УДК 51-76
doi: 10.18101/2304-5728-2017-1-78-85

ARIMA-МОДЕЛЬ ПУЛЬСОВОГО СИГНАЛА

© Раднаев Базар Баирович
магистр Института математики и информатики
Бурятский государственный университет
Россия, 670000, Улан-Удэ, ул. Смолина, 24а
E-mail: bakumandiamond@gmail.com

© Цыбиков Анатолий Сергеевич
кандидат педагогических наук
заведующий кафедрой информационных технологий
Бурятский государственный университет
Россия, 670000, Улан-Удэ, ул. Смолина, 24а
E-mail: cas313@rambler.ru

© Хабитуев Баир Викторович
старший преподаватель кафедры информационных технологий
Бурятский государственный университет
Россия, 670000, Улан-Удэ, ул. Смолина, 24а
E-mail: bairinc0@mail.ru

В статье рассматривается один из подходов моделирования пульсовой волны человека, представленного в виде временного ряда, по методологии Бокс-Дженкинса. Построена ARIMA-модель (модель авторегрессии и проинтегрированного скользящего среднего) сфигмограммы пульсовой волны лучевой артерии человека. Модели данного типа могут иметь практическое применение в области функциональной диагностики.
***Ключевые слова*:** моделирование пульсовой волны, ARIMA, модель авторегрессии и скользящего среднего.

Введение

В течение многих лет, восточная медицина доказывала высокую эффективность в лечении различных хронических заболеваний. Наиболее перспективной областью в восточной биометрии считается математический анализ пульсового сигнала (сфигмограммы), как главного источника информации о функциональном состоянии организма человека [1].

Анализ и моделирование пульсового сигнала позволяет получить множество информативных параметров, интерпретируемых с точки зрения восточной медицины.

Рассмотренны известные попытки математического моделирования пульсовой волны такие как метод сплайн-аппроксимации, основанный на применении интерполирующих сплайнов разных порядков – кубических и локальных В-сплайнов, модель Акулова, представляющая собой произведение экспоненциальной и тригонометрической функции с тремя параметрами, модель Самарского, основанная на законе сохранения энергии и импульса и др. Каждая из этих моделей и подходов имеют свои достоинства и недостатки, и все же остаются далеко не совершенными в виду сложности данного биофизического явления в организме человека [2].

Одним из перспективных подходов статистического моделирования пульсового сигнала, представленного в виде временного ряда, является применение методологии Бокс-Дженкинса (ARIMA-модель авторегрессии и проинтегрированного скользящего среднего) [3, 4, 5].

Математическое описание

Временной (динамический) ряд – последовательность наблюдений некоторого признака X в последовательные моменты t. Отдельные наблюдения называются уровнями ряда и обозначаются x_t, $t=1,..., n$.

ARIMA-модель представляет собой синтез двух различных методов моделирования временного ряда: скользящего среднего и авторегрессии [3].

Модель скользящего среднего q-го порядка $MA(q)$ временного ряда X_t имеет вид:

$$X_t = \sum_{j=1}^{q} b_j \varepsilon_{t-j} + \varepsilon_t \qquad (1)$$

где ε_t – белый шум, b_j – параметры модели. Модель содержит $q+1$ параметров (b_1, b_2, ..., b_q, и σ^2), значения которых оценивается по принципу максимального правдоподобия.

Модель авторегрессии *p-го порядка* $AR(p)$ временного ряда X_t имеет несколько иной вид:

$$X_t = c + \sum_{i=1}^{p} a_i X_{t-i} + \varepsilon_t \qquad (2)$$

где ε_t – белый шум, a_i – параметры модели, c – константа. Модель содержит $p+1$ параметров (a_1, a_2, ..., a_p, и σ^2), значения которых, как правило, оценивается по методу наименьших квадратов.

