ISSN 2226-3780

2. Гордієнко, І. А. Дослідження метрологічних характеристик вологомірів природного газу різних типів в умовах експлуатації на об'єктах ДК «Укртрансгаз» [Текст] : зб. наук. пр. / І. А. Гордієнко, А. І. Лур'є, В. М. Козій, А. Г. Івков та ін. // Питання розвитку газової промисловості України. — Харків: УкрНДІгаз, 2006. — Вип. XXXIV. — С. 187—195.
3. Швейкін, О. Приклад інструментального визначення температури початку утворення кристалогідратів в природному газі [Текст] / О. Швейкін // Питання розвитку газової промисловості України. — Харків, 2009. — С. 131—133.
4. Швейкін, О. Л. Інструментальне визначення температур утворення рідкої та твердої фази компонентів природного газу в автоматичному режимі [Текст] / О. Швейкін // Метрологія та прилади. — № 4. — 2008. — С. 37—39.
5. Иберла, К. Факторный анализ [Текст] / К. Иберла. — М.: Статистика, 1980. — 380 с.
6. Harman, H. H. Modern factor analysis [Text] / H. H. Harman. — Chicago, 1976. — 508 p.
7. Ким, Дж. О. Факторный, дискриминантный и кластерный анализ [Текст] / Дж. О. Ким, Ч. У. Мьюллер. — М.: Финансы и статистика, 1989. — 215 с.
8. Шуметов, В. Г. Факторный анализ: подход с применением ЭВМ [Текст] / В. Г. Шуметов, Л. В. Шуметова. — Орел: ОрелГТУ, 1999.
9. Боровиков, В. П. Statisitica: искусство анализа данных на компьютере [Текст] / В. П. Боровиков. — 2-е изд. — С-Пб.: Питер, 2003 — 688 с.
10. Хвостова, О. В. Вологометрія природного газу [Текст] / О. В. Хвостова, А. Й. Лур'є, О. Л. Швейкін. — Харків: Курсор, 2011. — С. 74—80.

СТАТИСТИЧЕСКОЕ ИССЛЕДОВАНИЕ СИГНАЛОВ СИСТЕМЫ ОПРЕДЕЛЕНИЯ ПОКАЗАТЕЛЕЙ КАЧЕСТВА ПРИРОДНОГО ГАЗА

В работе проведен факторный анализ сигналов измерительной системы. В частности, получены описательные статистики, корреляционная матрица, матрица факторной нагрузки и ее геометрическая интерпретация. Была получена наглядная структура факторов, которые влияют на процесс измерения. Доказана корректность исследования качественных показателей природного газа, которые выполняются измерительной системой с помощью изучения влияния этих двух факторов.

Ключевые слова: факторный анализ, корреляционная матрица, влажность природного газа, статистические характеристики, фактор.

Швейкін Олександр Леонідович, *кандидат технічних наук, старший науковий співробітник, Український науково-дослідний інститут природних газів, Харків, Україна,* ***e-mail: shveykin_al@meta.ua****.*
Прокопенко Олена Олександрівна, *кандидат технічних наук, доцент, кафедра систем управління, Українська інженерно-педагогічна академія, Харків, Україна,* ***e-mail: digaz@i.ua****.*

Швейкин Александр Леонидович, *кандидат технических наук, старший научный сотрудник, Украинский научно-исследовательский институт природных газов, Харьков, Украина.*
Прокопенко Елена Александровна, *кандидат технических наук, доцент, кафедра систем управления, Украинская инженерно-педагогическая академия, Харьков, Украина.*

Shveikin Alexandr, *Ukrainian Research Institute for Natural Gases, Kharkiv, Ukraine,* ***e-mail: shveykin_al@meta.ua****.*
Prokopenko Olena, *Ukrainian Engineering Pedagogics Academy, Kharkiv, Ukraine,* ***e-mail: digaz@i.ua***

УДК 519.246

Щелкалин В. Н.

