*ISSN 1993-9981 Методи та прилади контролю якості, № 19, 2007*

УДК 681.121:389

# СТАТИСТИЧЕСКИЕ ХАРАКТЕРИСТИКИ ВРЕМЕННЫХ РЯДОВ ПРИХОДА И РАСПРЕДЕЛЕНИЯ ГАЗА МАГИСТРАЛЬНОГО ГАЗОПРОВОДА

***Я.С. Марчук [1)], М.П. Андриишин [1)], Е.А. Игуменцев [2)], В.Л. Добров [2)]***

*[1)] УМГ «Киевтрансгаз», пр. Комарова, 44, г. Киев, 01001, раб. тел. (8-044) 461-26-79*
*[2)] Украинская инженерно-педагогическая академия, ул. Демченко, 4, к. 4, г. Харьков, 61002, раб. тел. (8-0572) 706-12-45*

*Проанализированы причины дисбаланса между приходом и распределением газа двух автономных систем транспорта газа в Управлении магистральных газопроводов «Киевтрансгаз», получены автокорреляционные и взаимные корреляционные функции между приходом и распределением газа указанных систем. Показано, что указанные функции включают случайные и периодические составляющие, дисбаланс между приходом и распределением газа представляет собой стационарный случайный процесс и характеризуется остаточной дисперсией.*

*There are analyzed the reasons of unbalance between arrival and distributing of gas of two off-line systems of gas transport in Management of main gas pipelines of «Kievtransgaz», auto-correlationand mutual cross-correlation functions are got between arrival and distributing of gas of the specified systems. It is rotined that the indicated functions are included by casual and periodic constituents, an unbalance between arrival and distributing of gas is a stationary casual proces and described remaining dispersion.*

Газотранспортные предприятия допускают погрешности вследствие корректируемого измерения расхода газа, что не позволяет с достаточной точностью вычислить дисбаланс газа между приходом и его распределением. Причиной неточностей служат систематические погрешности измерений, а также наличие утечек в отдельных нитках газопровода [1]. Выявление объема и мест корректируемых результатов измерений и утечек в газопроводе является нетривиальной задачей, для которой в настоящее время не разработано функциональных методик. Практика применения статистических методов в разных областях науки и народного хозяйства позволяет предположить, что их применение позволит решить вышеуказанные проблемы и в газовой промышленности.

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 высказывалась в научной литературе [2, 3]. Однако каких-либо определенных результатов в данном направлении получено не было и для выявления дисбалансов газа в газопроводе применяются известные своей простотой методы [1], что не позволяет распознать корректируемые измерения.

Цель исследований данной работы заключается в том, чтобы разработать универсальный алгоритм обнаружения признаков корректируемого измерения расхода газа путем статистического анализа случайных временных рядов суточного измерения прихода и распределения газа.

Представим разность суточного расхода $D(t_i)$ между приходом $X(t_i)$ и распределением $Y(t_i)$ в $t_i$-ые сутки измерений в следующем виде:

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

где $n$ — число суток измерений временного ряда прихода и распределения.

В газотранспортной системе каждые сутки ($t_i$) существует изменение запаса газа

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

где $Z(t_i)$, $Z(t_{i-1})$ — запас газа в газопроводе в $i$-ые и $(i-1)$-ые сутки измерений, связанное с колебаниями давления и температуры. С учетом (2) дисбаланс $(t_i)$ между приходом и распределением определим из соотношения: $\beta(t_i) = D(t_i) - \alpha(t_i)$; $(i = 1,2,...,n)$. Несложно показать, используя (2), что среднее изменение запаса газа $\overline{\alpha}(t_i) = \frac{1}{n}\sum_{i=1}^{n}\alpha(t_i)$ равно: $\overline{\alpha} = [Z(t_n) - Z(t_1)]/n \approx 0$. Таким образом, для вычисления среднего значения дисбаланса $\overline{\beta}$ при достаточно длинных временных рядах $n \to \infty$ необходимо определить среднюю разность $\overline{D} = \frac{1}{n}\sum_{i=1}^{n} D(t_i)$. При этом дисперсия разности равна $\sigma_D^2 = \sigma_\beta^2 + \sigma_\alpha^2$, где $\sigma_\beta^2$ и $\sigma_\alpha^2$ — дисперсии дисбаланса и изменения запаса газа.