Итак, объединенив две вышеуказанные модели получаем модель авторегрессии и скользящего среднего порядка (p, q) - $ARMA(p, q)$:

$$X_t = c + \sum_{i=1}^{p} a_i X_{t-i} + \sum_{j=1}^{q} b_j \varepsilon_{t-j} + \varepsilon_t \qquad (3)$$

которая содержит p+q+1 параметров.

В случае если наблюдаемый ряд X_t имеет признаки нестационарности (например, какие-либо детерминированные тренды – полиномиальный, линейный и т.д.), то модель не может является адекватной. Тем не менее в этом случае, некоторая разность наблюдаемого процесса порядка d, может оказаться стационарной: $\Delta^d X_t$, где Δ оператор разности, $\Delta X = X_t - X_{t-1}$ – разность первого порядка (аналог дифференцирования), Δ^d – последовательное взятие первой разности d раз. Теперь для описания полученного процесса $\Delta^d X_t$ уже можно эффективно применить *ARMA* – модель. В итоге, получили модель авторегрессии и проинтегрированного скользящего среднего порядка. При этом p – параметр *AR*-части, d – степень интеграции, q – это параметр *MA*-части. *ARIMA(p, d, q)* (от английского — *«Auto-Regressive Integrated Moving Average»*):

$$\Delta^d X_t = c + \sum_{i=1}^{p} a_i X_{t-i} + \sum_{j=1}^{q} b_j \varepsilon_{t-j} + \varepsilon_t \qquad (4)$$

Алгоритм построения

I этап. Идентификация параметров.

1. Установить порядок разности интеграции d, то есть добиться стационарности ряда, взяв достаточное количество последовательных разностей $(X_t = X_t - X_{t-1})$. Порядок d устанавливается исходя из поведения автокорреляционной функции ряда.

2. К полученному стационарному ряду Y_t, подбираем *ARMA(p,q)*. Исходя из поведения автокорреляционной и частной автокорреляционной функции, устанавливаем параметры p и q. Если ряд имеет сезонную составляющую с определенным периодом, в модель необходимо внести сезонную корректировку. Сезонные модели *ARIMA* являются обобщением обычных моделей *ARIMA*. Полная сезонная модель может быть представлена в виде *ARIMA(p, d, q)(Ps, Ds, Qs)* где к параметрам модели добавлены сезонные параметры: сезонный параметр авторегрессии – *Ps*, сезонная разность – *Ds*, сезонный параметр скользящего среднего – *Qs*.

II этап. Оценивание параметров.

С помощью специальных численных процедур по известным данным на данном этапе оцениваются коэффициенты $a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q$ и σ^2 при условии, что мы уж знаем p и q. Оценка значений параметров проводится на основе метода наименьших квадратов и принципа максимального правдоподобия.

III этап. Проверка адекватности модели.

Исходной информацией для анализа адекватности модели служат остатки (в данном случае остаточные ошибки – это предполагаемые значения ε_t). Проверяем качество модели при помощи критерия AIC. Мы предполагали, что ε_t является белым шумом, поэтому, проверяем некоррелированность остатков.

Кроме того существуют более формализованные критерии, например, критерий предложенный Акаике *AIC* для модели *ARMA(p,q)* выглядит следующим образом:

$$AIC(p,q) = \ln\sigma^2 + 2\frac{p+q}{T} \tag{5}$$

$$\sigma^2 = \frac{RSS}{T-p-q}$$

где T – число наблюдений, RSS – остаточная сумма квадратов.

Выбор подкласса моделей производится на основе сравнение показателей качества с различными значениями параметров p и q.

IV этап

Практическое применение модели. Анализ различных путей модификации модели, в том числе необходимости предобработки исходного сигнала.

Модель пульсового сигнала

Для реализации данного подхода на эмпирических данных выбран реальный пульсовой сигнал человека, полученного с помощью специального аппаратно-программного комплекса (рис.1.).

