

**ЦЕНТР НАУК О ЖИЗНИ НАЗАРБАЕВ УНИВЕРСИТЕТА  
ЦЕНТР ИЗУЧЕНИЯ ГЛОБАЛЬНОГО ЗДОРОВЬЯ ЦЕНТРАЛЬНОЙ АЗИИ  
НАЦИОНАЛЬНЫЙ ЦЕНТР ПРОБЛЕМ ТУБЕРКУЛЕЗА МЗ РК**

«Согласовано»

Руководитель Центра научных исследований, экспертизы и инновационной деятельности в здравоохранении РГП «РЦРЗ» МЗиСР РК

 В. Койков

« 20 » октября 2014 г.

«Утверждаю»

Директор Департамента науки и человеческих ресурсов Министерства здравоохранения и социального развития Республики Казахстан

 А. Сыздыкова

« 19 » октября 2014 г.



**Прогнозирование заболеваемости туберкулёзом  
в Республике Казахстан**

Методические рекомендации

Астана-Алматы, 2014

**УДК:616.24-002.5**

**ББК 55.4**

**П 78**

**Рецензенты:**

**Слажнева Т.И.** Главный научный сотрудник отдела науки Национального центра проблем формирования здорового образа жизни МЗ РК, д.м.н., профессор

**Аденов М.М.** главный врач Национального центра проблем туберкулеза МЗ РК, к.м.н.

**Авторы:**

**Жусупов Б.С.** Зам.Генерального менеджера Центра изучения глобального здоровья в Центральной Азии, MSc

**Абилдаев Т.Ш.** Директор НЦПТ МЗ РК, д.м.н., профессор

**Исаева Р.Б.** Зам.Генерального директора ЧУ «Центр наук о жизни» Назарбаев университета, д.м.н

**Терликбаева А.М.** Генеральный менеджер Центра изучения глобального здоровья в Центральной Азии, MSc

**Даришева М.А.** Координатор проекта Центра изучения глобального здоровья в Центральной Азии, к.м.н.

**Гуляев П.И.** ИТ специалист Центра изучения глобального здоровья в Центральной Азии

**Берикова Э.А.** зам. директора по научно-исследовательской работе НЦПТ МЗ РК, к.м.н.

**Аленова А.Х.** г.н.с. НЦПТ МЗ РК, д.м.н., профессор

**Исмаилова А.Т.** Ученый секретарь НЦПТ МЗ РК, к.м.н.

**Прогнозирование заболеваемости туберкулезом в Республике Казахстан:**  
Методические рекомендации // Б.С.Жусупов, Т.Ш.Абилдаев, Р.Б.Исаева и др. – Алматы, 2014. – 21 с.

**ISBN 978-601-305-019-5**

В разработанных методических рекомендациях представлены методы прогнозирования заболеваемости туберкулезом в Республике Казахстан с использованием моделей ARIMA. Методические рекомендации предназначены для широкого круга специалистов противотуберкулезных учреждений, организаторов здравоохранения.

**УДК 616.24-002.5**

**ББК 55.4**

© ЦЕНТР НАУК О ЖИЗНИ НАЗАРБАЕВ УНИВЕРСИТЕТА, 2014  
© ЦЕНТР ИЗУЧЕНИЯ ГЛОБАЛЬНОГО ЗДОРОВЬЯ ЦЕНТРАЛЬНОЙ АЗИИ, 2014  
© НАЦИОНАЛЬНЫЙ ЦЕНТР ПРОБЛЕМ ТУБЕРКУЛЕЗА МЗ РК, 2014

## Содержание

1. Основные термины и сокращения	4
2. Введение	5
3. Основная часть	6
3.1. Краткое теоретическое описание моделирования временных рядов ARIMA	6
3.1.1 Формулы и основные термины ARIMA	6
3.1.2 Определение оптимальной модели ARIMA	7
3.1.3 Оценка модели ARIMA	7
3.2 Пример прогноза новых случаев туберкулеза с использованием ARIMA	8
3.2.1 Оценка сезонности заболеваемости туберкулезом	8
3.2.2 Модель ARIMA на основе национальных данных о регистрации новых случаев туберкулеза	10
3.2.3 Модель ARIMA на основе областных и городских данных о регистрации новых случаев туберкулеза	11
3.2.4 Сравнение прогнозов ARIMA моделей с реальными данными о регистрации новых случаев туберкулеза в 2013 году	14
3.3 Практический алгоритм прогнозирования новых случаев туберкулеза с использованием ARIMA моделей	16
4. Заключение	20
Литература	21

## 1. Основные термины и сокращения

Туберкулез - инфекционное заболевание, вызываемое микобактериями туберкулёза, характеризующееся образованием патологических изменений в любых тканях, главным образом в легких.

Прогнозирование — это разработка прогноза; специальное научное исследование конкретных перспектив развития какого-либо процесса.

ARIMA - англ. *autoregressive integrated moving average* - интегрированная модель авторегрессии — скользящего среднего — модель и методология анализа временных рядов, иногда называемые моделями (или методологией) Бокса-Дженкинса.

## 2. Введение

Математическое моделирование является одним из самых эффективных инструментов для прогнозирования эпидемиологической ситуации по туберкулезу, которое используется для разработки программ по профилактике и контролю туберкулеза и оценки их воздействия. Процесс распространения туберкулезной инфекции существенно различается в разных регионах мира. Это обусловлено разным эпидемическим фоном, географическими и климатическими условиями, экономическими, демографическими и социальными характеристиками.

Одним из эффективных методов прогнозирования является интегрированная модель авторегрессии и скользящего среднего (Autoregressive integrated moving average - ARIMA). Эта модель анализа временных рядов широко используется в мире для предсказания случаев различных заболеваний, в том числе туберкулеза. Модель прогнозирования ARIMA использует исторические данные о заболеваемости. Она способна учесть сезонные особенности, то есть уменьшение или увеличение регистрации случаев в зависимости от месяца. Развитие компьютерной техники, появление программ обработки данных, находящихся в свободном доступе, создают благоприятные условия для применения ARIMA в целях прогнозирования количества новых случаев туберкулеза в краткосрочной и среднесрочной перспективе.