Поскольку дисперсия $\sigma_D^2$ характеризует меру рассеивания мгновенных значений $D(t_i)$ от среднего значения $\overline{D}$ и используется для сравнения показателей отчетности в различных временных периодах, необходима правильная методика вычисления дисперсии. Точное определение дисперсии зависит от установления статистической природы временных рядов $X(t_i)$, $Y(t_i)$, $D(t_i)$ путем применения современных статистических методов исследований. Здесь необходимы: спектральный анализ для установления периодичности; дисперсионный анализ для классификации нестабильности измерений; определение наличия и подбор трендов оценки сползания средних значений; установление с помощью корреляционных функций соотношения сигнал-шум (случайных и детерминированных составляющих); наилучшие кривые аппроксимации.

Для статистического анализа временных рядов используем данные УМГ «Киевтрансгаз», где эксплуатируются две автономные системы транспорта газа: «Киевская система (КС)» и «Экспортный газопровод (ЕГ)». Замеры $X(t_i)$ и $Y(t_i)$ проводились в течение пяти лет раздельно по ЕГ и КС ежесуточно $(t_i)$.

Определим взаимную корреляционную функцию $R_{XY}(\tau)$ между $X(t)$, $Y(t)$ для произвольного сдвига времени $\tau$ следующим образом [4]:

$$\left.\begin{aligned} R_{XY}(\tau) &= \lim_{T\to\infty}\frac{1}{T}\int_0^T [X(t)-\overline{X}]\,[Y(t+\tau)-\overline{Y}]dt; \\ \overline{X} &= \frac{1}{n}\sum_{i=1}^{n} X(t_i); \quad \overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y(t_i); \quad T = t_n. \end{aligned}\right\} \qquad (3)$$

В частном случае, когда $X(t) = Y(t)$, из (3) получим автокорреляционные функции $R_{XX}(\tau)$, $R_{YY}(\tau)$. Вычисление автокорфункций $R_{XX}(\tau)$ и $R_{YY}(\tau)$ для системы транспорта газа УМГ «Киевтрансгаз» выполнено с использованием функций пакета MathCad [5].

Полученные автокорфункции прихода и распределения КС по виду напоминают собой зависимости, образованные суммой случайного и периодического рядов [4] с периодом в один год (рис. 1). Аналогичную зависимость представляют автокорфункции ЕГ. Автокорфункции разности $R_{DD}$ ЕГ и КС (см. рис. 1) также образованы случайными и периодическими рядами несколько меньшего периода.

Взаимные автокорфункции $R_{XY}$ КС и ЕГ аналогичны $R_{XX}$ и $R_{YY}$, содержат случайные и периодические (период один год) составляющие и подтверждают выводы по автокорфункциям (рис. 2).

