i \in [0, s-1]$  — сезонный профиль,  $r_t$  — параметр тренда,  $a_t$  — параметр прогноза, очищенный от влияния тренда и сезонности. Соответственно коэффициент  $\alpha \in [0, 1]$  указывает, как сильно величины зависят от предыдущих значений,  $\beta \in [0, 1]$  показывает значимость сезонности, а  $\gamma \in [0, 1]$  — есть ли у данных ярко выраженный тренд.

Оптимальные параметры  $\alpha, \beta, \gamma$  предлагается находить экспериментальным путем.

Для построения модели Хольта–Уинтерса применяется функция *Holt-Winters* из стандартного пакета *stats* программы *R*, а для дальнейшего ее прогнозирования — пакет *forecast*. Оптимальные параметры  $\alpha, \beta, \gamma$  находятся путем минимизации квадратов отклонений (формула 1). Однако частоту повторения значений приходится задавать вручную, опираясь на вид данных. Для выявления сезонности предлагалось составить график автокорреляций со значимым уровнем 0,4.

На рис. 5 показаны графики, построенные для исследуемых показателей методом Хольта–Уинтерса. Из них видно, что в отличие от метода скользящего среднего модель не сгладила пики и спады, а включила их в тенденцию.

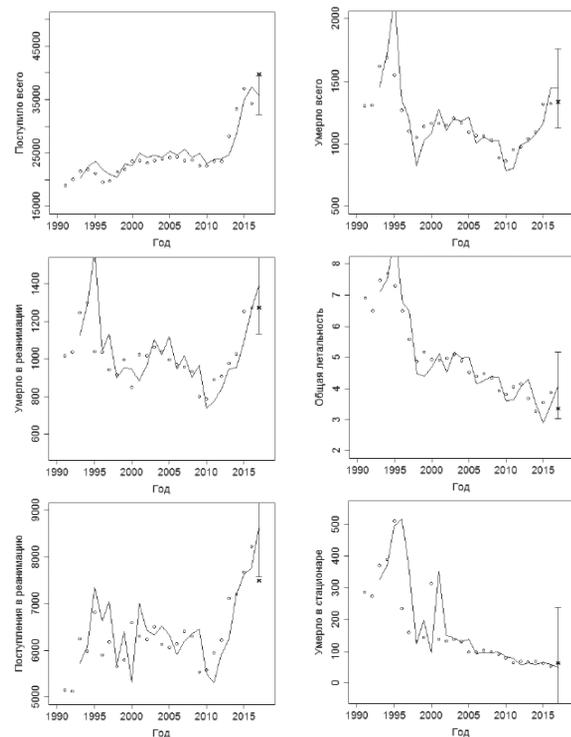


Рис. 5. Модель Хольта–Уинтерса для данных по обращениям, умершим в стационаре, умершим в реанимации, общей летальности, поступившим в реанимацию, умершим в отделениях. Точками показаны реальные данные, крестиками — значения за 2017 г.

Fig. 5. The Holt-Winters model for data on admissions, in-hospital deaths, intensive care deaths, general mortality, admission to intensive care, deaths in various departments. Dots show real data, crosses show values for 2017

Таблица 4

## Характеристики моделей Хольта–Уинтерса для каждого набора данных

Table 4

## Characteristics of Holt-Winters models for each data set

Прогнозная величина	$R^2$	AIC
Всего поступивших	0,93	479,3
Всего умерших	0,89	343,5
Умершие в реанимации	0,88	338,8
Общая летальность	0,91	46,8
Поступившие в реанимацию	0,91	417,6
Умерли в стационаре	0,68	294,6

Как видно из табл. 4, применение метода Хольта–Уинтерса к нашим данным позволило получить хорошие результаты. Все случайные выбросы (открытие, закрытие отделений и реанимаций, перегруз или недогруз и т.д.) сгладились, но сезонные колебания остались, поэтому прогнозные значения наиболее близки к реальным данным.

## АВТОРЕГРЕССИОННАЯ МОДЕЛЬ (AR)

С помощью *AR*-моделей моделируется сезонность временного ряда. Данная модель строится из предположения о том, что каждый член временного ряда образуется при помощи  $p$  предыдущих членов, то есть модель имеет некоторое запаздывание данных:

$$y_t = c + \sum_{i=1}^p a_i y_{t-i} + \varepsilon_t, \quad (\text{Формула 3})$$

где  $y$  — члены ряда,  $c=const$  — постоянный уровень,  $a$  — коэффициенты авторегрессии,  $\varepsilon_t$  — белый шум.

