УДК 519.233.5:519.87

АНАЛИЗ ВРЕМЕННЫХ РЯДОВ ПОВЕДЕНИЯ СЛОЖНЫХ ДИНАМИЧЕСКИХ СИСТЕМ

В. С. Микшина, Н. Б. Назина, Л. А. Денисова
Сургутский государственный университет, mikshinavs@gmail.com

В работе представлены результаты исследования биомедицинских данных, а также данных, получаемых с датчиков в области нефтегазовой промышленности, методами математической статистики. В качестве биомедицинских данных, использовались показатели кардиоинтервала сердечного ритма, а в качестве технических данных использовались показатели датчиков водяной трубы. В исследовании определялась корреляционная зависимость между последовательными уровнями временного ряда, со сдвинутыми во времени рядами данных. Проведен анализ автокорреляционной функции и коррелограммы.

Ключевые слова: кардиоинтервал, показания счетчика расхода сетевой воды, временной ряд, коэффициент автокорреляции, коррелограмма.

TIME SERIES ANALYSIS OF COMPLEX DYNAMIC SYSTEMS BEHAVIOR

V. S. Mikshina, N. B. Nazina, L. A. Denisova
Surgut State University, mikshinavs@gmail.com

The paper presents the results of a study of biomedical data, as well as data obtained from receivers using mathematical statistics in the oil and gas industry. As biomedical data, the RR intervals were used for cardiac rhythm, and the parameters of water pipe sensors were used as characteristic data. The study determined the correlation dependence between progressive levels of the time series with time-shifted data series. An analysis of the autocorrelation function and the correlogram is carried out.

Ключевые слова: RR interval, meter reading of delivery water supply rate, time series, autocorrelation coefficient, correlogram.

Введение

На сегодняшний день исследование сложных динамических систем является актуальной и трудоемкой задачей, так как в таких системах состояние принимает новое значение в каждый момент времени, отличное от предыдущего. Определить положение вектора состояния системы в следующий момент времени затруднительно. Если представить вектор состояния системы $x = (x_1, x_2, \dots, x_m)^T$ в многомерном фазовом пространстве состояний, то конкретное значение данного вектора x_i не несет информационной значимости, так как в следующий момент времени эта точка сместится. Таким образом, вектор состояний системы совершает непрерывное хаотическое движение в фазовом пространстве состояний [1, 5, 6].

Сложность заключается так же в том, что для однотипных временных рядов функция распределения случайной величины параметров сложных динамических систем, вероятно, будет различна. Обязательным условием детерминистско-стохастического подхода является неоднократное воспроизведение начального состояния системы x в момент времени t_0, наличие возможности стационарных режимов и точек покоя. С точки зрения детерминистского подхода, многократное повторение процесса обеспечивает идентификацию модели динамической системы в фазовом пространстве состояний, а в стохастическом подходе – статистической функции распределения. Но возможность получения функции распределения для динамики поведения любой стохастической системы не снимает проблемы определенности конечного состояния динамических систем в фазовом пространстве состояний [1, 5, 6].

Для построения математической модели процесса необходимо определить, какой подход использовать: детерминистско-стохастический или хаотический. Для этого проводятся исследования поведения параметров сложных динамических систем, которые эти процессы описывают [3]. К таким исследованиям относится анализ временных рядов [4]. Используя представление о характеристике объекта исследования, математическую модель можно построить, имея данные, характеризующие один объект, взяв их в качестве ряда последовательных моментов времени. При этом будем понимать, что на временной ряд оказывают влияние составляющие: формирующие тенденцию ряда (возрастающая тенденция), циклические колебания ряда (сезонные) и случайные компоненты (хаотичность) [3].

Целью данной работы является выявление корреляционной зависимости между уровнями временных рядов двух сложных динамических систем; построение автокорреляционной функции, которая характеризует структуру временного ряда, а также правильно подобрать тип функции распределения, которая теоретически опишет объект исследования.

Объект и методы исследования

Объектом настоящего исследования являются параметры состояния сложных динамических систем, на примере, биомедицинской системы (параметр сердечно-сосудистой системы – кардиоритм), а также газовая котельная установка нефтегазодобывающего комплекса (показатель датчика сетевой воды) [3, 9].