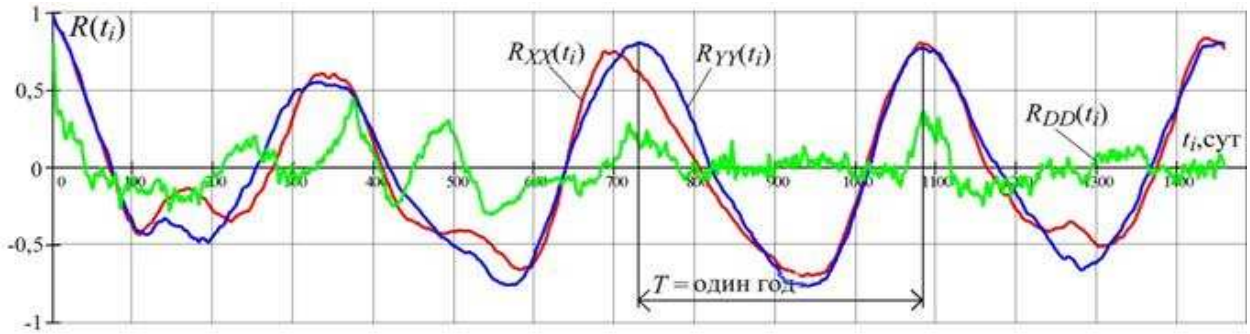



**Рисунок 1 – Автокорреляционные функции $R_{XX}$, $R_{YY}$, $R_{DD}$ по КС**

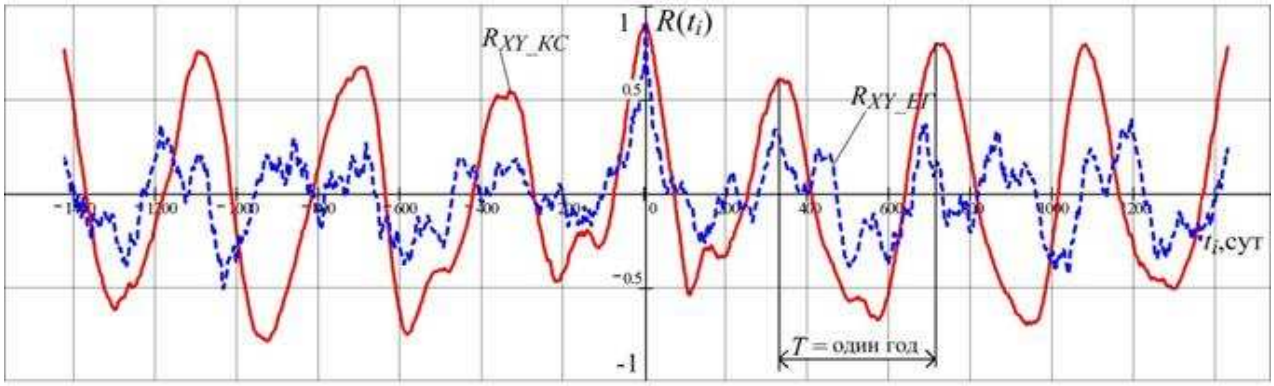



**Рисунок 2 – Взаимные корреляционные функции $R_{XY}$ КС и ЕГ**

Более наглядное представление о периодических и случайных составляющих временных рядов получим путем их спектрального анализа. В общем случае эта операция выполняется с помощью преобразования Фурье [4]. Так как в расчетах используются данные, представляющие собой дискретную конечную последовательность ежесуточных значений расхода газа, то применяется дискретная аппроксимация финитного быстрого преобразования Фурье [5]

$$F\left[X\left(k/n\right)\right]=\sum_{i=1}^{n} X\left(t_i\right)\cdot\exp\left(-\,j2\pi ki/n\right);$$
$$\left(k=1,2,\text{K},N\right), \qquad (4)$$

где $j^2=-1$, $n$ — объем временной выборки, $N$ — число гармоник спектра.

Данные статистической отчетности в УМГ охватывают промежутки времени меньшие или соизмеримые с обнаруженной периодичностью $X(t_i)$, $Y(t_i)$, $D(t_i)$ в один год (рис. 1, рис. 2), поэтому необходимо зафиксировать постоянные периоды отчетности в весенне-летние и осенне-зимние периоды.

Применяемые при анализе спектрально-корреляционные 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 в один год: $T=2\cdot128+2\cdot64=384$. Оценка минимально допустимой статистической выборки анализа равна $n=50$ [6] и позволяет использовать предложенные рекомендации. Спектральный анализ временных рядов выборки $n=128$ и $n=64$ (рис. 3) показывает, что кроме дискретных периодических составляющих и широкополосного шума, спектр содержит два пика на малых частотах (первая и вторая гармоники), что обусловлено наличием линейного тренда временных рядов $X(t_i)$ и $Y(t_i)$ [6]. Поэтому кроме компоненты, описываемой периодическими функциями с сезонными (один месяц — $k=4$) и технологическими периодами (восемь дней — $k=16$), а также случайного процесса, в рядах наблюдается плавное долгосрочное движение, описываемое некоторой непрерывной функцией, обусловленной поставками газа. На спектре представлены относительные значения $k$ - тых гармоник $F\left[X\left(k/n\right)\right]/\sigma_X$ .