МОДЕЛЬ АВТОРЕГРЕССИИ — СПЕКТРАЛЬНО ПРОИНТЕГРИРОВАННОГО СКОЛЬЗЯЩЕГО СРЕДНЕГО

В статье рассматривается развитие метода Бокса-Дженкинса, основанное на совместном использовании идей методов «Гусеница»-SSA и Бокса-Дженкинса. Предложена модель авторегрессии — спектрально проинтегрированного скользящего среднего, реализующая трендовый подход, который заключается в моделировании процесса как отклонения фактических значений относительно трендовой составляющей, в роли которой выступает линейная рекуррентная формула метода «Гусеница»-SSA.

Ключевые слова: *прогнозирование временных рядов, структурная идентификация, метод Бокса-Дженкинса, метод «Гусеница»-SSA.*

1. Введение

В методе Бокса-Дженкинса уделено особое внимание проблеме выбора модели и ее оцениванию. В это методе используются идеи, что нестационарные временные ряды можно преобразовать в стационарные путем перехода от исходного временного ряда к его разностям соответствующего порядка $d = 1, 2$ и т. д. [1—6]. В [7] указывается, что игнорирование длинной памяти, когда она в действительности имеет место, приводит к более серьезному ухудшению результатов, чем ее наложение при отсутствии таковой.

2. Анализ литературных данных и постановка проблемы

Анализ литературы позволяет сделать вывод о том, что дальнейшим распространением моделей метода Бокса-Дженкинса для моделирования нестационарных временных рядов являются модели авторегрессии — дробно интегрированного скользящего среднего [8] и модели метода ОЛИМП.

Теория прогнозирования, которая рассматривается в данной статье, основана на методах Бокса-Дженкинса и «Гусеница»-SSA. Предлагается модель авторегрессии —

спектрально проинтегрированного скользящего среднего (АРСПСС).

3. Модель прогнозирования

Модель АРСПСС (p^+, L, r, q^*) можно записать следующим образом:

$$y_t = \frac{\theta^*_{q^*}(B)}{g'(B)\Phi^+_{p^+}(B)} a_t, \quad (1)$$

где $g'(B) = 1 - g(B)$, $g(B) = \sum_{i=1}^{L-1} g_{L-i} B^i$ — полином от оператора задержки B и его коэффициенты определяются, как и коэффициенты линейной рекуррентной формулы (ЛРФ) метода «Гусеница»-SSA; L — длина окна метода «Гусеница»-SSA, r — количество первых собственных троек сингулярного разложения траекторной матрицы в. р., отобранных на этапе группировки метода «Гусеница»-SSA [9]. Остальные обозначения модели (1) приведены в [4].

Удобство введения оператора задержки по времени B заключается в том, что во многих случаях с ним можно работать также, как с обыкновенным действительным или комплексным числом x. Так, например, для модели АРПСС (p, 1, q) имеем:

$$\varphi_p(B)(1-B)y_t = \theta_q(B)a_t; \quad \varphi_p(B)y_t = \frac{\theta_q(B)}{(1-B)} a_t.$$

Используя формулу геометрической прогрессии

$$\frac{1}{(1-x)} = 1 + x + x^2 + \ldots$$

получим:

$$\varphi_p(B)y_t = \theta_q(B)(a_t + a_{t-1} + a_{t-2} + \ldots).$$

Отсюда и в названии модели понятие — «проинтегрированного скользящего среднего».

