

УДК (681.3.06 ÷ 681.518.54):622.276

М.П. АНДРИИШИН<sup>1</sup>, Е.А. ИГУМЕНЦЕВ<sup>2</sup>, Е.А. ПРОКОПЕНКО<sup>2</sup>

<sup>1</sup>УМГ «Киевтрансгаз», Киев, Украина

<sup>2</sup>Украинская инженерно-педагогическая академия, Харьков, Украина

## ЛИНЕЙНЫЕ ТРЕНДЫ В ДИАГНОСТИКЕ БАЛАНСА ГАЗА

Рассматривается статистический подход к определению наличия коррекции и погрешности измерения расхода и распределения газа в магистральном газопроводе. В его основе лежит алгоритм выявления признаков коррекции путем статистического анализа временных рядов. Разработана математическая модель баланса газа в магистральном газопроводе. Приведены результаты диагностики для двух автономных систем газопроводов: «Киевской системы» и «Экспортного газопровода» Управления магистральных газопроводов (УМГ) «Киевтрансгаз», полученные на основе разработанной модели.

**диагностика, измерение расхода газа, статистический анализ, изменение запаса газа, суточный расход газа, тренд, линейная регрессия**

### Общая постановка проблемы и ее связь с научно-практическими задачами

Газотранспортные предприятия допускают неточности из-за корректируемого измерения расхода газа, что не позволяет точно вычислить дисбаланс газа между приходом и распределением. Причиной этого являются погрешности измерений и утечки в отдельных нитках газопровода [1]. Выявление объема и мест коррекции измерений и утечек является нетривиальной задачей, для решения которой в настоящее время не разработано функциональных методов. Практика применения статистических методов в разных областях науки и народного хозяйства позволяет предположить, что их применение позволит решить вышеуказанные проблемы и в газовой промышленности.

**Обзор публикаций и анализ нерешенных проблем.** В настоящее время сбор и обработка статистической информации прихода и распределения газа по газотранспортной системе Украины ведется с помощью централизованной базы данных «АРМ диспетчера Трансгаза», основанной на измерениях современными расходомерами и вычислении физических характеристик газа [1]. Идея применения новых технологий (в частности статистических методов) для решения указанных проблем ранее вы-

сказывалась в научных публикациях [2, 3], но каких либо определенных результатов в данном направлении получено не было. Для выявления дисбалансов расхода газа в газопроводе в настоящее время применяются известные своей простотой методы [1], не позволяющие распознать коррекцию измерений.

**Цель исследований** – разработка универсального алгоритма выявления признаков коррекции измерения расхода газа, путем статистического анализа случайных временных рядов суточного измерения прихода и распределения.

### Результаты исследований

Представим разность суточного расхода  $D(t_i)$  между приходом  $X(t_i)$  и распределением  $Y(t_i)$  в газотранспортной системе за  $i$ -е сутки измерений в следующем виде:

$$D(t_i) = X(t_i) - Y(t_i); \quad (i = 1, 2, \dots, n), \quad (1)$$

где  $n$  – количество суток измерений временного ряда прихода и распределения;  $t_i$  – время.

Каждые сутки ( $t_i$ ) существует изменение запаса газа связанное с колебаниями давления и температуры

$$\alpha(t_i) = Z(t_i) - Z(t_{i-1}), \quad (2)$$

где  $Z(t_i)$ ,  $Z(t_{i-1})$  – запас газа в газопроводе в  $i$ -е и  $(i - 1)$  сутки измерений.

С учетом (2) дисбаланс  $\beta(t_i)$  между приходом и распределением определим из соотношения:  $\beta(t_i) = D(t_i) - \alpha(t_i)$ , ( $i = 1, 2, \dots, n$ ). Из (2), следует, что среднее изменение запаса газа  $\bar{\alpha}(t_i) = \frac{1}{n} \sum_{i=1}^n \alpha(t_i)$  равно:  $\bar{\alpha} = [Z(t_n) - Z(t_1)]/n \approx 0$ .