Кардиоинтервалы сердечно-сосудистой системы показывают ритм работы сердца [1, 2]. Данные кардиоритма были получены с помощью пульсоксиметра ЭЛОКС-01 М, регистрирующего пульсовую волну с одного из пальцев испытуемого в положении сидя.

Основным параметром котельной установки нефтегазодобывающего комплекса выступил показатель датчика сетевой воды [9].

На рис. 1, 2 представлены примеры сигнала кардиоритма, а так же показателя сигнала датчика сетевой воды.

В процессе исследований было выяснено, что при наличии основных компонентов временного ряда, значения каждого последующего уровня ряда зависят от предыдущих значений, полученных в эксперименте. Корреляционную зависимость между последовательными уровнями временного ряда рассчитывали с помощью коэффициента автокорреляции уровней ряда. Коэффициент автокорреляции показывает тесноту связи текущего и предыдущего уровней ряда и варьируется в диапазоне от -1 до 1. Чем ближе значение к -1 или 1, тем связь более сильная. Чем ближе к 0, тем связь слабее.

$$r_{x_1 x_2} = \frac{\sum(x_1 - \overline{x_1}) * (x_2 - \overline{x_2})}{\sqrt{\sum(x_1 - \overline{x_1})^2 * \sum(x_2 - \overline{x_2})^2}} \qquad (1)$$

где x_1 – значение временного ряда;
x_2 – значение сдвинутого временного ряда;
$\overline{x_1}$ и $\overline{x_2}$ – среднее значение временных рядов.