Уравнения $\varphi_p(B) = 0$, $\theta_q(B) = 0$ называются характеристическими уравнениями авторегрессии и скользящего среднего соответственно. Корни характеристического уравнения $\varphi_p(B) = 0$ определяются из уравнений $1 - \varphi_1 B^{-1} - \varphi_2 B^{-2} - \ldots - \varphi_p B^{-p} = 0$ или $B^p - \varphi_1 B^{p-1} - \varphi_2 B^{p-2} - \ldots - \varphi_p = 0$ [3, 10]. Для того чтобы найти корни характеристического уравнения авторегрессии $B^p - \varphi_1 B^{p-1} - \varphi_2 B^{p-2} - \ldots - \varphi_p = 0$ представим его левую часть в виде разложения на множители:

$$\varphi_p(B) = (1 - h_1 B)(1 - h_2 B)\ldots(1 - h_p B).$$

Говорят, что характеристическое уравнение $\varphi_p(B) = 0$ имеет корень $B = h^{-1}$, если его разложение имеет множитель $(1 - hB)$. Аналогично определяются корни характеристического уравнения скользящего среднего. Алгебраическое уравнение $\varphi_p(B) = 0$ степени p имеет ровно p корней. Если все p корней характеристического уравнения различны, то общее решение уравнения $\varphi_p(B)y_t = 0$ имеет вид:

$$y_t = \sum_{i=1}^{p} c_i h_i^t,$$

где c_i — произвольные постоянные.

Если же какой-либо корень j имеет кратность m, то вместо слагаемого $c_j h_j^t$ в сумму войдет комбинация вида:

$$\left(c_j + t c_{j+1} + t^2 c_{j+2} + \ldots + t^{m-1} c_{j+m-1}\right) h_j^t.$$

Если среди корней содержатся пары комплексно-сопряженных чисел, то в сумму:

$$y_t = \sum_{i=1}^{p} c_i h_i^t$$

должна входить комбинация:

$$ch^t + c^* h^{*t} = |c||h|^t \left(e^{i(\omega t + \varphi)} + e^{-i(\omega t + \varphi)}\right) = 2|c||h|^t \cos(\omega t + \varphi),$$

где $c = |c| e^{i\varphi}$ и $*$ — знак комплексного сопряжения; $c^* = |c| e^{-i\varphi}$, ω — фаза комплексного числа [3].

В таких моделях минимальный по модулю корень характеристического полинома является индикатором поведения прогнозирующей функции при больших l (l — глубина прогноза). Прогнозирующая функция убывает экспоненциально, если все корни характеристического уравнения по модулю больше единицы. В случае, когда по крайней мере один, в том числе и кратный, корень по модулю равен единице, прогнозирующая функция ведет себя как $l^{m-1}\cos(\omega t + \varphi)$, m — кратность корня. В частном случае действительного корня $B = 1$ прогнозирующая функция ведет себя как l^{m-1}. Если минимальный корень по модулю равен единице, то такой ряд называется слабонестационарным [3]. Если хотя бы один корень по модулю меньше единицы, то прогнозирующая функция экспоненциально растет и может быть применима для прогнозирования нестационарных процессов. Такие процессы называются сильнонестационарными. Именно такие случаи могут быть рассмотрены в методе «Гусеница»-SSA.

Полином в (1) $g'(B)$ можно рассматривать как полином порядка L относительно B, поэтому $g'(B)$ можно представить в виде:

$$g'(B) = (1 - h_1 B) \times (1 - h_2 B)\ldots(1 - h_L B).$$

Для получения характеристических уравнений, как было сказано выше, требуется замена оператора сдвига назад на комплексную переменную. Тогда представление для обратного оператора можно получить разлагая $\left[g'(B)\right]^{-1}$ на элементарные дроби:

$$\left[g'(B)\right]^{-1} = \frac{k_1}{1 - h_1 B} + \frac{k_2}{1 - h_2 B} + \ldots + \frac{k_L}{1 - h_L B}.$$

Следовательно,

$$y_t = \left[g'(B)\right]^{-1} a_t = \sum_{i=0}^{L} \frac{k_i}{1 - h_i B} a_t;$$

$$\frac{k_i}{1 - h_i B} = k_i \sum_{n=0}^{\infty} (h_i B)^n;$$

$$y_t = \sum_{i=1}^{L} \left(k_i \sum_{j=0}^{\infty} h_i^j B^j\right) a_t.$$

Отсюда и происходит в названии понятие — «спектрально проинтегрированного скользящего среднего».

