Обобщенный алгоритм прогнозирования технического состояния объектов наземных комплексов с использованием комбинированных моделей авторегрессии – скользящего среднего

Generalized algorithm of forecasting technical state of ground complex elements with application of combined autoregression models – moving average

Ключевые слова: техническое состояние – technical state; стохастическая модель – stochastic model; модель авторегрессии – autoregression model.

Рассмотрен обобщенный алгоритм построения модели временных рядов, описывающих изменение параметров технического состояния объектов наземных комплексов во времени, с использованием комбинированных моделей авторегрессии – скользящего среднего. Обобщенный алгоритм включает в себя частные алгоритмы: ввода исходной информации и преобразования временного ряда, идентификации стохастической модели, предварительного оценивания параметров стохастической модели, выбора оптимальной стохастической модели, оценивания параметров стохастической модели и вывода результатов вычислений. Применение алгоритма наиболее целесообразно при организации эксплуатации техники, обслуживание которой производится по фактическому техническому состоянию с контролем параметров в системе мониторинга.

Described is the generalized algorithm of making the time-series model describing the parameters of ground complex elements technical state change in time with application of the combined autoregression models – moving average. The generalized algorithm includes the particular algorithms: input of initial information and transformation of time series; identification of stochastic model; preliminary evaluation of stochastic model parameters; selection of optimum stochastic model; evaluation of stochastic model parameters and output of calculation results. Application of the algorithm is most reasonable at organization of the equipment operation, maintenance of which is carried out by actual technical state with checking parameters in the monitoring system.

БЕССОНОВ / BESSONOV P.
Павел Евгеньевич
(bessonov.pe@szte.ru)
старший научный сотрудник,
филиал ФГУП «ЦЭНКИ» – «ЦЭНКИ – Северо-Запад»,
Санкт-Петербург

ПИВОВАРОВ / PIVOVAROV O.
Олег Григорьевич
(command@usst3.spb.ru)
начальник ФГУП «ГУССТ № 3 при Спецстрое России»,
Санкт-Петербург

Математические модели, которые могут быть использованы для описания процесса изменения параметров объектов контроля в системе мониторинга объектов наземных комплексов (ОНК), содержат ряд величин, определяемых по результатам наблюдений. Задача состоит в построении формализованных процедур, применяемых к статистическим данным для выявления моделей, адекватно описывающих процесс, и оценивания их неизвестных параметров. Для построения моделей число наблюдений N выбирается из условия $N \geq 20–25$.

Основой построения модели временного ряда является представление исходного процесса x_t в виде линейного преобразования последовательности независимых импульсов a_t с нормальным распределением при нулевом среднем и дисперсией σ_a^2 (белого шума) $x_t = a_t+\psi_1 a_{t-1}+... = \psi(B)a_t$, где B – оператор сдвига назад, определяемый как $B^m x_t = x_{t-m}$ $(m = 0,1,2,...)$, $\psi(B) = 1+\psi_1 B+\psi_2 B^2+...$ – передаточная функция фильтра, преобразующего a_t в x_t.

Наиболее часто используются модели [1]:

1. Модель авторегрессии (АР) порядка p: $x_t = \Phi_1 x_{t-1}+\Phi_2 x_{t-2}+...+\Phi_p x_{t-p}+a_t$, или $\Phi(B)x_t = a_t$, где $\Phi(B) = 1-\Phi_1 B-\Phi_2 B^2-...-\Phi_p Bp$ – оператор авторегрессии.

2. Модель скользящего среднего (СС) порядка q: $x_t = a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}$, или $x_t = \theta(B)a_t$, где $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$ – оператор СС порядка q.

3. Комбинированная модель авторегрессии – скользящего среднего АРСС порядка p, q: $x_t = \Phi_1 x_{t-1} + \Phi_2 x_{t-2} + \ldots + \Phi_p x_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}$, или $(B)x_t = \theta(B)a_t$. В данной модели имеется $p+q+2$ неизвестных параметра, которые оцениваются по наблюдениям.

4. Модель авторегрессии – проинтегрированного скользящего среднего (АРПСС). Данная модель применима для однородного нестационарного процесса (процесса, который можно трансформировать в стационарный при помощи взятия разностей): $\varphi(B) = \Phi(B)(1-B)^d$, где $\Phi(B)$ – стационарный оператор, $(1-B)^d$ – оператор взятия разности порядка d. Тогда процесс АРПСС порядка p, d, q представляется в виде $\Phi(B)(1-B)^d x_t = \theta(B)a_t$.