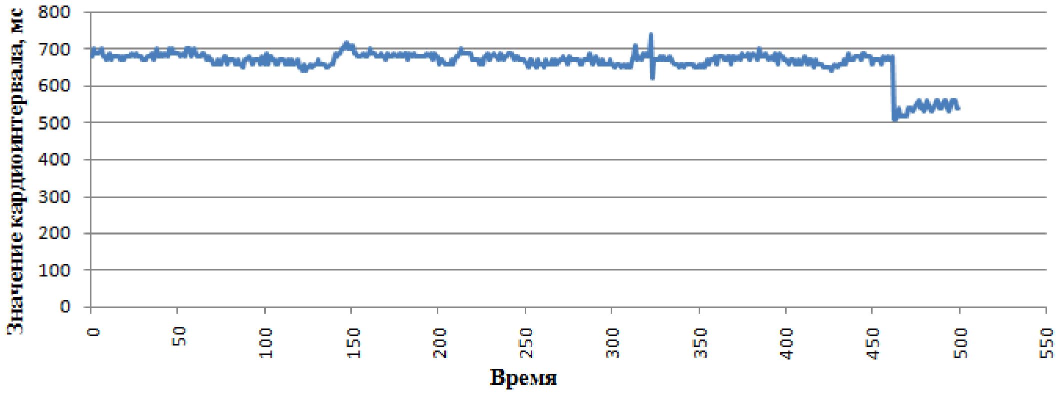


Рис. 1. Пример сигнала кардиоритма

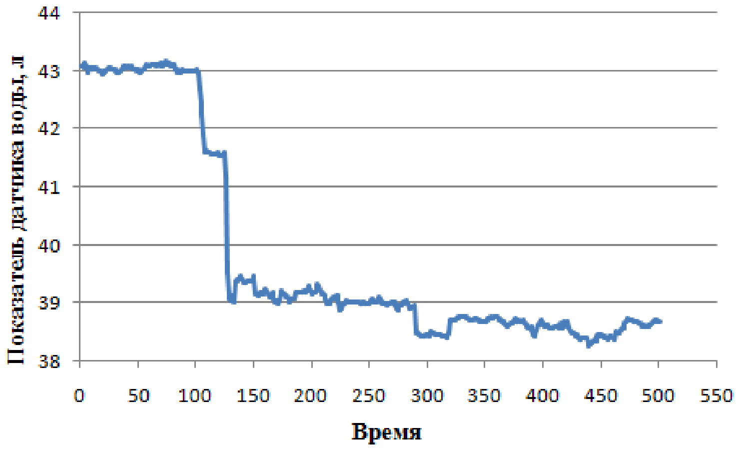


Рис. 2. Пример показателя сигнала датчика сетевой воды

Число периодов, по которым рассчитывается коэффициент автокорреляции, называется лагом (τ). С увеличением лага число пар значений, параметров уменьшается. Для обеспечения статистической достоверности коэффициентов автокорреляции было принято использование правила «максимальный лаг не должен превышать $n/4$» [8]. Количество сдвигов временных рядов в данной работе составило 50 ($\tau = 50$). Общее количество точек данных для каждого ряда составило 500 ($n = 500$) Результаты расчета коэффициента автокорреляции данных кардиоритма, а так же данных показателя счетчика сетевой воды, и соответствующие этим данным лаги представлены в табл. 1.

Таблица 1

Значение коэффициента корреляции временных рядов

Значение коэффициента корреляции для временного ряда биомедицинской системы				Значение коэффициента корреляции для временного ряда для технической системы			
№ лага,τ	Коэффициент корреляции, r	№ лага, τ	Коэффициент корреляции, r	№ лага,τ	Коэффициент корреляции, r	№ лага, τ	Коэффициент корреляции, r
1	0,91	65	0,04	1	0,99	65	0,48
5	0,80	70	0,03	5	0,97	70	0,44
10	0,65	75	0,00	10	0,94	75	0,39
15	0,54	80	-0,01	15	0,91	80	0,35
20	0,41	85	-0,03	20	0,88	85	0,31
25	0,28	90	-0,04	25	0,84	90	0,26
30	0,18	95	-0,05	30	0,79	95	0,22
35	0,09	100	-0,06	35	0,75	100	0,17

Окончание табл. 1

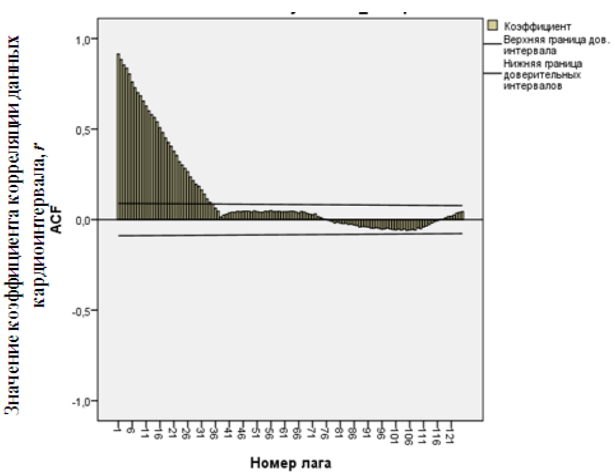
Значение коэффициента корреляции для временного ряда биомедицинской системы				Значение коэффициента корреляции для временного ряда для технической системы			
№ лага,τ	**Коэффициент корреляции, *r***	**№ лага, τ**	**Коэффициент корреляции, *r***	**№ лага,τ**	**Коэффициент корреляции, *r***	**№ лага, τ**	**Коэффициент корреляции, *r***
40	0,03	105	-0,05	40	0,71	105	0,13
45	0,04	110	-0,05	45	0,66	110	0,10
50	0,05	115	-0,01	50	0,63	115	0,07
55	0,05	120	0,02	55	0,58	120	0,04
60	0,04	125	0,04	60	0,53	125	0,01

Как видно из табл. 1, для данных биомедицинской системы, с увеличением лага значение коэффициента корреляции существенно уменьшается, что говорит о слабой связи между данными временного ряда, т. е. о наличии случайной компоненты. Во временном ряде при лаге τ =1…10 прослеживается высокая корреляция r = 0,6, что говорит о наличии тенденции. Так же с увеличение величины лага, увеличивается количество отрицательных значений. Одним из свойств коэффициента автокорреляции является то, что по его знаку невозможно определить наличие возрастающей или убывающей тенденции в уровнях временного ряда. Для данных временных рядов характерно нелинейное поведение, так как значения автокорреляции первого и последующего уровней отличаются относительно друг от друга.