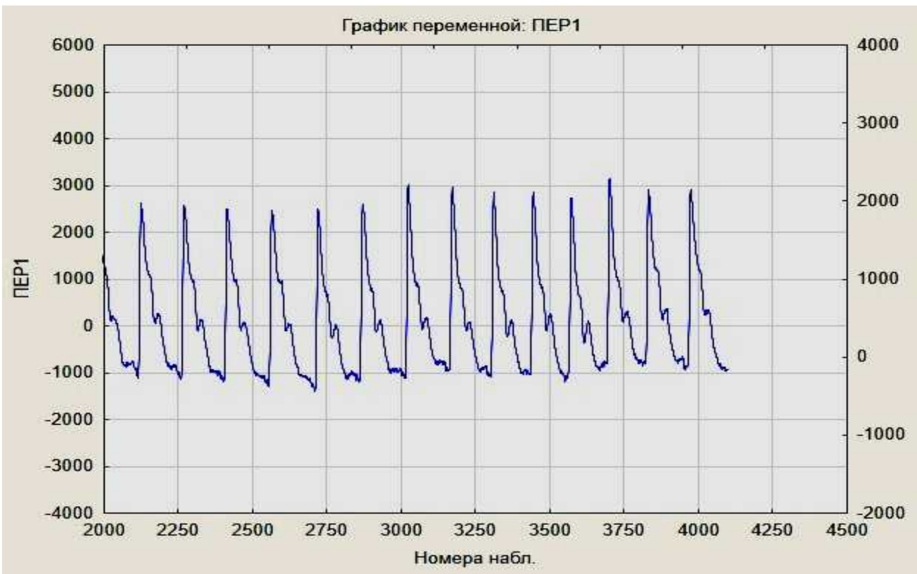

Рис. 1. График сфигмограммы пульсовой сигнала

Идентификация и оценка параметров модели производится с применением программного пакета STATISTICA 10. С помощью модуля «*Time Series Analysis*» производится оценка параметров модели по принципу максимального правдоподобия, качественный и количественный анализ автокорреляционных и частных автокорреляционных функций, анализ остатков, критериальная и визуально-графическая оценка адекватности модели, а также прогнозирование. В результате проведенных расчетов и экспериментов в рамках методологии получено три класса моделей реальной пульсовой волны, отвечающих требованиям адекватности(Табл. 1).

Таблица 1.

Классы моделей и их характеристики качества

№	Модель	AIC	W	RSS
1	ARIMA(2, 1, 0)(1, 1, 0)	8,0334239058	0,87201	12165760
2	ARIMA(1, 1, 0)(1, 1, 0)	8,1158696188	0,86185	13211282
3	ARIMA(2, 0, 0)(1, 1, 0)	8,1996723144	0,87821	14639777

Исходя из количественных характеристик, анализа остатков (в т.ч. остаточной коррелограммы) и полученного прогноза определена модель под №1 с лагом 146 как наиболее адекватная среди рассматриваемых.

Итак, выбранная модель *ARIMA*(2, 1, 0)(1, 1, 0) с оцененными параметрами приобретает следующую аналитическую форму:

$$\Delta^1 X_t = c + 1{,}33699646176872\left(\Delta^1 X_{t-1}\right) - 0{,}5328006151115\left(\Delta^1 X_{t-2}\right) -$$
$$-0{,}525157536913316\left(\Delta^1 (146) X_{t-2}\right) + \varepsilon_t$$

$$AIC(2,0) = \ln \frac{12165760}{9353-2} + 2\frac{2}{3953} = 8{,}0334239058$$

Графики остатков и остаточной автокорреляционной функции приведены ниже (рис. 2, 3).