Таким образом, в модели уже заложен прогноз на  $p$  шагов вперед при имеющихся начальных значениях  $\{y_i\}$ . Наиболее рациональным способом определения параметров уравнения авторегрессии является использование метода наименьших квадратов (формула 1).

Однако главная задача сводится к поиску порядка авторегрессии  $p$ .

Результаты применения модели авторегрессии показаны на рис. 6. Поскольку моделирование этим методом предусматривает сохранение предыдущих периодов, при прогнозировании общей летальности и числа умерших в стационаре модель учла первоначальные скачки значений и в дальнейшем не смогла правильно отобразить снижение этих показателей.

Таблица 5 показывает, что данная модель хуже аппроксимирует данные, чем модель Хольта-Уинтерса.

Поскольку наши данные имеют резкие изменения значений, этот метод прогноза оказался малопродуктивным. При равномерном количестве поступлений (до 2012 г.) модель обеспечивала адекватный прогноз, но при резком возрастании числа госпитализаций с 2012 г. прогнозируемые данные стали значительно отличаться от реальных.

**МОДЕЛЬ АВТОРЕГРЕССИИ – СКОЛЬЗЯЩЕГО СРЕДНЕГО (ARMA)**

Как видно из названия, данная модель представляет собой сумму авторегрессии (формула 3) — запаздывания данных — и скользящего окна (формула 2):

$$y_t = c + \sum_{i=1}^p a_i y_{t-i} + \varepsilon_t + \sum_{j=0}^q b_j \varepsilon_{t-j}.$$

Поскольку мы уже оценивали параметры  $p$  и  $q$  для двух последних моделей, для расчетов были взяты те же значения, что и в них. Методом наименьших квадратов оценили коэффициенты при независимых переменных модели.

На рис. 7 и в табл. 6 показаны результаты прогнозирования изучаемых показателей с помощью модели ARMA. Эта модель хорошо подходит для прогнозирования данных, стабильно сохраняющих свою динамику на всем рассматриваемом временном периоде. Поскольку для наших данных это условие не выполняется, величины, спрогнозированные этим методом, получились отличными от реальных.

Полученные методом ARMA прогнозные значения хуже значений, полученных простым методом авторегрессии, поскольку часть реальных данных имеет большую амплитуду, принятую моделью за основу дальнейшего прогнозирования. Это хорошо видно на графике «Умершие в стационаре».

**МОДЕЛЬ БОКСА-ДЖЕНКИНСА (ARIMA)**

Так называемая интегрированная модель авторегрессии — скользящего среднего, расширенная версия модели ARMA. ARIMA строится на предположении о том, что данные имеют авторегрессию (формула 3), шумовой эффект (формула 2) и интеграцию.

$$\Delta^d x_t = c + \sum_{i=1}^p \Delta^d x_{t-i} + \sum_{j=1}^q b_j \varepsilon_{t-j} + \varepsilon_t.$$

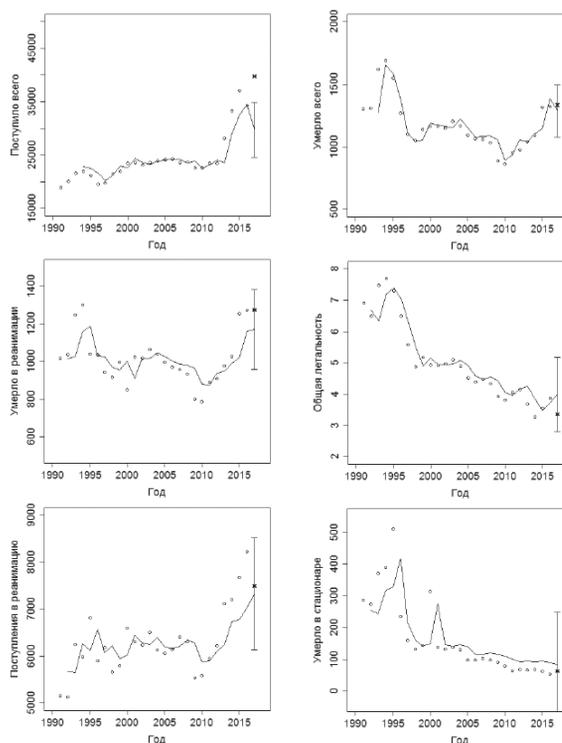


Рис. 6. Модель авторегрессии для данных по обращениям, умершим в стационаре, умершим в реанимации, общей летальности, поступившим в реанимацию, умершим в отделениях. Точками показаны реальные данные, крестиками — значения за 2017 г.