ISSN 2226-3780

Корреляционная функция процесса АРСПСС$(0,L,r,0)$ определяется следующим образом:

$$\gamma_0=\sum_{j=0}^{\infty}\left(\sum_{i=1}^{L}k_i g_i^j\right)^2\sigma_a^2,$$

$$\gamma_k=\left(-\left(\sum_{i=1}^{L}k_i\right)\left(\sum_{i=1}^{L}k_i g_i^k\right)+\right.$$
$$\left.+\sum_{j=1}^{\infty}\left(\left(\sum_{i=1}^{L}k_i g_i^j\right)\left(\sum_{i=1}^{L}k_i g_i^{k+j}\right)\right)\right)\sigma_a^2.$$

Тогда автокорреляционная функция имеет вид:

$$\rho_k=\frac{-\left(\sum_{i=1}^{L}k_i\right)\left(\sum_{i=1}^{L}k_i g_i^k\right)+\sum_{j=1}^{\infty}\left(\left(\sum_{i=1}^{L}k_i g_i^j\right)\left(\sum_{i=1}^{L}k_i g_i^{k+j}\right)\right)}{\sum_{j=0}^{\infty}\left(\sum_{i=1}^{L}k_i g_i^j\right)^2},$$

$k=1,2,\ldots$

4. Результаты применения предложенной модели прогнозирования

Продемонстрируем применение модели АРСПСС на примере прогноза временного ряда потребления природного газа (объемы ежедневного потребления (рис. 1)). Длина окна является основным параметром базового алгоритма SSA. Наиболее детальное разложение и лучшая разделимость детерминированной составляющей от шумовой достигается при выборе длины окна, приблизительно равной половине длины ряда $n=950$. Выберем $L=469$. На рис. 2 изображены единичная окружность и корни характеристического уравнения $g'(B)=0$. Анализ собственных значений траекторной матрицы временного ряда показал, что к детерминированной составляющей можно отнести первые пять собственных троек $r=5$ сингулярного разложения траекторной матрицы ряда. В результате получена модель:

$$g'(B)y_t=\frac{\left(1+0{,}103B^{11}\right)\left(1+0{,}052B^{21}\right)\left(1-0{,}288B^{23}\right)\left(1+0{,}072B^{56}\right)\left(1+0{,}257B^{365}\right)}{(1-B)\left(1-0{,}938B+0{,}088B^2\right)\left(1+0{,}046B^7\right)\left(1+0{,}047B^8\right)\left(1+0{,}195B^{11}\right)\left(1-0{,}135B^{23}\right)}a_t$$

и прогноз с глубиной $l=145$, представленный на рис. 1.

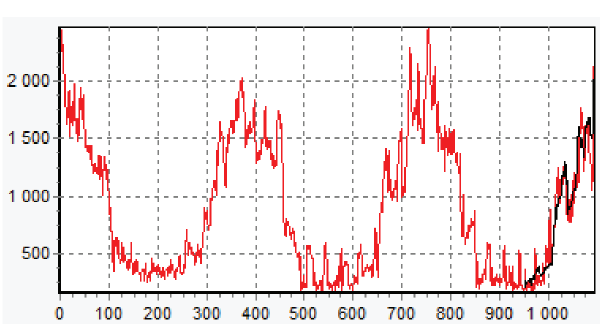

Рис. 1. Временной ряд потребления природного газа и прогноз

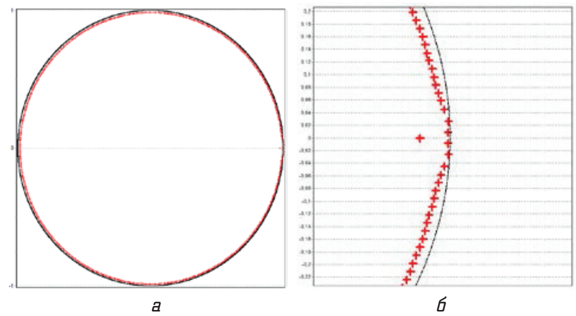

Рис. 2. Единичная окружность и корни характеристического уравнения $g'(B)=0$ (*а*), единичная окружность и корни характеристического уравнения $g'(B)=0$ вблизи точки (1, 0) (*б*)