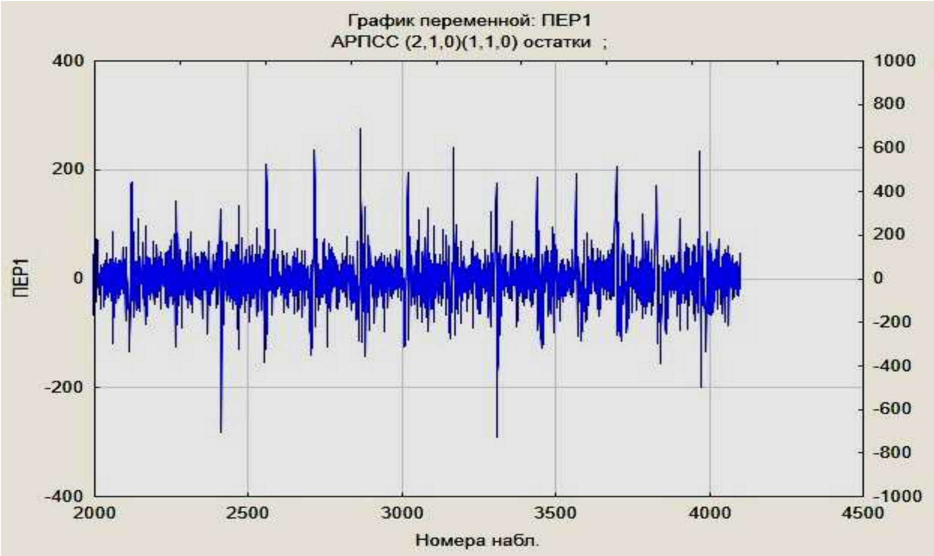

Рис. 2. График остатков *ARIMA*(2, 1, 0)(1, 1, 0)

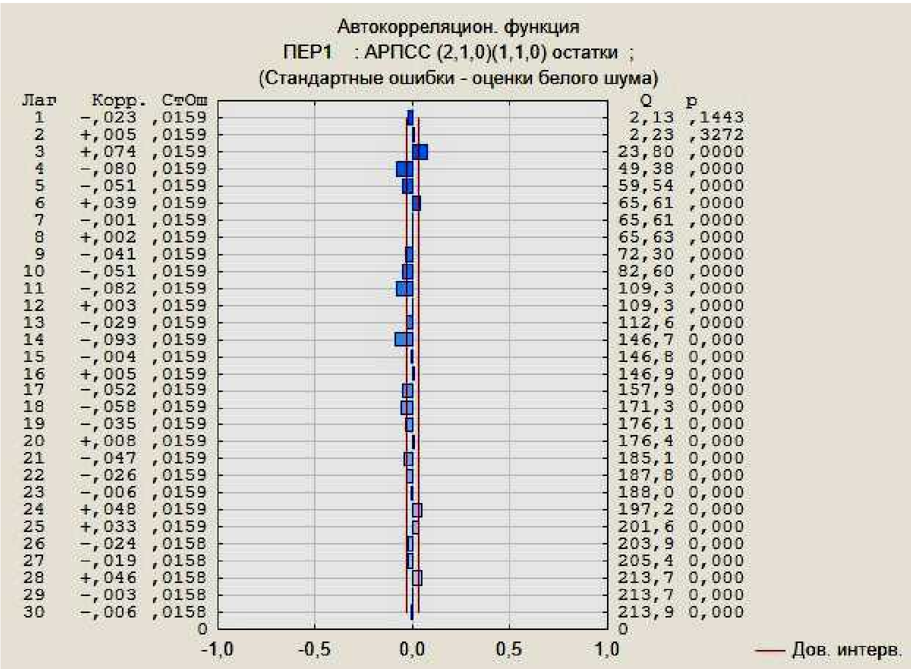

Рис. 3. Коррелограмма остатков *ARIMA*(2, 1, 0)(1, 1, 0)

На основе полученной модели построен прогноз на два сердечных цикла пульсового сигнала (рис. 4). График прогноза считаем вполне удовлетворительным.