С целью установления наличия тренда проведена проверка случайности временных рядов с помощью метода поворотных точек и метода разностей [7]. Применение данных методов позволило обнаружить наличие тренда в рядах $X(t_i)$ и $Y(t_i)$ КС и ЕГ. Дополнительное применение метода Фостера-Стьюдента [7] позволило принять гипотезу о наличии тренда в дисперсии.

Для выбора модели предположим, что временные ряды $X(t_i)$ и $Y(t_i)$ описываются аддитивной трендовой моделью вида [7]:

$$X\left(t_i\right)=f\left(t_i,\bar{a}\right)+\varepsilon_X\left(t_i\right), \qquad (5)$$

где $f\left(t_i,\bar{a}\right)$ — функция, описывающая основную тенденцию поставок газа в зимний или летний период времени (тренд); $\varepsilon_X\left(t_i\right)$— случайные независимые помехи.

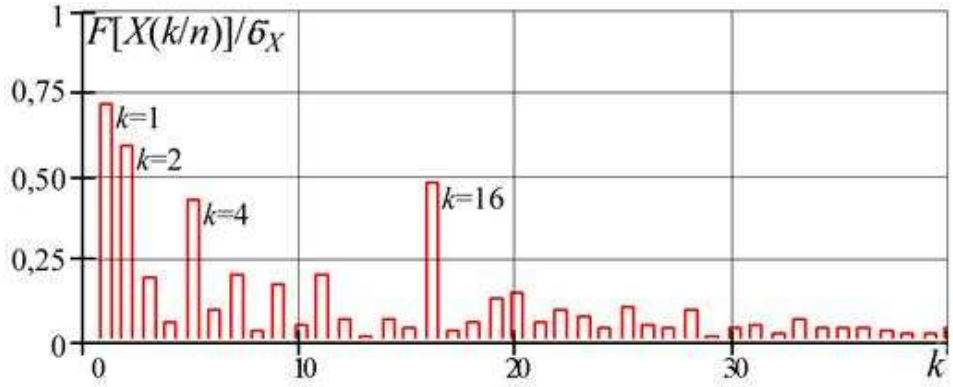



**Рисунок 3 – Спектр временного ряда прихода газа $X(t_i)$ (КС)**

Определение трендов с помощью простой линейной регрессии [7] не привело к положительным результатам из-за малого значения коэффициента детерминации. Поскольку любую непрерывную кривую, в силу теоремы Вейерштрасса [6,7], можно описать с полиномом $P_m(t_i)$ некоторой степени $m$, представим, что $f(t_i,\bar{a}) = P_m(t_i)$ и является таким полиномом. В соответствии с [7], характеристикой полинома $m$-ой степени есть $m$-ые разности, которые являются постоянными величинами. При этом $(m+1)$-ые разности и разности более высокого порядка равны нулю. В этом случае за оценку дисперсии ошибок $\sigma_e^2$ возьмем значение $V_m$:

$$V_m = \mathrm{var}\left(\Delta^m \varepsilon_X(t_i)\right) \Big/ C_{2m}^m = \sigma_e^2 = \mathrm{const}, \quad (6)$$

где $\mathrm{var}(!)$ — оператор дисперсии; $\Delta^m(!)$ — оператор взятия разности порядка $m$; $C_n^m$ — число сочетаний из $n$ по $m$.