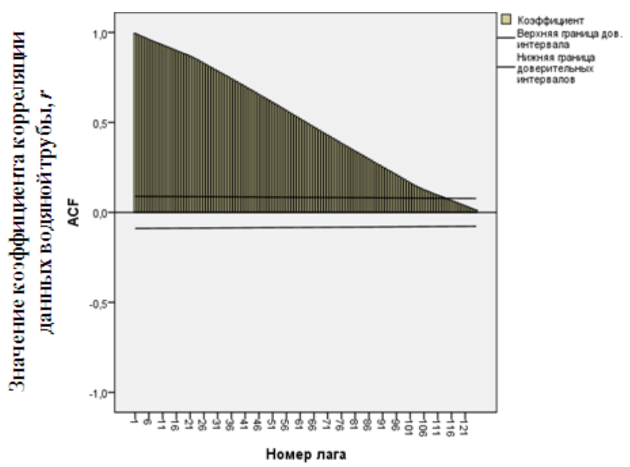
Для данных технической системы можно заметить, что с увеличением уровня лага уменьшается коэффициент корреляции, но не так существенно, как для биомедицинских данных. Во временном ряде при лаге τ = 1…50 прослеживается высокая корреляция r = 0,6, что говорит о наличии тенденции, следовательно, что связь между временными рядами сильная.

Далее был проведен анализ автокорреляционной функции и коррелограммы позволяющей определить лаг, при котором автокорреляция определяется как высокая либо слабая, и, следовательно, связь между текущим и последующим уровнями ряда более или менее тесная. То есть при помощи коррелограммы можно выявить линейность или хаотичность в наборе данных [2, 3].

На рис. 2 *а, б* представлены примеры коррелограмм для данных биомедицинской и технической системы.



а



б

Рис. 2. Коррелограмма «Расчет коэффициента корреляции от величины лага данных»:
а – биомедицинской системы; *б* – технической системы

На рис. 2*а* привалируют значения коэффицинтов автокорреляции близкие к нулю, что говорит о хаотичности процесса для данного временного ряда биомедицинской системы, в отличии от коррелограммы технической системы (рис. 2*б*), где временному ряду характерно линейное поведение.

На рис. 3*а*, *б* представлены примеры графиков аналитических функций и трендовых составляющих временных рядов биомедицинской и технической системы.

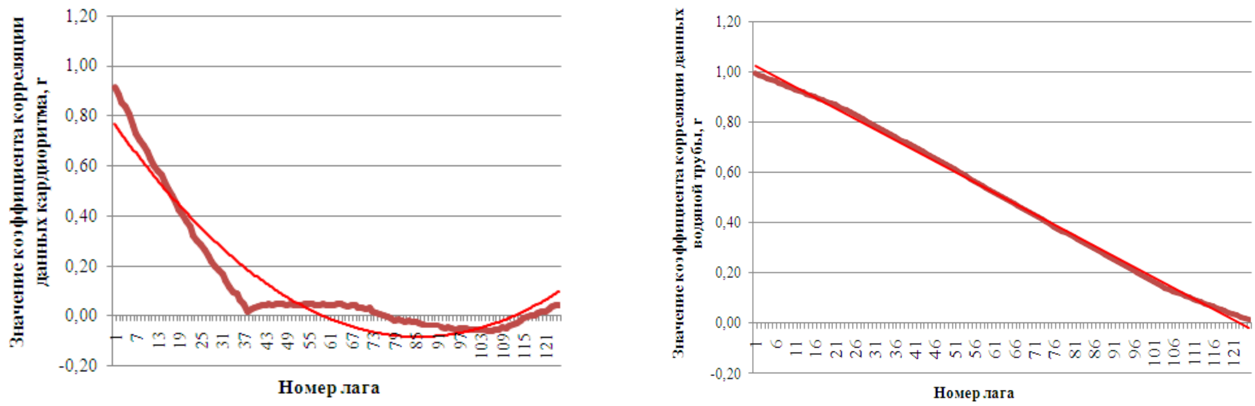
Для анализа тенденции временного ряда были получены математические модели в виде полиномиальных функций, характеризующих зависимость уровней ряда от величины лага. Адекватность моделей проверялась с помощью коэффициента детерминации R^2. Чем ближе значение коэффициента R^2 к 1, тем менее различаются экспериментальная и аналитическая функции [7, 10]. В табл. 2 представлены результаты аппроксимации автокорреляционной функции для временных рядов биомедицинской и технической системы, а также выявлены трендовые составляющие.