Fig. 6. Autoregressive model for data on admissions, in-hospital deaths, intensive care deaths, general mortality, admission to intensive care, deaths in various departments. Dots show real data, crosses show values for 2017

Таблица 5

**Характеристики авторегрессионных моделей для каждого набора данных**

Table 5

**Characteristics of autoregressive models for each data set**

Прогнозная величина	$R^2$	AIC
Всего поступивших	2 150 786	464,22
Всего умерших	8744	319,5
Умершие в реанимации	9491	318,55
Общая летальность	0,1532	35,61
Поступившие в реанимацию	244 620	403,5
Умершие в стационаре	5919	306,48

Интеграция данных подразумевает наличие стабильной разности некоторого порядка. То есть  $\Delta^d x_t = x_t - x_{t-d} \forall t$  сохраняет свой вид и поведение.

Модель достаточно часто применяется для прогнозирования временных рядов и имеет много улучшений и вариаций. Метод считается достаточно точным, чтобы получать как краткосрочные, так и долгосрочные прогнозы, которые не требуют отдельного оценивания. Однако для построения модели типа ARIMA требуется больший набор данных и всеобъемлющий их анализ [11].

Параметрами модели являются  $p, d, q$ , то есть порядок запаздывания, порядок сохранения остатков и размер скользящего окна.

Результаты прогнозирования методом ARIMA показаны на рис. 8 и в табл. 7. Согласно табл. 7, данный метод оказался статистически слабо значимым из всех

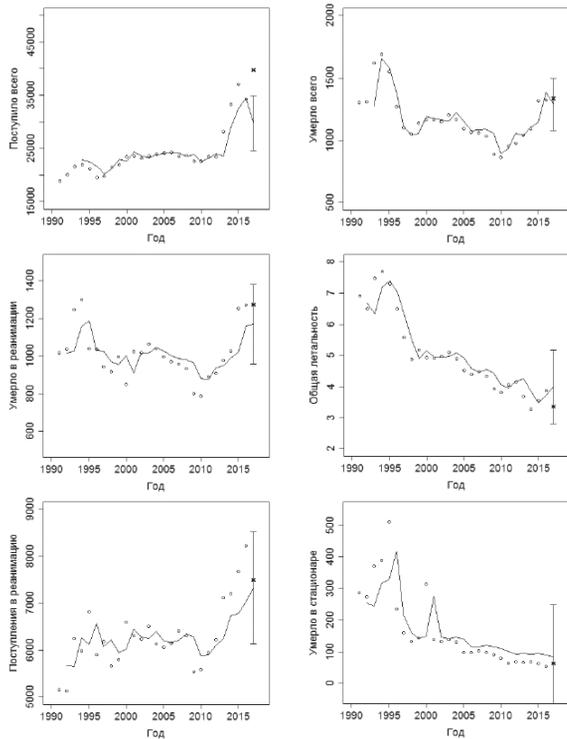


Рис. 7. Модель авторегрессии и скользящего среднего для данных по обращениям, умершим в стационаре, умершим в реанимации, общей летальности, поступившим в реанимацию, умершим в отделениях. Точками показаны реальные данные, крестиками — значения за 2017 г.  
Fig. 7. The model of autoregression and moving average for data on admissions, in-hospital deaths, intensive care deaths, general mortality, admission to intensive care, deaths in various departments. Dots show real data, crosses show values for 2017

Таблица 6  
Характеристики авторегрессионных моделей для каждого набора данных

Table 6  
Characteristics of autoregressive models for each data set

Прогнозная величина	$R^2$	AIC
Всего поступивших	1 607 957	466,84
Всего умерших	5606	319,61
Умершие в реанимации	8142	321,59
Общая летальность	0,1331	38,96
Поступившие в реанимацию	213 385	406,43
Умершие в стационаре	4553	308,57

выше рассмотренных, поскольку предусматривает не менее 40 временных точек для прогнозирования. На графиках видно, что смоделированные данные имеют большое отличие от реальных.

Из-за малого количества анализируемых значений применение метода *ARIMA* дало плохие результаты прогнозирования, хотя он считается универсальным, обоснованным и достоверным методом прогнозирования. В дальнейшем мы планируем применить его к большей выборке, разбив данные каждого года по квартально.

## РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

На примере проведенного исследования можно утверждать:

1. **Линейный метод наименьших квадратов** применяется только в случае, если нужно узнать общую

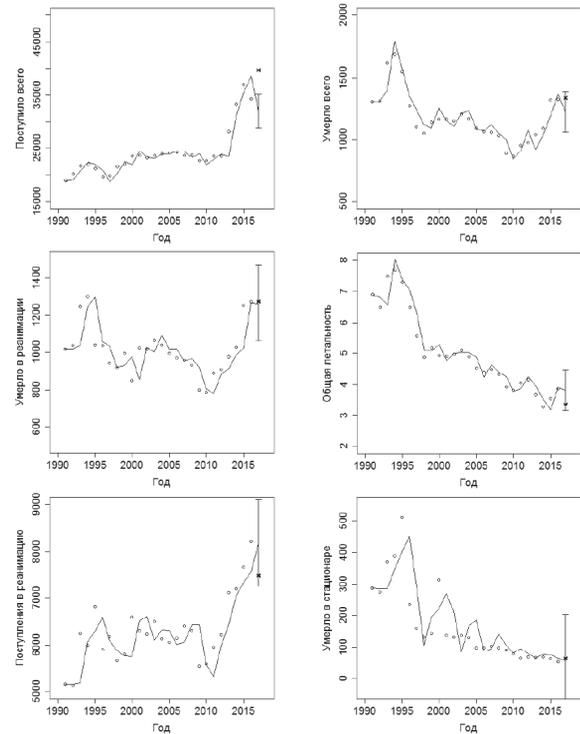


Рис. 8. Модель авторегрессии и скользящего среднего для данных по обращениям, умершим в стационаре, умершим в реанимации, общей летальности, поступившим в реанимацию, умершим в отделениях. Точками показаны реальные данные, крестиками — значения за 2017 г.  
Fig. 8. The model of autoregression and moving average for data on admissions, in-hospital deaths, intensive care deaths, general mortality, admission to intensive care, deaths in various departments. Dots show real data, crosses show values for 2017

Таблица 7  
Характеристики авторегрессионных моделей для каждого набора данных

Table 7  
Characteristics of autoregressive models for each data set

Прогнозная величина	$R^2$	AIC
Всего поступивших	1 843 255	446,71
Всего умерших	6893	306,98
Умершие в реанимации	10767	313,15
Общая летальность	0,1069	31,31
Поступившие в реанимацию	222 159	389
Умершие в стационаре	5466	296,82

долгосрочную тенденцию ряда к росту или снижению. Если данные имеют колебания, то метод не сможет их уловить. Важное преимущество метода наименьших квадратов состоит в том, что он выравняет данные, которые имели некоторую погрешность вычисления, и полученная с его помощью оценка лучше любой другой [12]. Метод используется для решения задач сглаживания данных, интерполяции и экстраполяции.

Таким образом, плюсы метода наименьших квадратов линейной модели применительно к работе медицинского учреждения заключаются в определении тенденции развития к росту или снижению с возможностью оценки ее выраженности, но основанный на нем прогноз имеет недостаточную точность.

2. **Метод скользящего среднего** хорошо подходит для равномерно колеблющихся данных и может дать хорошие результаты даже при неравномерности дан-

ных. Скользящее окно широко применяется для предобработки данных в прогнозировании и других видах анализа, поскольку позволяет исключить влияние случайной составляющей. Однако при сильных скачках ближе к «будущему» он может давать сбои. Недостаток метода состоит в его локальности — он не реагирует на данные в целом и не может прогнозировать резкое изменение поведения данных, опираясь лишь на ближайшие точки [13].

**3. Метод Брауна** подходит для данных с зависимостью от предыдущих значений и без сильных амплитудных колебаний. Экспоненциальное сглаживание — это наиболее распространенный метод прогнозирования различных временных рядов. Его главные достоинства заключаются в довольно простых вычислениях и гибкости описания разных изменений. Метод экспоненциального сглаживания позволяет получать оценки параметров тренда, описывающих не средний уровень явления, а тенденцию, сложившуюся к моменту последнего наблюдения. Для данного метода очень важно выбрать параметры сглаживания и начальные условия. Рассматриваемый метод прогнозирования является достаточно эффективным и надежным. Но он дает возможность спрогнозировать процесс только в краткосрочном периоде, т.е. всего лишь на 1–2 года вперед [14, 15].

**4. Метод Хольта–Уинтерса** универсален для переносимых выше особенностей, однако для данных нужно строгое определение сезонности [16]. Возможно, этот метод дал бы лучшие результаты, если бы мы разбили данные поквартально.

Метод Хольта–Уинтерса может применяться:

— при стратегическом планировании: построение основной тенденции развития (тренда) дает возможность учитывать восходящую или нисходящую динамику исследуемого явления;

— при оперативном и тактическом планировании: выявленная сезонная составляющая позволяет отметить неравномерность распределения объемов по годам по отношению к данной динамике.