Среднее значение дисбаланса равно  $\bar{D} = \bar{X} - \bar{Y}$ , где  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X(t_i)$  и  $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y(t_i)$  — средние значения прихода и распределения.

При довольно длинных временных рядах  $n \rightarrow \infty$  достаточно определить среднюю разность

$$\bar{D} = \frac{1}{n} \sum_{i=1}^n D(t_i),$$

где  $\sigma_D^2 = \sigma_\beta^2 + \sigma_\alpha^2$ ,  $\sigma_\beta^2$  и  $\sigma_\alpha^2$  — дисперсии разности, дисбаланса и изменения запаса газа.

Дисперсия  $\sigma_D^2$  характеризует меру рассеивания

мгновенных значений  $D(t_i)$  от среднего значения  $\bar{D}$  и используется для сравнения показателей отчетности в различных временных периодах. Точное определение дисперсии зависит от определения статистической природы временных рядов  $X(t_i)$ ,  $Y(t_i)$ ,  $D(t_i)$ .

Для статистического анализа временных рядов использованы данные УМГ «Киевтрансгаз», где эксплуатируются две автономные системы транспорта газа: «Киевская система» (КС) и «Экспортный газопровод» (ЭГ). Замеры  $X(t_i)$  и  $Y(t_i)$ , ( $i = 1, 2, \dots, n$ ) проводились в течение пяти лет отдельно по ЭГ и КС ежедневно ( $t_i$ ). Относительный приход газа  $X_o(t_i) = X(t_i)/\bar{X}$  по КС (рис. 1) содержит характерные участки тренда, связанные с сезонной неравномерностью поставок газа. Аналогичный вид имеет временной ряд распределения, а также временные ряды прихода и распределения по ЭГ.

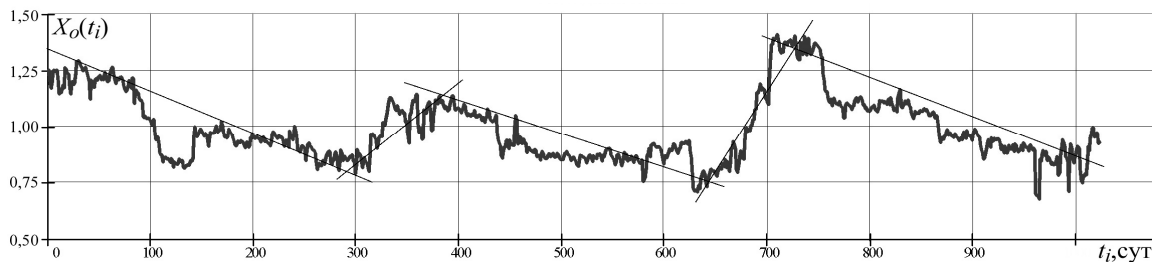


Рис. 1. Временной ряд относительного прихода газа по КС

Исследуем периодические и случайные составляющие временных рядов путем их спектрального анализа. В общем случае для этого используется преобразование Фурье [4]. Т.к. в расчетах используются данные, представляющие собой дискретную конечную последовательность ежедневных значений расхода газа, то применяется дискретная аппроксимация финитного быстрого преобразования Фурье [5]:

$$F[X(k/n)] = \sum_{i=1}^n X(t_i) \cdot \exp(-j2\pi ki/n) \quad (k = 1, 2, \dots, N), \quad (3)$$

где  $n$  — объем временной выборки;  $N$  — число гармоник спектра.

Спектральный состав временных рядов КС и ЭГ

выборки  $n = 1024$  одинаков и представляет собой сумму случайного и периодического сигналов с большим числом амплитуд отдельных  $k$ -х гармоник  $F = [X(k/n)]$ . На рис. 2 представлены относительные значения гармоник  $F[X(k/n)]/\sigma_X$ .