В данном методическом руководстве представлены краткое описание методологии ARIMA, принципы выбора оптимальной модели, а также показаны практические шаги, необходимые для внедрения этого метода прогнозирования в противотуберкулезную службу.

### 3. Основная часть

#### 3.1 Краткое теоретическое описание моделирования временных рядов ARIMA

##### 3.1.1 Формулы и основные термины ARIMA

*Модели временных рядов* — это математические модели прогнозирования, которые стремятся найти зависимость будущего значения от прошлого внутри самого процесса и на этой зависимости вычислить прогноз. Эти модели универсальны для различных предметных областей и активно используются для прогнозирования случаев заболевания, в том числе туберкулеза [1].

Несезонная модель ARIMA обычно представляется в виде «ARIMA(p,d,q)», где p, d, q - целые числа, задающие порядок модели:

- p – порядок авторегрессии;
- d – порядок интегрирования;
- q – порядок скользящего среднего.

Модель ARIMA для ряда  $X_t$  имеет следующий вид [2]:

$$\Delta^d X_t = c + \sum_{i=1}^p a_i \Delta^d X_{t-i} + \sum_{j=1}^q b_j \varepsilon_{t-j} + \varepsilon_t$$

где,

$\Delta^d$  - оператор разности временного ряда порядка d;

$c, a_i, b_j$  – параметры модели (коэффициенты);

$\varepsilon_t$  – ошибка, которая представляет собой стационарный временной ряд, чьи характеристики не меняются со временем: с математическим ожиданием, равным 0, и неизменной дисперсией.

Сезонные модели ARIMA (seasonal ARIMA – SARIMA) построены на тех же принципах [3]. Отличием является тот факт, что для прогноза переменной, в нашем случае – количества новых случаев туберкулеза, принимаются во внимание не только значения, относящиеся к предыдущим месяцам, но также значения переменной из предыдущего цикла. Например, для прогнозирования количества случаев в марте 2015 года используются прогнозные значения в феврале 2015 года (предшествующий месяц) и прогнозные или фактические значения в марте 2014 года (тот же месяц предшествующего года). Сезонные модели ARIMA представляются в виде «SARIMA(p,d,q)(p<sub>12</sub>,d<sub>12</sub>,q<sub>12</sub>)», где порядки p<sub>12</sub>,d<sub>12</sub>,q<sub>12</sub> относятся к сезонной (циклической) составляющей с размером цикла равным 12 (количество месяцев в году).

### 3.1.2 Определение оптимальной модели ARIMA

Нахождение оптимальной модели ARIMA представляет собой сложную задачу. На первом этапе на основе значений автокорреляционной функции и функции частичной автокорреляции определяется порядок интегрирования d и порядок авторегрессии p. Основная цель – достижение стационарности ошибки. Затем на основе Информационного критерия Акаике (AIC) из нескольких статистических моделей выбирается лучшая, то есть лучше всего предсказывающая значения временного ряда по методу максимального правдоподобия [4]. В настоящее время существуют программы, которые автоматически подбирают оптимальную модель ARIMA [5], что значительно упрощает процедуру и позволяет осуществлять прогнозирование на основе ARIMA специалистам без специального статистического образования.

### 3.1.3 Оценка модели ARIMA

Оценки модели включает два основных этапа. Во-первых, это проверка распределения временного ряда, состоящего из ошибок прогноза (residuals). Этот временной ряд должен быть стационарным, то есть его характеристики, такие как среднее и дисперсия, должны быть постоянными. Ошибка должна быть подчинена нормальному распределению и значения функции автокорреляции должны быть низкими. Во-вторых, оценивается точность прогноза путем сравнения прогнозных и фактических значений. Результат сравнения отражается в нескольких показателях – среднее абсолютной

разницы (mean absolute deviation - MAD), корень средней квадратической ошибки (root-mean-square error - RMSE) и среднее абсолютной процентной ошибки (mean absolute percentage error - MAPE) [6].

## 3.2 Пример прогноза новых случаев туберкулеза с использованием ARIMA

### 3.2.1 Оценка сезонности заболеваемости туберкулезом

Количество месячной регистрации новых случаев заболевания туберкулезом в 2007-2012 годах представлено на рисунке 1.