Экспоненциальное сглаживание учитывает внутренние спады и подъемы в ряде динамики. Его можно использовать при выявлении крупных спадов и подъемов заблаговременно (при применении тактического планирования) и быть к ним готовым. Таким образом, метод имеет достаточно большую сферу применения. Данный метод основан на использовании большого объема статистических данных, что не всегда может быть актуально. Метод Хольта–Уинтерса может применяться при комбинированном прогнозировании одновременно с экспертными методами прогнозирования.

**5. Авторегрессионная модель** схожа с моделью Брауна, большое количество скачков не дает модели настроиться. Применение авторегрессионных моделей основано на предварительном анализе, когда известно, что изучаемый процесс в значительной степени

зависит от его развития в предыдущие периоды. В некоторых случаях они используются для нахождения простого преобразования, приводящего к последовательности независимых случайных величин. Область применения авторегрессионной модели ограничена, поскольку помимо сезонных изменений она ничего не описывает. Она редко используется в чистом виде. В большинстве случаев применяется более гибкая модель, включающая скользящее окно [15].

**6. Модель авторегрессии** — скользящего среднего может быть применима при предположении, что временной ряд стабилен, то есть его свойства не меняются во времени.

**7. Модель Бокса–Дженкинса** может быть применима к большому количеству типов данных, однако требует большей выборки (около 40 точек для хорошего прогноза) и тщательного исследования поведения временного ряда [11]. Недостаток заключается в том, что построение удовлетворительной модели *ARIMA* требует больших затрат ресурсов и времени.

До 1982 г. среди прогнозистов бытовало мнение, что модели *ARIMA* дают самые точные прогнозы, так как являются более общими для класса других моделей. Однако после проведения первых испытаний точности прогнозирования различных моделей в рамках «*M – Competition*» Международного института прогнозистов, в ходе которого модели *ARIMA* показали себя не лучше моделей экспоненциального сглаживания, это мнение сменилось на вполне логичное представление о том, что в каждом конкретном случае нужно использовать свою модель [16].

## ВЫВОДЫ

1. При использовании методов сглаживания по скользящей средней, наименьших квадратов, Бокса–Дженкинса, а также моделей Брауна и автокорреляции на результат прогноза влияют неоднородность и наличие скачкообразных изменений показателей, которые приводят к значительному снижению достоверности прогнозирования.

2. Применение модели Хольта–Уинтерса, учитывающей экспоненциальный тренд (тенденция изменения показателей временного ряда) и аддитивную сезонность (периодические колебания, наблюдаемые на временных рядах), является наиболее целесообразной для обработки статистических данных и прогнозирования на долгосрочные, среднесрочные и краткосрочные периоды с учетом специфики стационара, оказывающего скорую и неотложную помощь.

3. Выбор оптимального метода прогнозирования работы медицинского учреждения, основанного на выделении главных тенденций временного ряда, учитывая большинство особенностей при моделировании случайных процессов и случайных событий с некоторой случайной ошибкой, позволил уменьшить относительную ошибку прогноза.

## ЛИТЕРАТУРА

1. Карасев Н.А., Ермолов А.С., Турко А.П., Курилин Б.Л., Кислухина Е.В. Влияние реанимационной обеспеченности на результаты лечения острой хирургической патологии органов брюшной полости в многопрофильных больницах г. Москвы. *Московский хирургический журнал*. 2012; (1): 48–54.
2. Хубутя М.Ш., Карасев Н.А., Курилин Б.Л., Кислухина Е.В., Киселевская-Бабинина И.В., Молодов В.А. Развитие реанимационного коечного фонда в многопрофильных стационарах г. Москвы и его влияние на результаты лечебной деятельности. *Скорая медицинская помощь*. 2012; 13(3): 45–50.