Максимальные значения амплитуд соответствуют периодам в один год ( $k = 3$ ), шесть месяцев ( $k = 6$ ), четыре месяца ( $k = 8$ ) и один месяц ( $k = 34$ ) (рис. 2).

Кроме того, в спектре присутствуют два пика на малых частотах  $k = 1$  и  $k = 2$  (первая и вторая гармоники), что обусловлено наличием линейного тренда отдельных участков временных рядов (рис. 1) [6]. Поэтому кроме компонент, описываемых периодическими функциями с сезонными ( $k = 3$ ;  $k = 6$ ) и технологическими периодами ( $k = 8$ ;  $k = 34$ ), а также

случайного процесса с широкополосными составляющими, в рядах наблюдается плавное долгосрочное движение, описываемое непрерывной функцией тренда, обусловленное требованиями поставки газа.

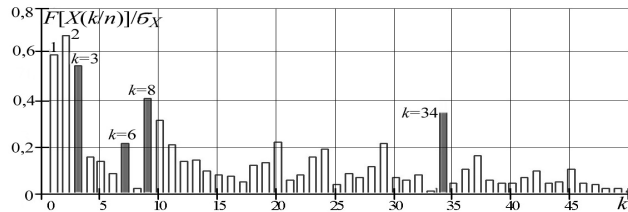


Рис. 2. Спектр временного ряда прихода газа  $X(t_i)$

Кроме того, в спектре присутствуют два пика на малых частотах  $k = 1$  и  $k = 2$  (первая и вторая гармоники), что обусловлено наличием линейного тренда отдельных участков временных рядов (рис. 1) [6]. Поэтому кроме компонент, описываемых периодическими функциями с сезонными ( $k = 3$ ;  $k = 6$ ) и технологическими периодами ( $k = 8$ ;  $k = 34$ ), а также случайного процесса с широкополосными составляющими, в рядах наблюдается плавное долгосроч-

ное движение, описываемое непрерывной функцией тренда, обусловленное требованиями поставки газа.

Оценку линейного тренда отдельных участков ряда проведем с помощью простой линейной регрессии [7]. Сложный спектральный состав временных рядов обладает своим специфическим законом плотности вероятности. Гистограммы  $X(t_i)$  и  $Y(t_i)$  КС показаны на рис. 3. Здесь  $N_{X_i}$ ,  $N_{Y_i}$  — число значений  $X(t_i)$  и  $Y(t_i)$ , попадающих в  $i$ -й фиксированный интервал наблюдений  $\Delta_{i+1} = (i + 1)\Delta - i\Delta$ ; ( $i = 1, 2, \dots, r$ ), где  $\Delta = (X_{max} - X_{min})/r$ ;  $r$  — число интервалов;  $i = 0$  соответствует  $X_{min}$ ;  $i = r$  соответствует  $X_{max}$ . Плотность вероятности рядов  $X(t_i)$ ,  $Y(t_i)$  (рис. 3) по виду напоминает закон Райса, [4]. Отличие от закона Райса состоит в том, что правый и левый максимумы плотности вероятности  $p[X(t_i)]$  и  $p[Y(t_i)]$  несимметричны относительно среднего (рис. 3). Таким образом, гистограммы подтверждают, что исследуемые ряды  $X(t_i)$  и  $Y(t_i)$  КС и ЭГ представляют сумму периодического и случайного рядов.