Последовательно вычислим разности $\Delta X(t_i)$, их дисперсии $\mathrm{var}\left(\Delta X^{(k)}(t_i)\right)$ и показатели $V_k$, $k = 1,2,\ldots$, пока на некотором шаге $k = k_0$ эти показатели «стабилизируются», т.е. $V_{k_0} = V_i$, $(i > k_0)$. Тогда за степень полинома принимаем $m = k_0$ и тренд описываем моделью:

$$P_m(t_i) = a_0 + a_1 t + a_2 t^2 + \ldots + a_m t^m, \quad (7)$$

где $m = k_0$.

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

$$B_X = \sum_{i=1}^{n} \left[f(t_i,\bar{a}) - \bar{X}\right]^2 \Big/ \sum_{i=1}^{n} \left[X(t_i) - \bar{X}\right]^2. \quad (8)$$

Коэффициент детерминации $B_X$ (8) указывает, какая часть полного рассеивания значений $X(t_i)$ обусловлена трендом. Чем большую долю в общей дисперсии $\sigma_X^2$ составляет дисперсия регрессии (тренда) $\sigma_{Xp}^2$, тем лучше выбранная функция регрессии соответствует эмпирическим данным. Остаточная дисперсия $\sigma_{Xo}^2$ определяется $\varepsilon_X(t_i)$ и используется для характеристики неопределенности или неточности регрессии. Объясняемая регрессией и остаточная дисперсии вычисляются с помощью коэффициента детерминации из следующего соотношения [7]:

$$\begin{cases} \sigma_{Xp}^2 = B_X \sigma_X^2; \\ \sigma_{Xo}^2 = (1 - B_X)\sigma_X^2. \end{cases} \quad (9)$$

Аппроксимация функций $X(t_i)$ и $Y(t_i)$ полиномами (7) с применением критериев (6) и (8) для системы транспорта УМГ «Киевтрансгаз» выполнена с использованием функций пакета MathCad [5]. Наилучшей функцией является полином шестой степени $m = 6$, что отображает участок длиной $n = 128$ временных рядов ЕГ и КС в летний и зимний периоды (табл. 1). На рис. 4 представлены относительные значения $X_o(t_i) = X(t_i)/\bar{X}$ и $P_6^o(t_i) = P_6(t_i)/\bar{X}$ КС в зимний период с коэффициентом детерминации $B_X = 0{,}74$. Здесь большую долю составляет $\sigma_{XP} = 0{,}86\sigma_X$, что

**Таблица 1 - Относительные коэффициенты детерминации и вариации прихода и распределения газа по КС и ЕГ**

| Газо-провод | Период | Функция | Относительные коэффициенты | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | $10^{12}a^o$ | $a_6^o$ | $B$ | $q$ | $q_p$ | $q_o$ |
| Экспортный газопровод | Зима | $X$ | 1.25 | 1.12 | 0.603 | 0.080 | 0.060 | 0.018 |
| | | $Y$ | 1.25 | 1.12 | 0.601 | 0.080 | 0.060 | 0.018 |
| | Лето | $X$ | -2.17 | 1.10 | 0.330 | 0.060 | 0.034 | 0.026 |
| | | $Y$ | -2.17 | 0.97 | 0.690 | 0.060 | 0.050 | 0.010 |
| Киевская система | Зима | $X$ | 2.20 | 0.89 | 0.740 | 0.080 | 0.069 | 0.011 |
| | | $Y$ | 30.30 | 0.89 | 0.710 | 0.080 | 0.067 | 0.013 |
| | Лето | $X$ | -0.67 | 1.19 | 0.690 | 0.080 | 0.064 | 0.014 |
| | | $Y$ | -0.94 | 1.19 | 0.650 | 0.080 | 0.064 | 0.014 |

обусловлено плановыми поставками газа с переменными трендами. Однако, для временного ряда разности $D(t_i)$ и дисбаланса $\beta(t_i)$ подобрать полином или линейную регрессию не удалось, т.к. $B \approx 0$.