- Ермолов А.С., Смирнов С.В., Карасев Н.А., Курилин Б.Л., Кислухина Е.В., Киселевская-Бабина И.В. и др. Анализ основных показателей работы Московского городского ожогового центра после модернизации. *Журнал им. Н.В. Склифосовского «Неотложная медицинская помощь»*. 2016;(1): 60–62.
- Хубутия М.Ш., Карасев Н.А., Кислухина Е.В., Васильев В.А., Курилин Б.Л., Медведева А.Б. и др. Анализ показателей клинической и организационной деятельности НИИ СП им. Н.В. Склифосовского в период 2005–2015 гг. *Журнал им. Н.В. Склифосовского «Неотложная медицинская помощь»*. 2016;(2): 59–63.
- Лукашин Ю.П. *Адаптивные методы краткосрочного прогнозирования временных рядов*. Москва: Финансы и статистика; 2003.
- Светушков С.Г. *Методы социально-экономического прогнозирования: в 2-х т.* Санкт-Петербург: Изд-во СПбГУЭФ; 2009.
- Линник Ю.В. *Метод наименьших квадратов и основы математико-статистической теории обработки наблюдений*. Москва: Физматгиз; 1958.
- Покровский В.И., Брико Н.И. (ред.) *Общая эпидемиология с основами доказательной медицины. 2-е изд., испр. и доп.* Москва: ГЭОТАР-Медиа; 2012.
- Савилов Е.Д., Мамонтова Л.М., Астафьев В.А., Жданова С.Н. *Применение статистических методов в эпидемиологическом анализе. 2-е изд., доп. и перераб.* Москва: МЕДпресс-информ; 2004.
- Winters P.R. Forecasting sales by exponentially weighted moving averages. *Management Science*. 1960; 6(3): 324–342.
- Орлов А.И. *Эконометрика*. Москва: Феникс; 2009.
- Погодин С.К. Методы оценки портфелей инвестиций, включающих ценные бумаги и недвижимость. Автореф. дис. ...канд. экон. наук. Москва, 2006. URL: <https://dlib.rsl.ru/viewer/01003277680#?page=1> [Дата обращения 12 августа 2019].
- Урбах В.Ю. *Математическая статистика для биологов и медиков*. Москва: Изд-во Академии наук СССР; 1963.
- Губинова Т.В. Применение экспоненциального сглаживания для прогнозирования выручки. *Материалы VII Междунар. студ. науч. форума-2015*. URL: <https://scienceforum.ru/2015/article/2015008484> [Дата обращения 12 августа 2019].
- Осипов Л.А., Кричевский А.М. Оценка и применение моделей временных рядов с долгой памятью в экономических задачах. *Информационно-управляющие системы*. 2007;(5):45–51.
- Makridakis S, Andersen A, Carbone R, Fildes R, Hibon M, Lewandowski R, et al. The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *J Forecast*. 1982;1(2):111–153.

## REFERENCES

- Karasev NA, Ermolov AS, Turko AP, Kurilin BL, Kislukhina EV. Vliyanie reanimatsionnoy obespechennosti na rezul'taty lecheniya ostroy khirurgicheskoy patologii organov bryuzhnoy polosti v mnogoprofil'nykh bol'nitsakh g. Moskvy. *Moskovskiy khirurgicheskij zhurnal*. 2012; (1): 48–54. (in Russ.)
- Khubutiya MSh, Karasev NA, Kurilin BL, Kislukhina EV, Kiselevskaya-Babinina IV, Molodov VA. Razvitie reanimatsionnogo koechnogo fonda v mnogoprofil'nykh stacionarakh g. Moskvy i ego vliyanie na rezul'taty lechebnoy deyatelnosti. *Skoraya meditsinskaya pomoshch'*. 2012; 13(3): 45–50. (in Russ.)
- Yermolov AS, Smirnov SV, Karasev NA, Kurilin BL, Kislukhina EV, Kiselevskaya-Babinina IV, et al. The Analysis of the Main Work Data in the Moscow City Burn Center after Remodeling. *Russian Sklifosovsky Journal "Emergency Medical Care"*. 2016; (1): 60–62. (in Russ.)
- Khubutiya MSh, Karasev NA, Kislukhina EV, Vasilyev VA, Kurilin BL, Medvedeva AB, et al. Analysis of Clinical and Organizational Activities in the N.V.Sklifosovsky Research Institute for Emergency Medicine in 2005–2015. *Russian Sklifosovsky Journal "Emergency Medical Care"*. 2016; (2): 59–63. (in Russ.)
- Lukashin YuP. *Adaptivnye metody kratkosrochnogo prognozirovaniya vremennykh ryadov*. Moscow: Finansy i statistika Publ.; 2003. (in Russ.)
- Svetun'kov SG. *Metody social'no-ehkonomicheskogo prognozirovaniya: in 2 vol.* Saint Peterburg: Izd-vo SpbGUEH Publ.; 2009. (in Russ.)
- Linnik YuV. *Metod naimen'shikh kvadratov i osnovy matematiko-statisticheskoy teorii obrabotki nablyudeniy*. Moscow: Fizmatgiz Publ.; 1958. (in Russ.)
- Pokrovskiy VI, Briko NI (ed.) Obshchaya epidemiologiya s osnovami dokazatel'noy meditsiny. 2-e izd., ispr. i dop.* Moskva: GEOTAR-Media; 2012. 2nd, rev. and ext. Moscow: GEOTAR-Media Publ.; 2012. (in Russ.)
- Savilov ED, Mamontova LM, Astaf'ev VA, Zhdanova SN. *Primenenie statisticheskikh metodov v ehpidemiologicheskoy analize. 2-e izd., dop. i pererab.* Moscow: MEDpress-inform; 2004. (in Russ.)
- Winters PR. Forecasting sales by exponentially weighted moving averages. *Management Science*. 1960; 6(3): 324–342.
- Orlov AI. *Ekonometrika*. Moscow: Feniks Publ.; 2009. (in Russ.)
- Pogodin SK. *Metody otsenki portfeley investitsiy, vkluchayushchikh tsennye bumagi i nedvizhimost'*. Cand. econ. sci. diss. Moscow, 2006. (in Russ.) Available at: <https://dlib.rsl.ru/viewer/01003277680#?page=1> [Accessed August 12, 2019].
- Urbakh VYu. *Matematicheskaya statistika dlya biologov i medikov*. Moscow: Izd-vo Akademii nauk SSSR Publ.; 1963. (in Russ.)
- Gubinova TV. *Primenenie eksponentsial'nogo sglazhivaniya dlya prognozirovaniya vyruchki. Materialy VII Mezhdunar. stud. nauch. foruma-2015*. Available at: <https://scienceforum.ru/2015/article/2015008484> (in Russ.)
- Osipov LA, Krichevsky AM. Estimation and application of time series models with long memory in economical problems. *Informatsionno-upravlyaiushchie sistemy*. 2007; (5): 45–51. (in Russ.)
- Makridakis S, Andersen A, Carbone R, Fildes R, Hibon M, Lewandowski R, et al. The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *J Forecast*. 1982; 1(2): 111–153.