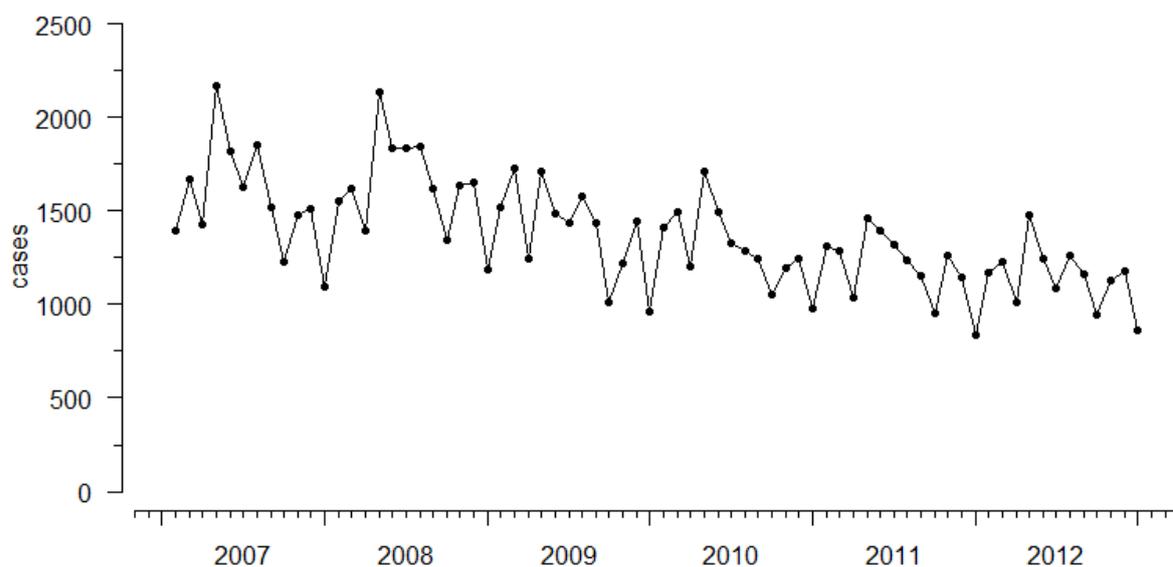


Рисунок 1. Количество зарегистрированных новых случаев туберкулеза в 2007-2012 годах по месяцам.

Рисунок 1 показывает значительные колебания уровня регистрируемой заболеваемости, что характерно для многих стран мира [7-11]. Чтобы понять, влияет ли фактор сезонности на регистрацию новых случаев, данные по отдельным месяцам были распределены по 12 группам. В первую группу вошли показатели января (январь 2007 года, январь 2008 года, ..., январь

2012 года), в последнюю группу – показатели декабря. Средние значения по группам показаны на рисунке 2.

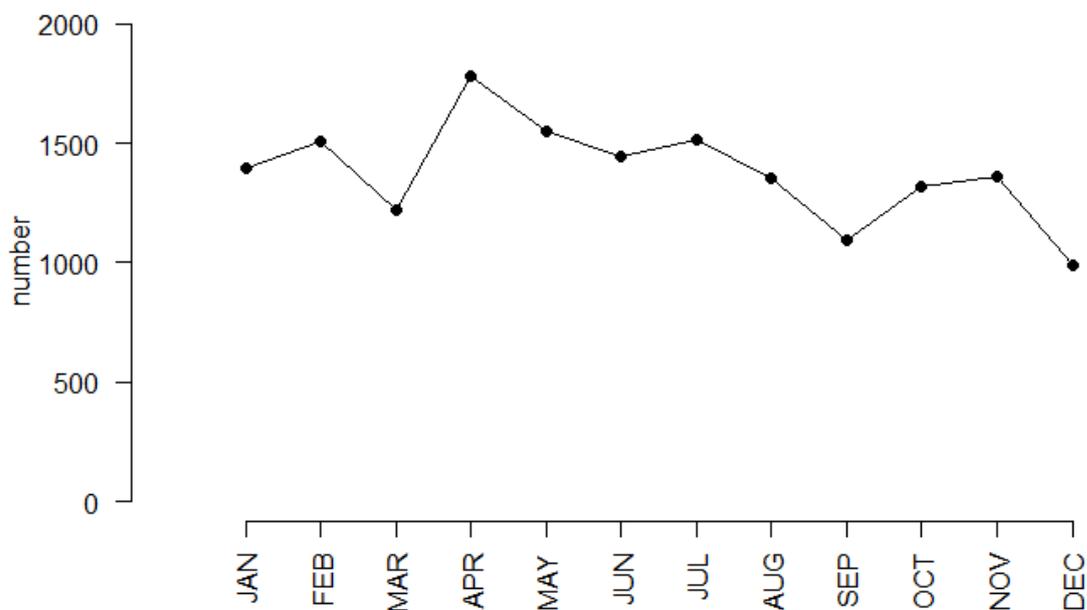


Рисунок 2. Средние значения числа зарегистрированных новых случаев туберкулеза по месяцам за 2007-2012 годы.

Сравнивая данные за 6 лет (2007-2012), пик регистрации наблюдался в апреле, когда в среднем регистрировалось 1778 новых случаев, минимальное значение в сентябре (1091 случаев) и декабре (986 случаев). Дисперсионный анализ показал статистически значимые различия в регистрации по месяцам -  $F_{11,60} = 5.752$  ( $p < 0.001$ ), что обуславливает использование сезонной модели  $ARIMA - SARIMA_{12}$ .

### 3.2.2 Модель ARIMA на основе национальных данных о регистрации новых случаев туберкулеза

Для построения модели были использованы данные о месячной регистрации новых случаев ТБ<sup>1</sup>, полученные из Национального регистра случаев туберкулеза с 2007 по 2012 годы. Эти данные были обработаны в программе R с использованием пакета forecast [5]. Более точно временной ряд описывает модель SARIMA(0,1,1)(1,0,1)<sub>12</sub> (рисунок 3).