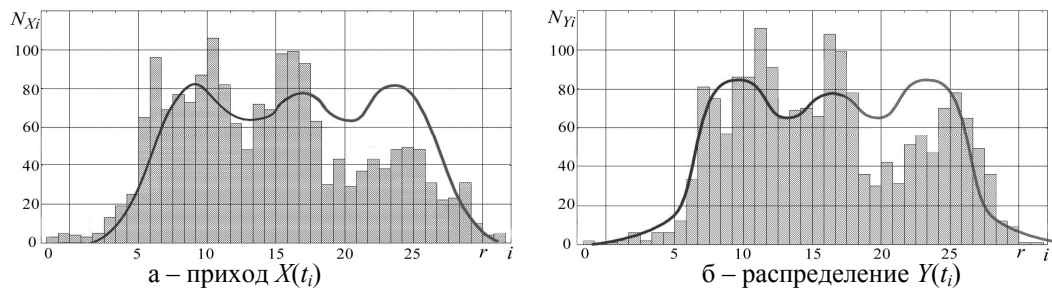


Рис. 3. Гистограммы рядов прихода и распределения газа КС

Данные статистической отчетности охватывают промежутки времени меньше или соизмеримые с обнаруженной периодичностью  $X(t_i)$ ,  $Y(t_i)$ ,  $D(t_i)$  в один год, поэтому необходимо зафиксировать постоянные периоды отчетности в весенне-летние и осенне-зимние периоды.

Применяемые при анализе спектральные соотношения, связанные с быстрым преобразованием Фурье, требуют выбирать длину выборки из соотношения  $n = 2^m$ . Здесь можно предложить в зимний и летний периоды выборку длиной  $n = 128$ ;  $m = 7$  (четыре месяца), а в весенний и осенний периоды  $n = 64$ ;  $m = 6$  (два месяца), что в сумме составляет период приблизительно один год:  $T = 2 \cdot 128 + 2 \cdot 64 = 384$ . Мини-

мально допустимая выборка анализа равна  $n = 50$  [6], а в спектре присутствует соответствующий технологический период  $n = 128$ ;  $k = 8$ , что позволяет использовать предложенные рекомендации.

Кроме того, установлено, что средние и дисперсии рядов с  $n = 128$  и  $n = 64$  являются также периодическими функциями с периодом в один год, возрастают в зимний и убывают в летний период. Таким образом, выборки длиной  $n = 128$  обладают зимой возрастающим, а летом — убывающим трендом, что обусловлено возрастанием потребления газа в зимний и убыванием в летний периоды. Спектральный анализ временных рядов выборки  $n = 128$  и  $n = 64$  показывает, что кроме дискретных периоди-

ческих составляющих и широкополосного шума спектр содержит два пика большой интенсивности на малых частотах, аналогично рис. 2, обусловленных линейным трендом.

Для выявления тренда проверялась случайность временных рядов на участках  $n = 1024, n = 128, n = 64$  с помощью метода поворотных точек и метода разностей [7]. Это позволило обнаружить наличие тренда в рядах  $X(t_i)$  и  $Y(t_i)$  КС и ЭГ при  $n = 128$  и  $n = 64$ . Применение метода Фостера–Стьюдента [7] позволило принять гипотезу о наличии тренда в дисперсии.

Оценка влияния трендов на  $X(t_i)$  и  $Y(t_i)$  проводилась с помощью простой линейной регрессии [7]:

$$\left. \begin{aligned} X(t_i) &= c_X + b_X t_i + \varepsilon_X(t_i); \\ Y(t_i) &= c_Y + b_Y t_i + \varepsilon_Y(t_i), \end{aligned} \right\} \quad (4)$$

где  $\varepsilon_X(t_i), \varepsilon_Y(t_i)$  – случайные составляющие рядов

(случайные помехи).

Параметры регрессии  $c_X, c_Y, b_X, b_Y$  получены [6]:

$$\left. \begin{aligned} b_X &= \frac{n \sum_{i=1}^n t_i X(t_i) - \sum_{i=1}^n t_i \sum_{i=1}^n X(t_i)}{n \sum_{i=1}^n t_i^2 - \sum_{i=1}^n t_i \sum_{i=1}^n t_i}; \\ c_X &= \frac{\sum_{i=1}^n X(t_i) \sum_{i=1}^n X(t_i)^2 - \sum_{i=1}^n t_i \sum_{i=1}^n t_i X(t_i)}{n \sum_{i=1}^n t_i^2 - \sum_{i=1}^n t_i \sum_{i=1}^n t_i}. \end{aligned} \right\} \quad (5)$$

Определив  $b_X, c_X$ , вычислим по выражению (4) значения регрессии для  $n = 128$  и проследим за изменением параметров относительной регрессии  $b^o_{Xi} = b_{Xi}/\bar{X}$  и  $b^o_{Yi} = b_{Yi}/\bar{Y}$  ( $i = 1, 2, \dots, 1024$ ), характеризующих линейный тренд за весь период измерений (рис. 4).

