УДК 528.11

*Г. А. ХАРКЕВИЧ*

# ЭМПИРИЧЕСКИЕ МОДЕЛИ ОСАДОК ИНЖЕНЕРНЫХ СООРУЖЕНИЙ НА ОСНОВЕ АВТОРЕГРЕССИИ С МИНИМАЛЬНОЙ ДИСПЕРСИЕЙ

Представим ход осадки инженерного сооружения совокупностью последовательностей конечной длительности

$$\{S_j(t_i)\}, \quad j \in [1, N], \quad i \in [1, K], \tag{1}$$

где $N$, $K$ — число осадочных марок и циклов наблюдений по каждой из них соответственно.

Последовательности (1) образуют процесс, в общем случае нестационарный по $i$ и неэргодический по $j$. Его вероятностные характеристики зависят от номера цикла и номера осадочной марки. Рассмотрим $S_j(t)$ раздельно по номеру $j$ и совокупность $S_j(t_i)$ — при фиксированном $i$, т. е. изучим ход каждой осадочной марки в отдельности, а затем профиль $S_j(t_i)$ — при фиксированном номере цикла.

В теории случайных процессов требование стационарности и эргодичности при определении вероятностных характеристик является основным. Поэтому на первом этапе анализа последовательностей (1) преобразуем их к стационарным. Составим для этого разность

$$X_j(t_i) = S_j(t_i) - y_j(t_i), \tag{2}$$

где $y_i(t_i)$ тренды последовательностей $S_j(t_i)$; $i=1, 2, \ldots, K$ при фиксированном $j$. Если $y_j(t_i)$ истинные, то $X_j(t_i)$ можно рассматривать как чисто случайные величины. Поскольку не имеется метода, приводящего к однозначному определению истинных $y_j(t_i)$, принимаем за $y_j(t_i)$ результаты нелинейного сглаживания $S_j(t_i)$ по семи ординатам с помощью полиномов третьей степени [4]. Получившийся в итоге процесс $X_j(t_i)$ назовем квазистационарным. Если в ансамбле $S(t)$ значение осадки для какой-либо марки в отдельном цикле оказалось пропущенным, то соответствующий уровень $S_j(t_i)$ следует «восстановить» с помощью интерполяционного полинома Лагранжа, например, по пяти произвольно расположенным либо равноотстоящим узлам [3] в зависимости от календарного графика наблюдений.

Таким образом, получим совокупность последовательностей вида

$$S_j(t_i) = y_j(t_i) + X_j(t_i), \tag{3}$$

где $y(t)$, $X(t)$ — неслучайная и квазислучайная компоненты соответственно.

Представим квазислучайную компоненту $X(t)$ смешанной моделью авторегрессии — скользящего среднего [1]

$$(X_i - \bar{X}) = \alpha_1 (X_{i-1} - \bar{X}) + \alpha_2 (X_{i-2} - \bar{X}) + \cdots$$
$$+ \alpha_m (X_{i-m} - \bar{X}) + z_i + \beta_1 z_{i-1} + \beta_2 z_{i-2} + \cdots + g_i z_{i-i}, \tag{4}$$

где

$$\bar{X} = \frac{1}{K} \sum_{i=1}^{K} X_i; \tag{5}$$

$m$ — порядок уравнения авторегрессии; $z_i$ — случайные числа, являющиеся следствием погрешностей наблюдений и влияния других факторов.

В силу многозначности оценивания коэффициентов $\beta_i$ принимаем $\beta_i = 1/2^i$, которые обеспечивают устойчивость процедуры реализации $z_i$ по $X_i$, так как все $\beta_i$ лежат внутри единичного круга.

Ставим задачу выделить с помощью (4) из процесса $X_j(t_i)$ неслучайную компоненту и присоединить ее к $y_i(t_i)$. Получающиеся в результате случайные числа $z_i$ используем для оценки остаточной дисперсии $\sigma_z^2$ в процессе $S_j(t_i)$.