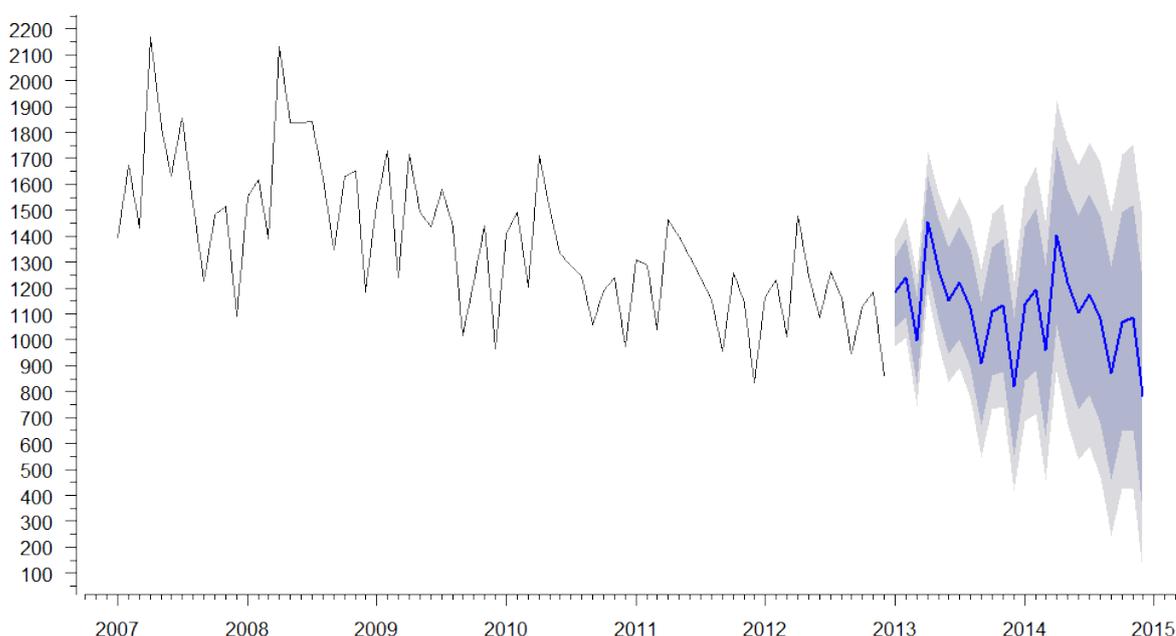


Рисунок 3. Количество зарегистрированных новых случаев ТБ по месяцам в 2007-2012 годах и прогноз на 2013-2014 годы с использованием национальных данных

Прогноз заболеваемости ТБ учитывает как сезонность, так и общий тренд на снижение количества новых случаев. Вместе с тем, обращает внимание большая величина размаха оценок. Так, 95% интервальная оценка новых случаев в декабре 2014 года находится в пределах от 338 до 1467 случаев.

<sup>1</sup> Отчетный период начинался с 26 числа предыдущего месяца и заканчивался 25 числом отчетного месяца.

### 3.2.3 Модель ARIMA на основе областных и городских данных о регистрации новых случаев туберкулеза

Чтобы повысить точность прогнозирования, были определены оптимальные модели ARIMA для каждого региона Казахстана: 14 областей и города Астана и Алматы. Поскольку регионы отличаются по климатическим условиям, социально-демографическим и экономическим характеристикам, динамика регистрации новых случаев туберкулеза в регионах ожидаемо представлена разными моделями ARIMA (таблица 1).

Таблица 1. Оптимальные модели ARIMA, представляющие динамику регистрации новых случаев туберкулеза в административных регионах Казахстана

Регион	Модель ARIMA
Акмолинская область	SARIMA(2,1,0)(2,0,0) <sub>12</sub>
Актюбинская область	SARIMA(2,1,1)(1,0,1) <sub>12</sub>
Алматинская область	SARIMA(2,0,2)(1,0,1) <sub>12</sub>
Атырауская область	SARIMA(1,1,1)(1,0,0) <sub>12</sub>
Западно-Казахстанская область	SARIMA(2,1,0)(0,0,1) <sub>12</sub>
Жамбыльская область	SARIMA(1,1,1)(2,0,0) <sub>12</sub>
Карагандинская область	SARIMA(0,0,0)(0,1,1) <sub>12</sub>
Костанайская область	SARIMA(2,1,0)(1,0,0) <sub>12</sub>
Кызылординская область	ARIMA(4,1,2)
Мангистауская область	SARIMA(1,1,1)(1,0,0) <sub>12</sub>
Южно-Казахстанская область	SARIMA(2,1,1)(2,0,0) <sub>12</sub>
Павлодарская область	SARIMA(1,1,1)(1,0,0) <sub>12</sub>
Северо-Казахстанская область	SARIMA(0,1,1)(1,0,1) <sub>12</sub>
Восточно-Казахстанская область	SARIMA(0,1,1)(2,0,0) <sub>12</sub>
г.Астана	SARIMA(0,1,1)(1,1,1) <sub>12</sub>
г.Алматы	SARIMA(1,0,0)(2,0,1) <sub>12</sub>
Республика Казахстан	SARIMA(0,1,1)(1,0,1) <sub>12</sub>

Графическое представление региональных моделей ARIMA вместе с прогнозами на 2013-2014 годы показано на рисунках 4-5.