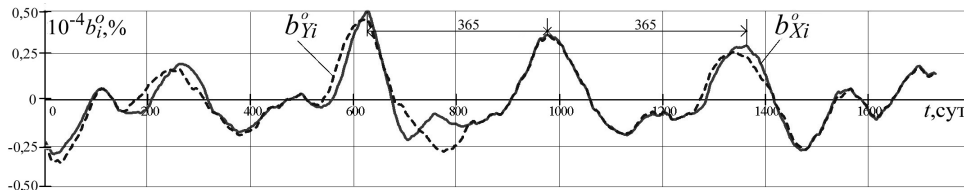


Рис. 4. Зависимость коэффициентов трендов  $b^o_{Xi}, b^o_{Yi}$  КС от времени

Из рис. 4 следует, что коэффициенты трендов (параметры регрессий)  $b^o_{Xi}$  и  $b^o_{Yi}$  являются периодическими функциями времени с периодом в один год, возрастают в зимний и убывают в летний периоды.

В качестве показателя интенсивности связи переменных  $X(t_i)$  и  $Y(t_i)$  с трендом используем коэффициент детерминации [7]:

$$B_X = \frac{\sum_{i=1}^n [(X(t_i) - \bar{X})(t_i - \bar{t})]^2}{\sum_{i=1}^n (X(t_i) - \bar{X})^2 \sum_{i=1}^n (t_i - \bar{t})^2}. \quad (6)$$

Коэффициент детерминации (6) показывает, какая часть полного рассеивания значений  $X(t_i)$  обусловлена трендом. Чем большую долю в общей дисперсии  $\sigma^2_X$  составляет дисперсия регрессии (тренда)  $\sigma^2_{Xp}$ , тем лучше выбранная функция регрессии соответствует эмпирическим данным. Остаточная дисперсия  $\sigma^2_{Xo}$  определяется  $\varepsilon_X(t_i)$  и характеризует неопределенность или неточность регрессии. Объяс-

няемая регрессией и остаточная дисперсии вычисляются с помощью коэффициента детерминации из следующего соотношения [7]:

$$\left. \begin{aligned} \sigma^2_{Xp} &= B_X \sigma^2_X; \\ \sigma^2_{Xo} &= (1 - B_X) \sigma^2_X. \end{aligned} \right\} \quad (7)$$

В табл. 1 представлены относительные значения коэффициентов регрессии  $b_o = b/\bar{X}$ ;  $c_o = c/\bar{X}$  и детерминации  $B$ , а также коэффициенты вариации общей  $q = \sigma/\bar{X}$ , остаточной  $q_o = \sigma_o/\bar{X}$  и объясняемой трендом  $q_p = \sigma_p/\bar{X}$  дисперсии КС и ЭГ в летний и зимний периоды. Временные ряды  $X(t_i)$  и  $Y(t_i)$  имеют близкие параметры трендов и дисперсий в летний и зимний периоды КС и ЭГ (табл. 1), возрастают зимой и убывают летом.