## ИНФОРМАЦИЯ ОБ АВТОРАХ

## Курилин Борис Леонидович

научный сотрудник лаборатории организации стационарной неотложной помощи ГБУЗ «НИИ СП им. Н.В. Склифосовского ДЗМ», <https://orcid.org/0000-0003-3019-1947>

## Киселевская-Бабина Виктория Ярославовна

инженер лаборатории АСУ ГБУЗ «НИИ СП им. Н.В. Склифосовского ДЗМ», <https://orcid.org/0000-0002-9057-2162>

## Карасёв Николай Александрович

заведующий лабораторией организации стационарной неотложной помощи ГБУЗ «НИИ СП им. Н.В. Склифосовского ДЗМ», <https://orcid.org/0000-0001-8370-6145>

## Киселевская-Бабина Ирина Викторовна

инженер лаборатории АСУ ГБУЗ «НИИ СП им. Н.В. Склифосовского ДЗМ», <https://orcid.org/0000-0003-4474-4469>

## Кислухина Евгения Викторовна

научный сотрудник лаборатории организации стационарной неотложной помощи ГБУЗ «НИИ СП им. Н.В. Склифосовского ДЗМ», <https://orcid.org/0000-0002-8980-4931>

## Васильев Владислав Андреевич

старший научный сотрудник лаборатории организации стационарной неотложной помощи ГБУЗ «НИИ СП им. Н.В. Склифосовского ДЗМ», <https://orcid.org/0000-0003-3205-1243>

Received on 19.11.2018

Accepted on 06.03.2019

Поступила в редакцию 19.11.2018

Принята к печати 06.03.2019

## Selection of Prediction Method of Basic Statistical Work Parameters of N.V. Sklifosovsky Research Institute for Emergency Medicine of the Moscow Healthcare Department

**B.L. Kurilin\*, V.Y. Kisselevskaya-Babinina, N.A. Karasyov, I.V. Kisselevskaya-Babinina, E.V. Kislukhkina, V.A. Vasilyev**

Laboratory for Organization of Inpatient Care  
N.V. Sklifosovsky Research Institute for Emergency Medicine of the Moscow Healthcare Department  
3 Bolshaya Sukharevskaya Square, Moscow 129090, Russian Federation

\* **Contacts:** Boris L. Kurilin, Researcher of the Laboratory for Organization of Inpatient Care, N.V. Sklifosovsky Research Institute for Emergency Medicine of the Moscow Healthcare Department. Email: kb111@yandex.ru

**BACKGROUND** The most important part of the state social and economic policy is optimization of the healthcare system, where the loss of public health leads to economic damage. Against this background, forecasting the work of medical institutions is the basis for the successful development of healthcare, despite the fact that the healthcare system, indicators and standards of medical and social welfare are still not stable, and a clear development strategy for the short- and long-term period has not been worked out.