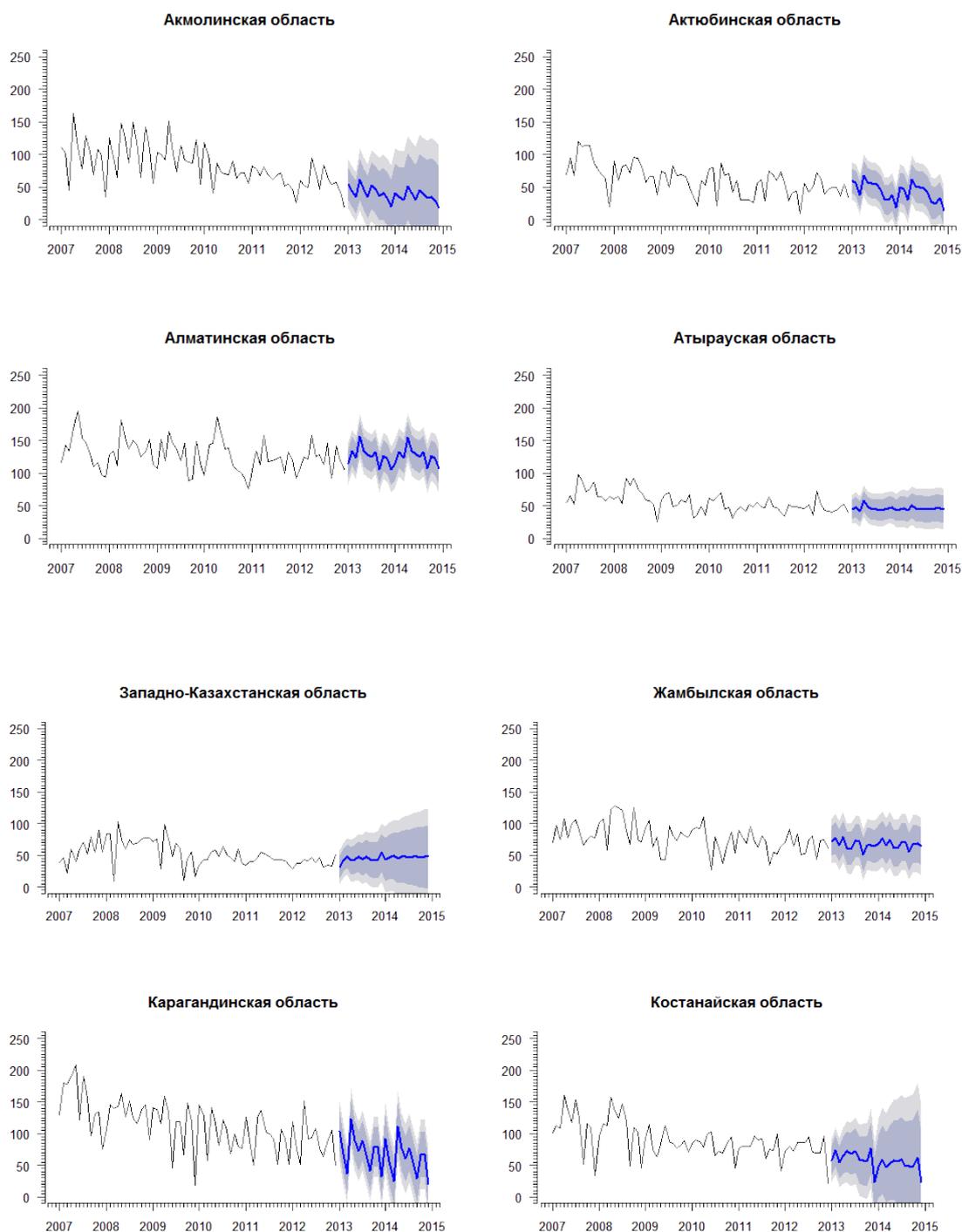


Рисунок 4. Месячная заболеваемость туберкулезом в регионах Казахстана в 2007-2012гг. и прогноз на 2013-2014гг. (8 областей)

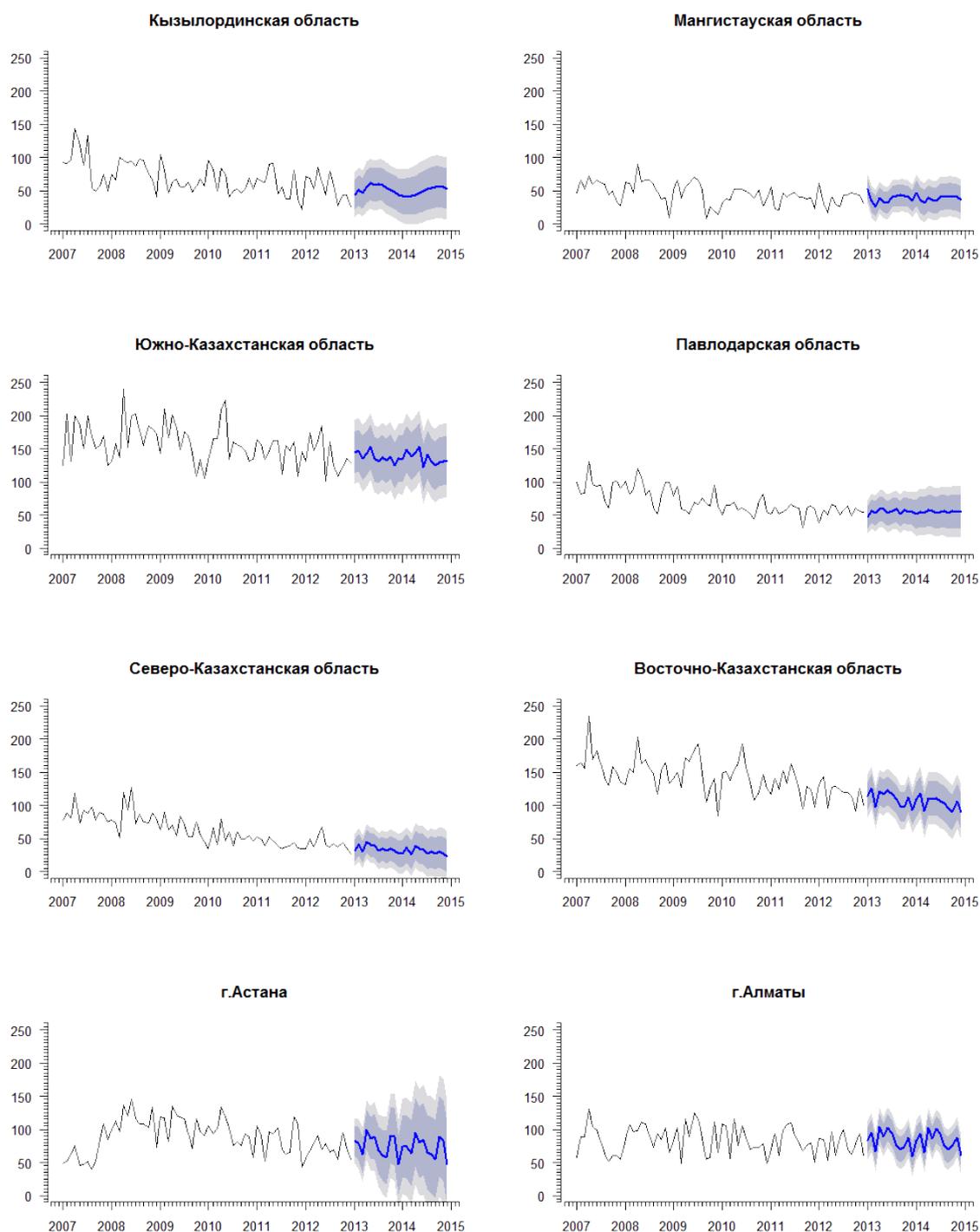


Рисунок 5. Месячная заболеваемость туберкулезом в регионах Казахстана в 2007-2012гг. и прогноз на 2013-2014гг. (6 областей, г. Астана и Алматы)

Уточненный вариант прогноза по Казахстану представляет собой обобщение региональных моделей, при этом его дисперсия ниже, чем у модели, построенной только на основе данных по республике (рисунок 6).