4. Выводы

Похожий результат можно получить и моделями АРСС высокого порядка. Однако, по количеству параметров модель становится низкосортной согласно критериям Акайке и Шварца. Предложенная же модель АРСПСС идентифицирует долгосрочную зависимость процесса лишь двумя параметрами L и r.

Литература

1. Бокс, Дж. Анализ временных рядов, прогноз и управление [Текст] : пер. с англ. / Дж. Бокс, Г. Дженкинс; под ред. В. Ф. Писаренко. — М.: Мир, 1974. — Кн. 1. — 406 с.
2. Бокс, Дж. Анализ временных рядов, прогноз и управление [Текст] : пер. с англ. / Дж. Бокс, Г. Дженкинс; под ред. В. Ф. Писаренко. — М.: Мир, 1974. — Кн. 2. — 197 с.
3. Безрукова, Е. Г. Прогнозирование статистических временных рядов [Текст] : учеб. пос. / Е. Г. Безрукова, Е. А. Руденчик. — Ярославль: Яросл. гос. техн. ун-т, 1997. — 94 с.
4. Евдокимов, А. Г. Оперативное управление потокораспределением в инженерных сетях [Текст] / А. Г. Евдокимов, А. Д. Тевяшев. — Х.: Вища школа, 1980. — 144 с.
5. Седов, А. В. Моделирование объектов с дискретно-распределенными параметрами: декомпозиционный подход [Текст] / А. В. Седов. — Южный научный центр РАН. — М.: Наука, 2010. — 438 с.
6. Бэнн, Д. В. Сравнительные модели прогнозирования электрической нагрузки [Текст] : пер. с англ. / Д. В. Бэнн, Е. Д. Фармер. — М.: Энергоатомиздат, 1987. — 200 с.
7. Andersson, M. K. On the Effects of Imposing or Ignoring Long Memory When Forecasting [Text] / M. K. Andersson // Working Paper Series in Economics and Finance. — 1998. — T. 225. — P. 1–14.
8. Hosking, J. R. M. Fractional differencing [Text] / J. R. M. Hosking // Biometrika. — 1981. — Vol. 68, No. 1. — P. 165—176.
9. Голяндина, Н. Э. Метод «Гусеница»-SSA: прогноз временных рядов [Текст] : учеб. пособие / Н. Э. Голяндина. — СПб.: С.-Петербургский государственный университет, 2004. — 52 с.
10. Горелова, В. Л. Основы прогнозирования систем [Текст] : учеб. пособ. для инж.-экон. спец. вузов / В. Л. Горелова, Е. Н. Мельникова — М.: Высш. шк., 1986. — 287 с.

МОДЕЛЬ АВТОРЕГРЕСІЇ — СПЕКТРАЛЬНО ІНТЕГРОВАНОГО КОВЗНОГО СЕРЕДНЬОГО

У статті розглядається розвиток методу Бокса-Дженкінса, яке засноване на спільному використанні ідей методів «Гу-

сениця»-SSA і Бокса-Дженкінса. Запропоновано модель авторегресії — спектрально інтегрованого ковзного середнього, що реалізує трендовий підхід, який полягає в моделюванні процесу як відхилення фактичних значень відносно трендової складової, в ролі якої виступає лінійна рекурентна формула методу «Гусениця»-SSA.