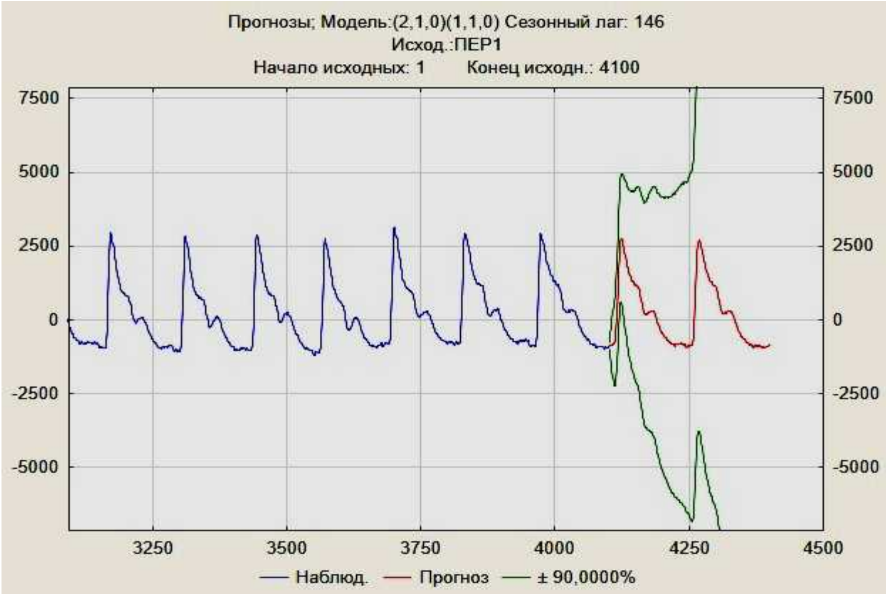

Рис. 5. Прогноз *ARIMA*(2, 1, 0)(1, 1, 0)

Заключение

Таким образом, в данной работе мы применили методологию Бокс-Дженкинса для моделирования дискретного пульсового сигнала человека. Построены адекватные модели ARIMA с применением пакета STATISTICA 10. Получен прогноз на несколько реализаций сердечного цикла.

Дальнейшее логическое продолжение исследования носит экспериментальный характер. Необходимо выявление взаимосвязей полученных параметров с различными физиологическими состояниями и типологическими особенностями организма человека.

Литература

1. Бороноев В. В. Пульсовая диагностика заболеваний в тибетской медицине: физические и технические аспекты. — Улан-Удэ: Изд-во БНЦ СО РАН, 2005. — 250 с.

2. Дармаев Т. Г., Цыбиков А. С., Хабитуев Б. В. Математическое моделирование пульсовых волн на основе теории солитонов и уравнения Кортевега Де Фриза // Вестник Бурятского государственного университета. — 2014. — Вып. 9(1). — С. 35 – 39.

3. Бокс Дж., Дженкинс Г. Анализ временных рядов прогноз и управление / Под ред. В. Ф. Писаренко. — Москва: Мир, 1974. — Кн. 1. — 406 с. — Кн. 2. — 197 с.

4. Безручко Б. П., Смирнов Д. А. Статистическое моделирование по временным рядам. — Учебно-методическое пособие. — Саратов: Изд-во ГосУНЦ «Колледж», 2000. — 23 с.

5. Халафян А. А. Статистический анализ данных. — 3-е изд. — Москва: ООО «Бином-Пресс», 2007. — 512 с.

ARIMA-MODEL OF PULSE WAVE

Bazar B. Radnaev
Master, Institute of Mathematics and Informatics
Buryat State University
24a Smolina St., Ulan-Ude 670000, Russia

Anatoliy S. Tsybikov
Cand. Sci. (Education), Department of Information Technologies
Buryat State University
24a Smolina St., Ulan-Ude 670000, Russia

Bair V. Khabituev
Senior Lecturer, Department of Information Technologies
Buryat State University
24a Smolina St., Ulan-Ude 670000, Russia

The article deals with one of the approaches to human pulse wave modeling represented in the form of time series according to the Box-Jenkins method. We have constructed the ARIMA-model (autoregressive integrated moving average model) of the sphygmogram of human radial artery pulse wave. Models of this type may have practical application in the field of functional diagnostics.
Keywords: pulse wave modeling, ARIMA, autoregressive integrated moving average model.