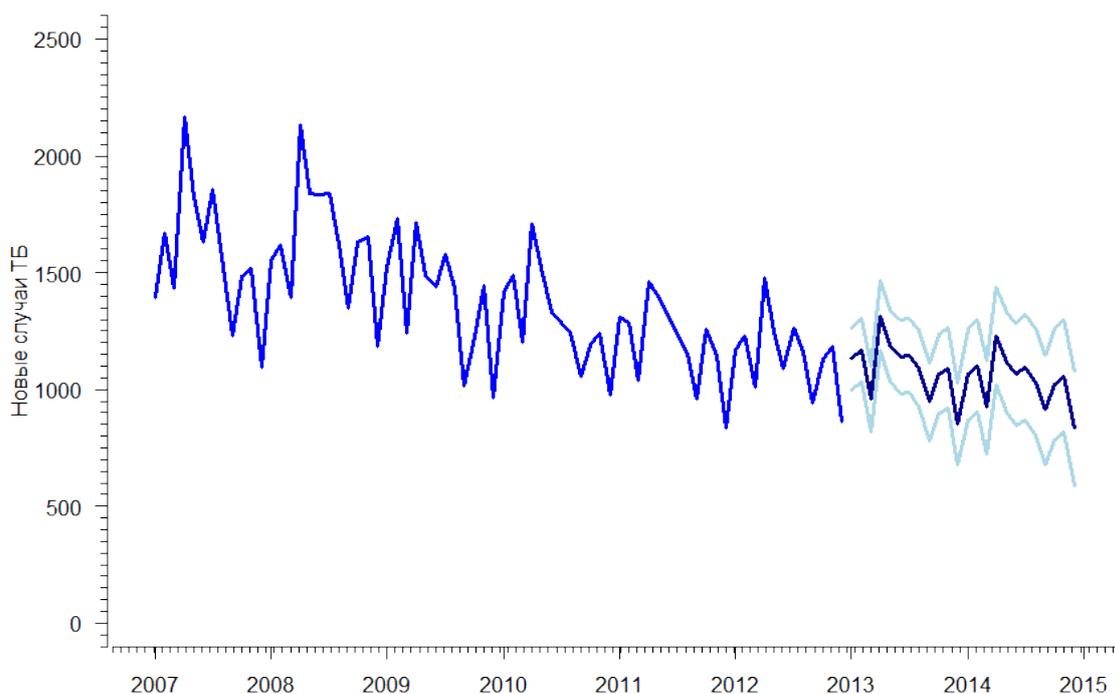


Рисунок 6. Месячная заболеваемость ТБ в 2007-2012 годах и уточненный прогноз на 2013-2014 годы (Республика Казахстан) с использованием региональных данных

### 3.2.4 Сравнение прогнозов ARIMA моделей с реальными данными о регистрации новых случаев туберкулеза в 2013 году

Для оценки правильности прогнозирования нами были сравнены результаты прогнозов трех моделей с фактическими данными регистрации за 2013 год. Первая модель – это повторение исторических данных. Данный способ прогнозирования чаще всего используется в здравоохранении. В основе него лежит предположение, что уровень заболеваемости в предстоящем году будет таким же, как и в нынешнем году. Вторая модель прогнозирования представляет собой модель ARIMA, построенную на основе

только республиканских показателей регистрации заболевания. Третья модель является суммированием региональных моделей ARIMA для получения национального прогноза. Сравнение результатов прогнозирования представлены в таблице 2 и на рисунке 7.

Таблица 2. Прогнозные и фактические значения количества зарегистрированных новых случаев туберкулеза в 2013 году

	Модель 1. Данные регистрации за 2012 год		Модель 2. Модель SARIMA (0,1,1)(1,0,1) <sub>12</sub> на основе республиканских данных		Модель 3. Суммирование региональных моделей ARIMA		Фактическое количество зарегистрированных новых случаев туберкулеза в 2013 году
	Прогноз	АПО*	Прогноз	АПО*	Прогноз	АПО*	
Январь	1166	10.5	1183	12.1	1132	7.3	1055
Февраль	1230	7.2	1239	8.0	1168	1.8	1147
Март	1012	10.7	998	9.2	957	4.7	914
Апрель	1480	13.8	1453	11.8	1317	1.3	1300
Май	1247	10.3	1270	12.3	1189	5.1	1131
Июнь	1086	5.0	1149	11.1	1139	10.2	1034
Июль	1262	8.8	1221	5.3	1145	-1.3	1160
Август	1163	19.8	1123	15.7	1090	12.3	971
Сентябрь	944	7.8	908	3.7	946	8.0	876
Октябрь	1128	17.5	1111	15.7	1068	11.3	960
Ноябрь	1182	22.6**	1131	17.3**	1091	13.2**	964
Декабрь	863	12.2	820	6.6	852	10.8	769