Для оценки разности  $D(t_i)$ , подставим (4) в (1):

$$D(t_i) = (b_X - b_Y)t_i + (c_X - c_Y) + \varepsilon_X(t_i) - \varepsilon_Y(t_i). \quad (8)$$

Коэффициенты трендов в зимний и летний периоды  $b_X \approx b_Y$ , что дает возможность представить

коэффициент тренда разности в виде:

$$b_D = b_X - b_Y \approx 0. \quad (9)$$

Таблица 1

Значения коэффициентов тренда

Газопровод	Период	Функция	Относительные коэффициенты					
			$c_0$	$10^6 b_0$	$10^2 B$	$10^2 q$	$10^2 q_p$	$10^2 q_0$
ЭГ	Зима	X	92	12,25	22,3	7,70	3,63	6,78
		Y	92	12,22	21,7	7,69	3,58	6,84
	Лето	X	101	-0,86	1,0	6,20	0,62	6,17
		Y	101	-0,86	1,0	6,24	0,62	6,21
КС	Зима	X	92	13,15	22,5	7,53	3,57	6,63
		Y	91	14,12	20,3	7,95	3,58	7,08
	Лето	X	110	-16,87	58,0	7,78	5,91	5,06
		Y	110	-16,72	55,0	8,04	5,95	5,39

Невыполнение условия (9) свидетельствует о корректируемых или ошибочных измерениях. С учетом (9) соотношение (8) представим в виде:

$$\left. \begin{aligned} D(t_i) &= c_D + \varepsilon_D(t_i); \\ c_D &= c_X - c_Y; \\ \varepsilon_D(t_i) &= \varepsilon_X(t_i) - \varepsilon_Y(t_i). \end{aligned} \right\} \quad (10)$$

Таким образом, временные ряды разности  $D(t_i)$  трендов не создают, что обусловлено равенством (9). Применение критериев наличия тренда (метода поворотных точек, метода разностей и метода Фостера–Стьюдента) свидетельствует об отсутствии трендов  $D(t_i)$ . Невыполнение равенства (9) КС в зимний период ( $b_D = -0,76 \cdot 10^{-6}$ ) (табл. 1) свидетельствует о корректировке измерений, однако тренды  $D(t_i)$  не возникают. Кроме того, завышенное значение общей дисперсии для этого случая также подтверждает корректировку измерений.

### Выводы

Показано, что временные ряды прихода и распределения состоят из случайных и периодических функций с периодом в 1,0; 0,5 и 0,33 года. Для данных отчетности необходимо использовать отдельные выборки в зимний и летний периоды равные 128 дням (0,33 года). Эти выборки обладают возрастающим трендом в зимний и убывающим в летний

периоды. Для классификации корректируемости измерений используются параметры линейной регрессии. При некорректируемых измерениях коэффициенты линейной регрессии прихода и распределения равны. Временные ряды разности  $D(t_i)$  трендов не создают.

**Перспективы дальнейших исследований.** Модель может быть усовершенствована при использовании нелинейной регрессии (сглаживания) прихода и распределения путем подбора оптимального коэффициента детерминации. Также представляет интерес анализ причин увеличения общей дисперсии разности при корректируемых измерениях.

### Литература

1. Вимірювання витрати та кількості газу: Довідник / М.П. Андрійшин і др.. – Івано-Франківськ: ПП. “Сімік”, 2004. – 160с.: іл.
2. Ігуменцев Є.О., Марчук Я.С., Андрійшин М.П. Виявлення перетоків газу в запірної арматурі КС за реєстрацією вібраційного сигналу // Нафтова і газова промисловість. – 2003. – № 2. – С. 13-18.
3. Корн Г., Корн Т. Справочник по математике. Для научных работников и инженеров. – М.: Наука, 1974. – 832с., ил.
4. Бендат Дж., Пирсон А. Применение корреляционного и спектрального анализа: Пер. с англ. – М.: Мир, 1983. — 312с.
5. Дьяконов В. Mathcad 8/2000: Специальный справочник. – С.-Пб.: Питер, 2000. – 592 с.
6. Иванов В.В. Анализ временных рядов и прогнозирование экономических показателей. – Х.: ХНУ, 1999. – 127 с.
7. Ферстер Э., Рёнц. Методы корреляционного и регрессионного анализа. – М.: Финансы и статистика, 1983. – 301с.

Поступила в редакцию 15.05.2008

**Рецензент:** д-р. техн. наук, проф. В.П. Себко, Национальный технический университет «ХПИ», Харьков.