Ключові слова: прогнозування часових рядів, структурна ідентифікація, метод Бокса-Дженкінса, метод «Гусениця»-SSA.

Щелкалин Виталий Николаевич, *ассистент, кафедра прикладной математики, Харьковский национальный университет радиоэлектроники, Украина,* ***e-mail: vitalii.shchelkalin@gmail.com****.*

Щелкалін Віталій Миколайович, *асистент, кафедра прикладної математики, Харківський національний університет радіоелектроніки, Україна.*

Shchelkalin Vitalii, *Kharkiv National University of Radioelectronics, Ukraine,* ***e-mail: vitalii.shchelkalin@gmail.com***

УДК 006.91:519.2

Шенгур С. В.

СТАТИСТИЧЕСКАЯ ОБРАБОТКА СВЕРХМАЛЫХ ВЫБОРОК СЛУЧАЙНЫХ УГЛОВ

Статья посвящена статистическому анализу выборок случайных углов очень малого объема. Основной целью работы является разработка и экспериментальное исследование новой методики статистической обработки угловых величин, полученных из различных источников и сфер деятельности. В основу положен метод бутстреп. Разработанная методика позволяет повысить точность оценивания выборочного кругового среднего и его доверительного интервала. Приведены результаты экспериментальных исследований.

Ключевые слова: *случайный угол, выборочные круговые характеристики, неопределенность, бутстреп.*

1. Введение

Результаты измерений в соответствии с национальными и международными стандартами [1, 2] должны сопровождаться показателями точности. В качестве таких показателей чаще всего используют расширенную неопределенность U либо доверительный интервал $\Delta_{д}$.

Распределение результата измерения случайной величины, основываясь на результатах предварительных измерений, как правило, принимают приближенным к гауссовскому [3]. При наличии достаточной статистики выполняют проверку на гауссовость по одному из известных критериев — Колмогорова, хи-квадрат.

Однако такие критерии предусматривают построение по данным наблюдения гистограммы, и, следовательно, имеют ограничения при применении к выборкам менее 35 значений. И даже при таком объеме данных имеют нестабильный результат. На практике не редко встречаются случаи, когда выборка, полученная по результатам наблюдений, не превышает 10 значений. Это обусловлено, прежде всего, высокой стоимостью проведения эксперимента либо невозможностью его повторного проведения. Общепринятого критерия сверхмалой выборки нет. Будем использовать этот термин при $n \leq 9$ [4]. Обработка сверхмалых выборок требует специфического подхода, за исключениями ситуаций, когда закон распределения случайного угла известен.

Традиционный метод расчета и представления результата измерений случайных угловых данных предусматривает получение по результатам измерений выборки углов определенного объема, определение выборочных круговых среднего и стандартного отклонения, оценки неопределенности измерений. Оценку неопределенности рассчитывают как симметричный относительно выборочного кругового среднего интервал значений углов, величина которого формируется как удвоенное произведение выборочного кругового стандартного отклонения и фактор покрытия для заданного уровня доверия и известного закона распределения случайной угловой величины. Применение традиционного метода к сверхмалым выборкам, закон распределения которых априорно неизвестен, необоснованно и может привести к грубым погрешностям.

2. Анализ литературных данных и постановка проблемы

В работе [5] приведена и экспериментально апробирована методика оценивания основных выборочных круговых характеристик — среднего, моды, медианы, по выборкам объемом до 100 значений.

Цель работы — разработать и экспериментально исследовать методику оценивания точечных и интервальных характеристик случайных угловых величин по выборкам сверхмалого объема, принадлежащим к априорно неизвестному закону распределения.

3. Результаты исследований

Задачу оценивания результатов измерения случайных углов по выборкам сверхмалого объема предложено решать на основе бутстреп метода [4—6]. Бутстреп методы статистической обработки данных известны с 70-х годов XX века. Их преимущество состоит в увеличении статистики без увеличения входных данных.