\* абсолютная процентная ошибка

\*\* самое большое значение АПО

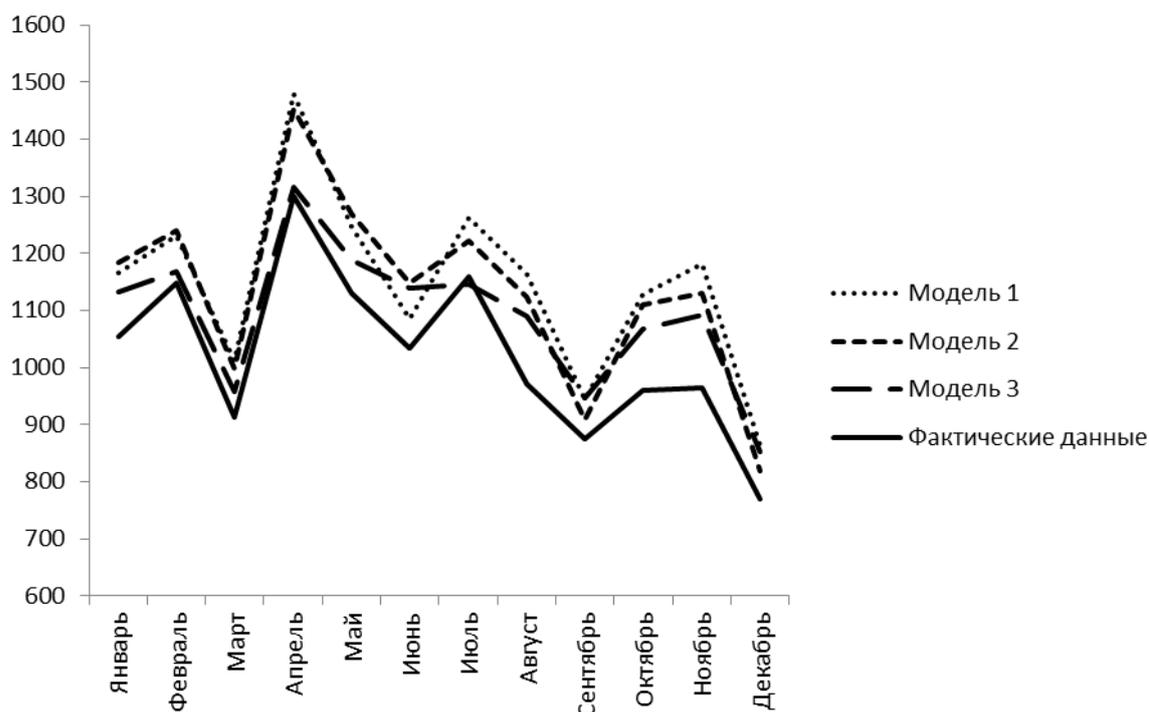


Рисунок 7. Прогнозные и фактические значения количества зарегистрированных новых случаев туберкулеза в 2013 году

Модель прогнозирования, составленная как сумма региональных моделей ARIMA, продемонстрировала лучшую пригодность, то есть ее прогнозы были ближе к фактической регистрации новых случаев заболевания. Ее применение позволяет почти в два раза уменьшить размер ошибки по сравнению с тем, когда в качестве прогноза используются исторические данные, то есть данные за прошлый год.

### 3.3 Практический алгоритм прогнозирования новых случаев туберкулеза с использованием ARIMA моделей

Вычисления для подбора оптимальных моделей ARIMA, определения ее параметров (коэффициентов) и прогнозных значений осуществляется в программе R, которая представляет собой язык программирования для статистической обработки данных и работы с графикой. Программа R

является бесплатной и имеет свободный доступ<sup>2</sup>. Ниже представлена последовательность действий для выполнения прогнозирования ARIMA в кратком виде:

1. Установка программы R
2. Загрузка и инициирование пакета forecast
3. Загрузка данных для прогнозирования
4. Определение оптимальной модели ARIMA
5. Получение результатов прогноза на основе оптимальной модели ARIMA

## 1. Установка программы R

Загрузить и установить R на компьютер можно с сайта <http://www.r-project.org/>, выбрав любой из зеркальных сайтов, например <http://cran.rstudio.com/>

Существует версия программы R на русском языке. Для ее выбора во время установки нужно указать русский язык в качестве рабочего языка. После установки программы необходимо зайти в директорию R и нажать на

соответствующую иконку 

## 2. Загрузка и инициирование пакета forecast

Для вычисления ARIMA моделей необходим пакет forecast. Его необходимо загрузить и инициировать.

В меню «Пакеты» выберите «Установить пакеты». Затем укажите любой зеркальный сайт, с которого будет производиться загрузка. Во всплывающем списке пакетов выберите forecast. Для инициирования пакета, то есть его загрузки в память компьютера в командной строке наберите:

```
library('forecast')
```

---

<sup>2</sup> Более подробно с этой программой можно ознакомиться на сайтах [http://ru.wikibooks.org/wiki/%D0%AF%D0%B7%D1%8B%D0%BA\\_%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D1%8F\\_R](http://ru.wikibooks.org/wiki/%D0%AF%D0%B7%D1%8B%D0%BA_%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D1%8F_R) и <http://soc-research.info/blog/rbook.html>

### 3. Загрузка данных для прогнозирования

Количество зарегистрированных новых случаев туберкулеза необходимо ввести в одну строку в программе MS Excel. Следующим шагом нужно выделить ячейки с данными и скопировать их.