На практике реальный процесс изменения диагностических параметров ОНК может быть представлен одной из рассмотренных моделей АР(р), СС(q), АРСС(p,q) и АРПСС(p,d,q), причем величины p, d, q обычно не превышают 1–2. Порядок идентификации следующий. Вначале выбирается порядок разности d столько раз, сколько необходимо для сведения нестационарного процесса к стационарному. Затем с помощью автокорреляционной и частной автокорреляционной функций определяются вид модели и приближенные оценки ее параметров. Структурная схема обобщенного алгоритма выбора модели включает в себя следующие частные алгоритмы: алгоритм 1 – ввод исходной информации и преобразование временного ряда; алгоритм 2 – идентификация стохастической модели; алгоритм 3 – предварительное оценивание параметров стохастической модели; алгоритм 4 – выбор оптимальной стохастической модели; алгоритм 5 – оценивание параметров стохастической модели, вывод результатов вычислений.

Алгоритм 1

Входная информация: значения временного ряда x_t, $t = \overline{1, N}$; число наблюдений N; параметр преобразования λ.

После ввода исходных данных ряд преобразуется следующим образом:

$$x'_t = \begin{cases} x_t, & \lambda = 1; \\ (x_t + m)^{\lambda}, & \lambda \neq 0, \lambda \neq 1; \\ \ln(x_t + m), & \lambda = 0, \end{cases}$$

где $m = 1{,}05\sup\{|x_t|\}$.

Выходная информация включает все входные данные, а также параметры преобразования λ, m и значения преобразованного ряда x'_t.

Алгоритм 2

Входная информация: выходная информация алгоритма 1; порядок операции взятия разности d; максимальная задержка автокорреляционной функции r_K(АКФ) и частной автокорреляционной функции $\tilde{\Phi}_{KK}$ (ЧАКФ) ($K =]T/4[$). По алгоритму 2 вычисляется разностный ряд, вычисляются величины: среднее значение ряда

$$\overline{\omega}_t = \frac{1}{N}\sum_{t=1}^{N}\omega_t$$

и дисперсия $\sigma_\omega^2 = c_0 = \frac{1}{N-1}\sum_{t=1}^{N}(\omega_t - \overline{\omega})^2$;

АКВФ по формуле

$$c_k = \frac{1}{N}\sum_{t=1}^{N-K}(\omega_t - \overline{\omega})(\omega_{t+k} - \overline{\omega}), \quad k = \overline{0, K};$$

АКФ $r_k = c_k / c_0$; ЧАКФ

$$\tilde{\Phi}_{kk} = \begin{cases} r_1, & \text{при } k = 1; \\ \dfrac{r_k - \sum_{j=1}^{K-1}\tilde{\Phi}_{k-1,j}r_{k-j}}{1 - \sum_{j=1}^{K-1}\tilde{\Phi}_{k-1,j}r_{k-j}}, & \text{при } k = 2,3,\ldots,K, \end{cases}$$

где $\tilde{\Phi}_{kj} = \tilde{\Phi}_{k-1,j} - \tilde{\Phi}_{k-1,k-j}$; $j = 1,2,\ldots,K-1$.

Выходная информация включает все входные данные, а также значения разностного преобразованного ряда ω_t, среднее значение ряда ω, дисперсию ряда c_0, АКВФ ряда c_k, $k = 0,1,\ldots,K$.

Алгоритм 3

Входная информация: число параметров авторегрессии p; число параметров скользящего среднего q; АКВФ ряда c_k.

После проверки условия $p > 0$ вычисляются оценки параметров авторегрессии Φ путем решения системы p уравнений $B \cdot \Phi = K$, где $B_{ij} = c_{|q+i-j|}$, $æ_i = c_{q+i}$; $i, j = 1,2,\ldots,p$. Затем вычисляются параметры скользящего среднего.