В табл. 1 представлены относительные значения коэффициентов полинома $a^o = a_0 / \overline{X}$; $a_6^o = a_6 / \overline{X}$ и детерминации $B$, а также коэффициенты вариации общей $q = \sigma / \overline{X}$, остаточной $q_p = \sigma_p / \overline{X}$ и объясняемая трендом $q_o = \sigma_o / \overline{X}$ дисперсии КС и ЕГ 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 ЕГ (см. табл.1), возрастают зимой и убывают летом. Временные ряды разности $D(t_i)$ трендов не создают, представляют собой стационарный случайный процесс и характеризуются остаточной дисперсией $\sigma_D^2 = \sigma_{Do}^2$.

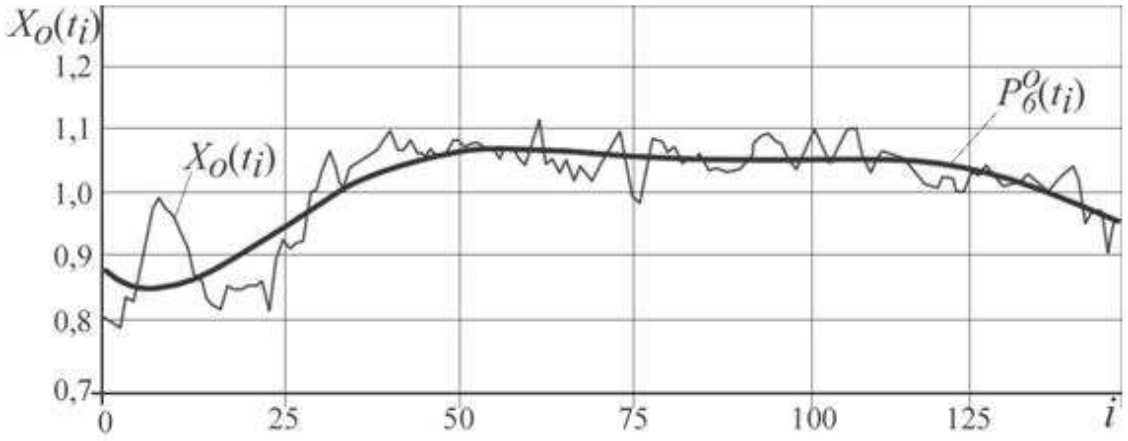



**Рисунок 4 – Аппроксимация временного ряда $X(t_i)$ КС в зимний период (полином шестой степени)**

Таким образом приведенные выше спектрально-корреляционные соотношения показали, что временные ряды прихода и распределения газа состоят из случайных и периодических функций с периодом в один год, один месяц и одну неделю. Для данных отчетности необходимо использовать отдельные выборки в зимний и летний периоды равные 128 дням. Указанные выборки $X(t_i)$ и $Y(t_i)$ обладают возрастающим трендом в зимний и убывающим трендом в летний периоды и аппроксимированы полиномами шестой степени, характеризующими поставки газа. Коэффициенты детерминации в этом случае близки к единице. Дисбаланс между приходом и распределением представляет собой стационарный случайный процесс и характеризуется остаточной дисперсией.

Представленную модель необходимо улучшить путем использования факторного анализа. Кроме того, провести детальный анализ причин увеличения общей дисперсии разности при корректируемых измерениях, над чем в настоящий момент работают авторы.

***Література***

*1. Андріїшин М.П. і ін. Вимірювання витрати та кількості газу: Довідник. – Івано-Франківськ: ПП "Сімик", 2004. – 160с.*

*2. Ігуменцев Є.О., Марчук Я.С., Андріїшин 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.*

*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. — СПб.: Изд- во «Питер», 2000. — 592с.*

*6. Иванов В.В. Анализ временных рядов и прогнозирование экономических показателей. — Харьков: изд-во Харьковского национального университета, 1999. — 127с.*

*7. Ферстер Э., Рёнц. Методы корреляционного и регрессионного анализа. — М.: Финансы и статистика, 1983. — 301с.*