В командной строке R набрать:

```
tb=scan('clipboard')
```

Далее необходимо преобразовать данные в формат временного ряда.

```
tb=ts(tb, start=c(2007,1), frequency=12)
tb
```

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2007	1395	1672	1431	2169	1821	1632	1856	1523	1228	1481	1515	1092
2008	1554	1618	1391	2133	1838	1836	1841	1621	1347	1633	1651	1186
2009	1522	1730	1242	1715	1490	1438	1580	1435	1016	1223	1442	964
2010	1414	1492	1202	1711	1495	1332	1287	1243	1056	1191	1241	975
2011	1308	1286	1038	1462	1397	1321	1240	1151	957	1258	1143	836
2012	1166	1230	1012	1480	1247	1086	1262	1163	944	1128	1182	863

### 4. Определение оптимальной модели ARIMA

Для определения оптимальной модели ARIMA необходимо использовать функцию `auto.arima()`

```
fit_tb1=auto.arima(tb)
summary(fit_tb)
```

```
Series: tb
ARIMA(0, 1, 1) (1, 0, 1) [12]
```

```
Coefficients:
      ma1      sar1      sma1
-0.494  0.975  -0.515
s.e.    0.130  0.019  0.157
```

```
sigma^2 estimated as 11240: log likelihood=-444
AIC=895 AICc=896 BIC=904
```

```
Training set error measures:
ME RMSE MAE MPE MAPE MASE ACF1
Training set -4.64 105 80.7 -0.494 5.95 0.628 0.0114
```

## 5. Получение результатов прогноза на основе оптимальной модели ARIMA

forecast(fit\_tb)

	Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2013		1182.5360	1046.6210	1318.451	974.6719	1390.400
Feb 2013		1239.0495	1086.7228	1391.376	1006.0859	1472.013
Mar 2013		998.1122	830.9777	1165.247	742.5020	1253.722
Apr 2013		1452.9977	1272.2645	1633.731	1176.5901	1729.405
May 2013		1269.7151	1076.3371	1463.093	973.9691	1565.461
Jun 2013		1149.3365	944.0914	1354.582	835.4412	1463.232
Jul 2013		1220.5265	1004.0638	1436.989	889.4754	1551.578
Aug 2013		1122.7369	895.6100	1349.864	775.3763	1470.097
Sep 2013		908.2924	670.9801	1145.605	545.3545	1271.230
Oct 2013		1111.2678	864.1895	1358.346	733.3941	1489.141
Nov 2013		1130.6920	874.2194	1387.165	738.4509	1522.933
Dec 2013		819.6726	554.1378	1085.207	413.5721	1225.773
Jan 2014		1137.6562	841.3878	1433.925	684.5527	1590.760
Feb 2014		1192.7623	879.9346	1505.590	714.3336	1671.191
Mar 2014		957.8248	629.2713	1286.378	455.3455	1460.304
Apr 2014		1401.3829	1057.8226	1744.943	875.9528	1926.813
May 2014		1222.6643	864.7259	1580.603	675.2448	1770.084
Jun 2014		1105.2834	733.5225	1477.044	536.7242	1673.843
Jul 2014		1174.7006	789.6131	1559.788	585.7601	1763.641
Aug 2014		1079.3461	681.3780	1477.314	470.7064	1687.986
Sep 2014		870.2417	459.7970	1280.686	242.5207	1497.963
Oct 2014		1068.1627	645.6096	1490.716	421.9235	1714.402
Nov 2014		1087.1032	652.7792	1521.427	422.8620	1751.344
Dec 2014		783.8287	338.0445	1229.613	102.0606	1465.597

Прогнозные значения указаны в первой строке.

#### 4. Заключение

Разработанная методика кратко- и среднесрочного прогнозирования заболеваемости туберкулезом и развития эпидемического процесса может использоваться для планирования профилактических и контрольных мероприятий. Актуальность применения предложенных методов прогнозирования в значительной степени обусловлена особенностями эпидемиологической ситуации в стране. Внедрение новых методов прогнозирования позволит определять и планировать направления и объем контрольных и превентивных мероприятий. Учитывая то, что эти методы являются для специалистов противотуберкулезных организаций новыми, в методических рекомендациях изложены все детали их применения.

## Литература

- 1 Allard R. Use of time-series analysis in infectious disease surveillance. *Bull World Health Organ.* 1998;76(4):327-33.
- 2 Andersen, A., and Weiss, A., (1984) "Forecasting: The Box-Jenkins Approach" in *The Forecasting Accuracy of Major Time Series Methods*, Makridakis et. al, John Wiley & Sons,
- 3 Bell, W.R. and Hillmer, M. (1983). Modeling Time Series with Calendar Variation, *Journal of the American Statistical Association*, 78, 526-534.
- 4 W. S. Wei, "Time Series Analysis Univariate and Multivariate Methods", Addison Wesley Publishing Company, USA, 1994.
- 5 Hyndman, R.J. and Khandakar, Y. (2008) "Automatic time series forecasting: The forecast package for R", *Journal of Statistical Software*, 26(3).
- 6 Permanasari AE, Rambli DR, Dominic PD. Performance of univariate forecasting on seasonal diseases: the case of tuberculosis. *Adv Exp Med Biol.* 2011;696:171-9.
- 7 Chichester, England. Mabaera B, Naranbat N, Katamba A, Laticevschi D, Lauritsen JM, Rieder HL. Seasonal variation among tuberculosis suspects in four countries. *Int Health.* 2009 Sep;1(1):53-60.
- 8 Nagayama N, Ohmori M. Seasonality in various forms of tuberculosis. *Int J Tuberc Lung Dis.* 2006 Oct;10(10):1117-22. PubMed PMID: 17044204.
- 9 Kumar V, Singh A, Adhikary M, Daral S, Khokhar A, Singh S. Seasonality of tuberculosis in Delhi, India: a time series analysis. *Tuberc Res Treat.* 2014;2014:514093.
- 10 Willis MD, Winston CA, Heilig CM, Cain KP, Walter ND, MacKenzie WR. Seasonality of tuberculosis in the United States, 1993-2008. *Clin Infect Dis.* 2012 Jun;54(11):1553-60.
- 11 Moosazadeh M, Nasehi M, Bahrapour A, Khanjani N, Sharafi S, Ahmadi S. Forecasting tuberculosis incidence in iran using box-jenkins models. *Iran Red Crescent Med J.* 2014 May;16(5):e11779.