**AIM OF STUDY** Determining the most optimal method for predicting the work of a medical institution, based on identification of the main trends in the time series when constructing a model of the dependence of parameters or determining the behavior of data as a stochastic series (i.e. modeling random processes and random events with some random error).

**MATERIAL AND METHODS** To predict the main statistical indicators of N.V. Sklifosovsky Research Institute for Emergency Medicine based on a retrospective analysis, data were used that were submitted to the City Bureau of Medical Statistics and entered into official reporting forms (form № 30, approved by Goskomstat of the Russian Federation dated September 10, 2002, № 175): the number of hospitalized patients and mortality rates in inpatient and intensive care units.

To select the optimal methodology for the experimental forecast model, data were used for the period from 1991 to 2016. Indicators for 2017 were taken as control values.

**RESULTS** As a result of the comparison of several methods (moving averages, least squares approach, Brown model, Holt–Winters method, autocorrelation model, Box–Jenkins method) as applied to the work of N.V. Sklifosovsky Research Institute for Emergency Medicine, the Holt–Winters model was chosen as the most appropriate one for the data characteristics.

**FINDINGS** 1. When using methods of moving averages, least squares, Box–Jenkins, as well as Brown model and autocorrelation, the forecast result is not always influenced by strictly straight-line indicators of the time series, due to the heterogeneity of the time series and the presence of outliers (often found in a medical institution providing emergency care), which lead to a significant decrease in the reliability of forecasting.

2. The application of the Holt–Winters model, which takes into account the exponential trend (the trend of time series indicators) and additive season (periodic fluctuations observed in the time series), is most suitable for processing statistical data and forecasting for long-term, medium-term and short-term periods taking the specifics of a hospital providing emergency care into account.

3. The choice of the optimal method for predicting the work of a medical institution, based on the identification of the main trends in the time series, taking most of the features in the modeling of random processes and events into account, allowed to reduce the relative forecast error.

**Keywords:** forecasting, alignment of time series, assessment of the reliability of differences in indicators

**For citation** Kurilin BL, Kisselevskaya-Babinina VY, Karasyov NA, Kisselevskaya-Babinina IV, Kislukhkina EV, Vasilyev VA. Selection of Prediction Method of Basic Statistical Work Parameters of N.V. Sklifosovsky Research Institute for Emergency Medicine of the Moscow Healthcare Department. *Russian Sklifosovsky Journal of Emergency Medical Care*. 2019;8(3):246–256. <https://doi.org/10.23934/2223-9022-2019-8-3-246-256> (in Russ.)

**Conflict of interest** Authors declare lack of the conflicts of interests

**Acknowledgments** The study had no sponsorship

### Affiliations

Boris L. Kurilin	Researcher of the Laboratory for Organization of Inpatient Emergency Care, N.V. Sklifosovsky Research Institute For Emergency Medicine, <a href="https://orcid.org/0000-0003-3019-1947">https://orcid.org/0000-0003-3019-1947</a>
Victoria Y. Kisselevskaya-Babinina	Engineer of the Laboratory of Automated Control Systems, N.V. Sklifosovsky Research Institute for Emergency Medicine, <a href="https://orcid.org/0000-0002-9057-2162">https://orcid.org/0000-0002-9057-2162</a>
Nikolay A. Karasev	Head of the Laboratory for Organization of Inpatient Emergency Care, N.V. Sklifosovsky Research Institute for Emergency Medicine, <a href="https://orcid.org/0000-0001-8370-6145">https://orcid.org/0000-0001-8370-6145</a>
Irina V. Kisselevskaya-Babinina	Engineer of the Laboratory of Automated Control Systems, N.V. Sklifosovsky Research Institute for Emergency Medicine, <a href="https://orcid.org/0000-0003-4474-4469">https://orcid.org/0000-0003-4474-4469</a>
Evgeniya V. Kislukhkina	Researcher of the Laboratory for Organization of Inpatient Emergency Care, N.V. Sklifosovsky Research Institute for Emergency Medicine, <a href="https://orcid.org/0000-0002-8980-4931">https://orcid.org/0000-0002-8980-4931</a>
Vladislav A. Vasilyev	Senior Researcher of the Laboratory for Organization of Inpatient Emergency Care, N.V. Sklifosovsky Research Institute for Emergency Medicine, <a href="https://orcid.org/0000-0003-3205-1243">https://orcid.org/0000-0003-3205-1243</a>