По известным АКВФ c_k ряда ω_t вычисляется модифицированная последовательность ковариаций c'_j

$$c'_j = \begin{cases} \sum_{i=0}^{p}\sum_{k=0}^{p}\tilde{\Phi}_i\tilde{\Phi}_k c_{|j+i-k|}, & p > 0\ (\tilde{\Phi}_0 = -1), \\ c_j, & p = 0,\ j = 0,1,\ldots,q. \end{cases}$$

Затем с помощью алгоритма Ньютона – Рафсона

[2]: $\tau^{i+1}=\tau^{i}-h^{i}$, где $T^i h^i = f^i$, вычисляется вектор τ^{i+1} на (i+1)-й итерации по его значению на i-й итерации, где:

$\tau^i=\left(\tau_0^i,\tau_1^i,...,\tau_q^i\right)$; $f_j'=\sum_{k=0}^{q-j}\tau_k^i\tau_{k+j}^i-c_j'$; $f^i=\left(f_0^i,f_1^i,...,f_q^i\right)$;

$$T^i=\begin{pmatrix}\tau_0^i\ \tau_1^i\ ...\ \tau_q^i\\ \tau_1^i\ \tau_2^i\ ...\ 0\\ \Lambda\ \Lambda\ \Lambda\ \Lambda\\ \tau_q^i\ \ 0...\ 0\end{pmatrix}+\begin{pmatrix}\tau_0^i\ \tau_1^i\ ...\ \tau_q^i\\ 0\ \ \tau_0^i\ ...\ \tau_{q-1}^i\\ \Lambda\ \Lambda\ \Lambda\ \Lambda\\ 0\ \ 0...\ \tau_0^i\end{pmatrix}$$

с начальными значениями

$$\tau_0^i=\sqrt{c_0'}\,,\tau_1^i=\tau_2^i=...=\tau_q^i=0\,.$$

Когда $\left|f_j^i\right|<\varepsilon,\ j=\overline{0,q}$ для некоторого заданного значения ε, итеративная процедура считается завершенной и по последнему значению $\tau=(\tau_0,\tau_1,...,\tau_q)$ находятся оценки параметров СС по формуле:

$$\widetilde{\theta}_{j0}=-\tau_j/\tau_0\,,\ \ j=\overline{1,q}\,.$$

Оценка дисперсии белого шума $\widetilde{\sigma}_a^2$ вычисляется по следующей формуле:

$$\widetilde{\sigma}_{a0}^2=\begin{cases}\tau_0^2, & \text{при } q>0,\\ c_0-\sum_{i=1}^{p}\widetilde{\Phi}_{i0}c_i, & \text{при } q=0.\end{cases}$$

Оценка общей константы $\widetilde{\mu}_0$ вычисляется по формуле:

$$\widetilde{\mu}_0=\begin{cases}\overline{\omega}_t\left(1-\sum_{i=1}^{p}\widetilde{\Phi}_{i0}\right), & \text{при } p>0,\\ \overline{\omega}_t, & \text{при } p=0.\end{cases}$$

Выходная информация включает все входные данные, а также начальные оценки параметров АР $\widetilde{\Phi}_0=\left(\widetilde{\Phi}_{10},\widetilde{\Phi}_{20},...,\widetilde{\Phi}_{p0}\right)$, начальные оценки параметров СС $\widetilde{\theta}_0=\left(\widetilde{\theta}_{10},\widetilde{\theta}_{20},...,\widetilde{\theta}_{q0}\right)$, начальную оценку дисперсии $\widetilde{\sigma}_{a0}^2$, начальную оценку общей константы $\widetilde{\mu}_0$.

Алгоритм 4

Входная информация: выходная информация программы 3; управляющий параметр M для общей константы μ.

Используя возвратное представление модели, вычисляют остаточные ошибки a_t для заданного набора параметров в два этапа:

1) $$e_t=(\omega_t-\mu)-\sum_{i=1}^{p}\Phi_i(\omega_{t-i}-\mu)+\sum_{j=1}^{q}\theta_j e_{t+j}\,,$$

где e_t – ошибки прогноза назад; $e_{t+j}=0$ при $t+j>0$;

2) $$a_t=e_t-\sum_{i=1}^{p}\Phi_i e_{t-i}+\sum_{j=1}^{q}\theta_j a_{t-j}\,,$$

где $a_{t-j}=0$ при $t-j\geq Q$; $t=Q, Q+1,..., N-d$ и Q – отрицательное значение времени t, при котором значения прогнозируемых величин неизмеримо малы.

Управляющий параметр

$$M=\begin{cases}1, & \text{если } \mu \text{ включено в модель;}\\ 0, & \text{если } \mu \text{ не включено в модель.}\end{cases}$$

Затем для каждого набора параметров (μ,Φ,θ) вычисляется сумма квадратов остаточных ошибок $S(\mu,\Phi,\theta)$ по формуле

$$S(\mu,\Phi,\theta)=\sum_{t=Q}^{N-d}a_t^2\,.$$

По минимальной величине $S(.)$ определяется оптимальная стохастическая модель. Выходная информация включает все входные данные, а также оценки параметров $\widetilde{\Phi}$ и $\widetilde{\theta}$ оптимальной модели, остаточные ошибки a_t, порядок операции взятия разности d, число параметров модели p и q, сумму квадратов остаточных ошибок S.