Коэффициенты $\alpha_1$, $\alpha_2$ уравнения авторегрессии в (4) имеют вид [1]:

$$\alpha_1 = \frac{r_1(1-r_2)}{1-r_1^2}, \quad \alpha_2 = \frac{r_2 - r_1^2}{1-r_1^2}, \tag{6}$$

где $r_1$, $r_2$ — коэффициенты корреляции. Так как $r_i = C_i/C_0$, имеем

$$\alpha_1 = \frac{C_1(C_0 - C_2)}{C_0^2 - C_1^2}, \quad \alpha_2 = \frac{C_0 C_2 - C_1^2}{C_0^2 - C_1^2}. \tag{7}$$

Здесь $C_i$ — коэффициенты автоковариации с лагом $i$:

$$C_i = \frac{1}{K-i} \sum_{j=1}^{K-i} (X_j - \bar{X})(X_{j+i} - \bar{X}). \tag{8}$$

Значения $\alpha_i$ при $i > 2$ находим по рекуррентной формуле

$$C_i = \alpha_1 C_{i-1} + \alpha_2 C_{i-2} + \cdots + \alpha_m C_{i-m}. \tag{9}$$

Выборочная оценка остаточной дисперсии $\sigma_m^2$ имеет вид

$$\sigma_m^2 = \frac{1}{K-2m-1} \sigma(\bar{X}, \alpha_1, \alpha_2, \cdots, \alpha_m), \tag{10}$$

где

$$\sigma(\bar{X}, \alpha_1, \alpha_2 \cdots, \alpha_m) = (K-m)(C_0 - \alpha_1 C_1 - \alpha_2 C_2 - \cdots \alpha_m C_m). \tag{11}$$

Так как $\sigma_m^2$ зависит от индекса $m$, то выбором

$$\sigma_{m^*}^2 = \min \sigma_m^2$$

определяем порядок $m^*$ уравнения авторегрессии в (4), обеспечивающий минимум остаточной дисперсии $\sigma_{m^*}^2$. Удержав в (4) слагаемые при $l \leqslant 4$, получим формулу для оценки случайных величин $z_i$ на основе результатов измерений:

$$z_i = (X_i - \bar{X}) - \alpha_1(X_{i-1} - \bar{X}) - \alpha_2(X_{i-2} - \bar{X}) - \cdots$$

$$- \alpha_{m^*}(X_{i-m^*} - \bar{X}) - \frac{1}{2} z_{i-1} - \frac{1}{4} z_{i-2} - \frac{1}{8} z_{i-3} - \frac{1}{16} z_{i-4}, \tag{12}$$

$$i = 1, 2, \ldots, K$$

и оценку дисперсии

$$\sigma_z^2 = \frac{1}{K-4} \sum_{i=1}^{K} z_i^2 \tag{13}$$

по каждой марке с номером $j = 1, 2, \ldots, N$. Здесь число степеней свободы принято равным $K-4$, так как сглаживание $S(t_i)$ выполнялось с помощью параболы третьей степени. Дисперсия $\sigma_z^2$ характеризует рассеивание наблюдений $S_j(t_i)$ относительно кривой $y_j(t_i) + \bar{X}_j$ по $j$-й осадочной марке, $i = 1, 2, \ldots, K$. При сохранении методики наблюдений в $(K+1)$ цикле можно принять

$$\sigma_z^2 = \frac{1}{K-4}\sum_{i=1}^{K} z_i^2 = \frac{1}{K-3}\sum_{i=1}^{K+1} z_i^2$$

и использовать $\sigma_z^2$ для оценки прогноза $Y_{k+1,j}$ осадки $S_j(t_{K+1})$ по марке $j$

$$S_{K+1,j} = Y_{K+1,j} \pm V_{K+1,j},$$

где

$$Y_{K+1} = y_{K+1} + \overline{X} + \alpha_1 (S_K - y_K - \overline{X}) + \alpha_2 (S_{K-1} - y_{K-1} - \overline{X}) + \cdots$$

$$+ \alpha_{m^*} (S_{K-m^*} - y_{K-m^*} - \overline{X}) + \frac{1}{2} z_K + \frac{1}{4} z_{K-1} + \frac{1}{8} z_{K-2} + \frac{1}{16} z_{K-3};$$

$$V_{K+1,j} = \sqrt{\sigma_{zj}^2 - \sigma_{yj}^2};$$

$\sigma_{yj}$ — дисперсия сглаженной кривой; $K$ — последний цикл наблюдений.

Величина $V_{K+1,j}$ определяет доверительный интервал для осадки $S_{K+1,j}$ относительно кривой $y_{K+1,j}+X_j$ в точке $K+1$.

На этом первый этап обработки результатов наблюдений по каждой марке в $K$ циклах заканчивается. Величины $Y_{K+1,j}$, $V_{K+1,j}$, $j=1, 2, \ldots, N$ являются исходными для анализа неоднородностей осадок в $K+1$ цикле наблюдений.