Таблица 2

Результаты аппроксимации автокорреляционной функции временных рядов

№	Название системы	Тип тренда	Уравнение	R²
1	Биомедицинская	Полиномиальный	$y = -0{,}0203x + 0{,}7887$	$R_1^2 = 0{,}92$
2	Техническая	Линейная	$y = -0{,}0084x + 1{,}0326$	$R_2^2 = 0{,}98$

Как видно из табл. 2 для обоих временных рядов характерен высокий коэффициент детерминации R^2, как для данных биомедицинской, так и для данных технической системы ($R_1^2 = 0{,}98$, $R_2^2 = 0{,}98$). Однако, для биомедицинской системы характерен полиномиальный тип тренда, в отличие от технической системы. На рис. 3*а, б* представлены примеры графиков аналитических функций и трендовых составляющих временных рядов биомедицинской и технической системы.



а *б*

Рис. 3. Результаты аппроксимации автокорреляционной функции временного ряда систем:
а – биомедицинской; б – технической

Таким образом, хаотический процесс, представленный временным рядом биомедицинской системы, может быть описан только полиномиальными функциями, а линейный процесс, на примере данных технической системы, может быть описан линейной функцией.

Вывод

Таким образом, в результате проведенных нами исследований было выявлено, что в зависимости от поведения временного ряда сложной динамической системы, возможно более грамотно подобрать математический аппарат для построения математических моделей процессов любой природы. Так, например, биомедицинские системы, характеризующие хаотический процесс, не могут исследоваться методами детерминистской-стохастического подхода в отличие от сложных технических систем.

Литература

1. Бетелин В. Б., Еськов В. М., Галкин В. А., Гавриленко Т. В. Стохастическая неустойчивость в динамике поведения сложных гомеостатических систем // Доклады академии наук. 2017. Т. 472. № 6. С. 642–644.

2. Григоренко В. В., Еськов В. М. Стохастический подход в анализе систем с хаотической динамикой на примере параметров сердечно-сосудистой системы // Мат-лы VI всерос. симпозиума с междунар. участием, посвященного 85-летию образования Удмурт. гос. ун-та. Ижевск, 2016. С. 111–115.

3. Григоренко В. В., Еськов В. М. Анализ временных рядов в исследовании процессов хаотической динамики // Естественные и технические науки. 2016. № 6 (96). С. 130–133.

4. Елисеева И. И. Эконометрика : учеб. СПб. : Финансы и статистика, 2003. 344 с.

5. Еськов В. М., Зинченко Ю. П., Филатов М. А., Еськов В. В. Эффект Еськова – Зинченко опровергает представления I. R. Prigogine, J A Wheeler и M. Gell-Mann о детерминированном хаосе биосистем – complexity // Вестник новых медицинских технологий. 2016. Т. 23. № 2. С. 34–43.

6. Еськов В. М., Хадарцев А. А., Козлова В. В., Филатов М. А. и др. Системный анализ, управление и обработка информации в биологии и медицине // Системный синтез параметров функций организма жителей Югры на базе нейрокомпьютинга и теории хаос-самоорганизации в биофизике сложных систем. Т. XI. Самара : Офорт, 2014. 192 с.

7. Заикин П. В., Погореловский М. А., Микшина В. С. Аппроксимация эмпирических функций полиномами высших порядков // Вестн. кибернетики. 2015. № 4 (20). С. 129–134.

8. Канторович Г. Г. Анализ временных рядов. // Экономический журнал ВШЭ. 2002. № 1. С. 85–116.

9. Максимюк Е. В., Микшина В. С. Математическое моделирование для поддержки принятия решений в области обеспечения энергетической эффективности // Качество. Инновации. Образование. 2014. № 8 (111). С. 54–63.

10. Погореловский М. А., Микшина В. С., Назина Н. Б., Заикин П. В. Агрегирование состава многокомпонентной смеси в математическом моделировании сложных динамических процессов // Север России: стратегии и перспективы развития : материалы III Всерос. науч.-практ. конф. Сургут, 2017. С. 127–133.