Алгоритм 5

Входная информация: выходные данные программы 4; параметры γ, ξ, ε, ограничивающие поиск минимума $S(\mu,\Phi,\theta)$.

Определение параметров μ,Φ,θ, минимизирующих $S(\mu,\Phi,\theta)$, проводится в четыре этапа:

1. Все параметры модели обозначаются через $\beta=(\beta_1,\beta_2,...,\beta_k)$, где $k=p+q+1$, т.е. $\beta=f(\mu,\Phi,\theta)$.

Затем делается поочередное возмущение каждого из параметров и для данной модели вычисляются рекуррентным способом значения остаточных ошибок: $a_t(\omega_t; \beta_{1,0}; ...; \beta_{i,0}+\delta_i; ...; \beta_{k,0})$, где $\beta_{i,0}$ – начальное значение i-го параметра модели, полученное в алгоритме 3; δ_i – приращение i-го параметра модели, $i = \overline{1,k}$.

По значению остаточных ошибок вычисляются производные

$$x_{i,t} = [a_t(\omega_t; \beta_{1,0}; ...; \beta_{i,0}; ...; \beta_{k,0}) - -a_t(\omega_t; \beta_{1,0}; ...; \beta_{i,0}+\delta_i; ...; \beta_{k,0})]\frac{1}{\delta_i}.$$

2. По известным $a_t, x_{i,t}$, найденным для текущих значений, находятся:

матрица размером $K\times K$: $\boldsymbol{A} = \{A_{ij}\}$,

где $A_{ij} = \sum_{t=0}^{N-d} x_{i,t} x_{j,t}$;

вектор $g = (g_1, g_2,...,g_k)$, где $g_i = \sum_{t=0}^{N-d} x_{i,t} a_t$;

нормирующие величины $D_i = \sqrt{A_{ii}}$.

3. Относительно вектора h^* решается система линейных уравнений: $A^*h^* = g^*$, где $A^*_{i,j} = A_{ij}/D_i D_j$, $i \neq j$; $A^*_{i,j} = 1+\gamma$, $g^*_i = g_i/D_i$, вычисляются новые значения параметров $\beta = \beta_0 + h$, $h_j = h^*/D_j$, оценивается сумма квадратов остаточных ошибок $S(\beta)$.

4. Если $S(\beta) \le S(\beta_0)$, то исследуются поправки параметров h. Если $h_i < \varepsilon$, сходимость достигнута и вычисляются стандартные ошибки $S_i = \left(A_{ii}^{-1} S(\beta)/(N-d-p-q)\right)^{1/2}$, корреляционная матрица $R = \{R_{ij}\}$,

где $R_{ij} = A_{ij}^{-1} \Big/ \left(A_{ij}^{-1} A_{jj}^{-1}\right)^{1/2}$,

χ^2 – статистика $\chi^2 = n\sum_{k=1}^{K} r_{aa}^2(k)$,

$k = 0,1,2,...,K$, где $r_{aa}(k) = c_{aa}(k)/c_{aa}(0)$,

$$c_{aa}(k) = \frac{1}{N-d}\sum_{t=1}^{N-d-k}(a_t - \overline{a})(a_{t+k} - \overline{a});$$

$$\overline{a}\,\frac{1}{N-d}\sum_{t=1}^{N-d} a_t .$$

В противном случае, $\beta_0 := \beta$; $\gamma := \gamma\xi$ и вычисления продолжаются с третьего этапа.

Выходная информация включает все входные данные, а также значения параметров модели (μ, Φ, θ), остаточные ошибки a_t, стандартные ошибки S_i, корреляционную матрицу оценок R, χ^2 – статистику.

Таким образом, в настоящей статье был рассмотрен алгоритм построения математической модели поведения параметров ОНК во времени по результатам их периодического контроля в процессе нормальной эксплуатации. Применение алгоритма наиболее целесообразно при организации эксплуатации техники, обслуживание которой производится по фактическому техническому состоянию с контролем параметров в системе мониторинга.

Литература

1. *Бокс Д., Дженкинс Г.* Анализ временных рядов. Прогноз и управление. – Вып. 1. – М.: Мир, 1974.

2. Статистические методы обработки результатов наблюдений / Под ред. Р.М. Юсупова. – МО СССР, 1984.