Для любого другого номера $i \leqslant K$ следует брать $Y_{i,j}$, $V_{i,j}$.

На втором этапе рассматриваем $Y_{K+1,j}$, $V_{K+1,j}$ как сигнал и помеху на входе фильтра, а $u_{K+1,j}$, $\mu_{K+1,j}$ их профильтрованные значения соответственно.

Таким образом, задача состоит в выявлении статистически достоверных величин $u_{K+1,j}$ в $Y_{K+1,j}$ на фоне помех $V_{K+1,j}$.

Известно [2], что наиболее полной характеристикой для отношения сигнал/помеха является энергетическое отношение $\rho$, максимизируемое по всей длине фильтра

$$\rho = u^T R_V^{-1} u, \tag{14}$$

где $u$ — вектор-столбец ординат $Y_{K+1}$; $R_V^{-1}$ — обратная корреляционная матрица помех.

В этом случае

$$u_{K+1,j} = \sum_{i=0}^{N-1} h_i Y_{K+1,j-i}; \tag{15}$$

$$\mu_{K+1} = \sum_{i=0}^{N-1} h_i V_{K+1,j-i}, \tag{16}$$

где $h_i$ — весовые коэффициенты, определяемые из решения уравнения

$$(R_Y - \lambda R_V) h = 0 \tag{17}$$

под условием $\sum_i h_i^2 = 1$. (18)

В этом уравнении $\lambda$ — собственные значения матрицы ($R_Y$—$R_V$); $h$ — собственный вектор весовых коэффициентов; $R_Y$, $R_V$ — корреляционные матрицы сигнала и помехи соответственно.

$$R_Y = \begin{vmatrix} R_{Y0} & R_{Y1} & R_{Y2} & \cdots & R_{Ym} \\ R_{Y1} & R_{Y0} & R_{Y1} & \cdots & R_{Vm-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ R_{Ym} & R_{Ym-1} & \cdots & \cdots & R_{Y0} \end{vmatrix},$$

$$R_V = \begin{vmatrix} R_{V0} & R_{V1} & R_{V2} & \cdots & R_{Vm} \\ R_{V1} & R_{V0} & R_{V1} & \cdots & R_{Vm-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ R_{Vm} & R_{Vm-1} & \cdots & \cdots & R_{V0} \end{vmatrix};$$

$m=N/2$ при $N$-четном и $m=(N+1)/2$ при $N$-нечетном; автоковариации вычисляются по известным формулам с лагом $j$:

$$R_{Yj} = \frac{1}{N-j} \sum_{i=1}^{N-j} (Y_i - \overline{Y})(Y_{i+j} - \overline{Y}); \tag{19}$$

$$R_{Vj} = \frac{1}{N-j} \sum_{i=1}^{N-j} (V_i - \overline{V})(V_{i+j} - \overline{V}), \tag{20}$$

где

$$\overline{Y} = \frac{1}{N} \sum_{j=1}^{N} Y_j, \quad \overline{V} = \frac{1}{N} \sum_{j=1}^{N} V_j. \tag{21}$$

Система уравнений (17) имеет ненулевое решение, если определитель матрицы ($R_Y$—$R_V$) равен нулю. В этом случае существует $m+1$ собственных значений $\lambda_i$ и соответствующие им $m+1$ собственных векторов $h$, выбирая $\lambda_{\max}$ и соответствующий ему собственный вектор $h_{\max}$, получаем наибольшее отношение $\rho_{\max}$ сигнал/помеха на выходе фильтра.

Профильтрованный по формуле (15) сигнал $u_j$ с весовой функцией $h_{\max}$ — главная компонента в $Y_j$. Вторая и последующая компоненты соответствуют меньшим вкладам в $Y_j$.

Точность вычисления $u_j$ определяется величиной $\mu_{K+1,j}$ на выходе фильтра (16). Поэтому вероятнейшая кривая $u_j(t_{K+1})$ будет находиться внутри доверительной полосы шириной $2\mu_{K+1,j}$.

1. *Дженкинс Г., Ваттс Д.* Спектральный анализ и его приложения. М., 1972. 2. *Никитин А. А.* Теоретические основы обработки геофизической информации. М., 1986. 3. *Справочник* по специальным функциям с формулами, графиками и математическими таблицами / Под. ред. М. Абрамовица и Н. Стигана. М., 1979. 4. *Четыркин Е. М., Калихман Н. Л.* Вероятность и статистика. М., 1982.

Статья поступила в редколлегию 26. 12. 88