СТАТИСТИЧЕСКИЙ МЕТОД ОЦЕНКИ БЕЗВОЗВРАТНЫХ ПОТЕРЬ ЯДЕРНЫХ МАТЕРИАЛОВ

А. В. Порошин

Федеральное государственное унитарное предприятие
«Производственное объединение "МАЯК"»

Изложен подход к оценке безвозвратных потерь ядерных материалов на основании набора статистических данных. Предложен алгоритм вывода уравнения нормы безвозвратных потерь, а также метод отбраковки «грубых» измерений при наборе статистики.

Введение

Одной из основных проблем учета и контроля ядерных материалов (ЯМ) является задача определения количества ЯМ, ушедших в безвозвратные потери. При этом, если потери за счет радиоактивного распада еще можно оценить по теоретическим формулам, то оценка безвозвратных технологических потерь вызывает определенные сложности.

Технологические потери определяются особенностями технологического процесса, устройством оборудования, свойствами самого материала и рядом других факторов.

Один из подходов оценки технологических потерь состоит в оценке этих самых факторов и влиянии их на количество потерь. Данный подход требует фундаментальных исследований и скрупулезных расчетов. При этом полученный результат не всегда бывает достаточно точным.

В большинстве случаев за безвозвратные потери принимается разница между теоретическим (документальным) количеством материала на конец межбалансового периода и количеством материала, определенным в результате физической инвентаризации, а норма безвозвратных потерь определяется, как максимально возможная разница между теоретическим и фактическим количеством материала. В этом случае инвентаризационная разница (ИР) всегда будет равна нулю. Однако, в случае возникновения систематической ошибки при измерениях в межбалансовый период, может возникнуть ситуация, когда теоретическое количество материала окажется меньше фактического, то есть потерь не будет, а будет необъяснимая положительная ИР. Чтобы избежать подобной ситуации, все-таки безвозвратные потери должны определяться по отдельной методике.

Один из возможных подходов к оценке безвозвратных потерь и норм безвозвратных потерь будет приведен ниже. Метод не требует фундаментальных исследований свойств материала и рассмотрения конструктивных особенностей установок. Метод основывается на статистической оценке потерь материала за некоторое количество периодов.

Метод применим к производствам с устоявшимся технологическим процессом.

Описание метода (набор статистики)

Во многих технологических процессах количество безвозвратных потерь, а также номы безвозвратных потерь, определяется, как функция от количества выпускаемой готовой продукции на данном технологическом переделе. Поэтому необходимо создать набор статистических данных по потерям в привязке к выпускаемой продукции. Предлагается определить следующие величины для каждого периода:

НК – количество материала на начало периода по данным физической инвентаризации;

УВ – поступления материала на технологический участок (в зону учета ЗУ) за отчетный период;

УМ – уменьшение материала в ЗУ за отчетный период, при этом:

УМ = Выд + Вып, где Выд – выдача материала из ЗУ в другие ЗУ; Вып – выдача материала из ЗУ в виде готовой продукции;

Пест – естественный распад материала (для материалов, у которых эта величина существенна);

Тогда конечное теоретическое количество материала может быть определено как КК = НК + УВ – УМ – Пест.

Безвозвратные потери можно будет определить следующим образом:

Пот = КК – ФК, где ФК – фактически измеренное количество материала на конец отчетного периода (НК следующего периода).

В итоге получаем набор Пот_i и соответствующих Вып_i.

Данную зависимость можно рассматривать как зависимость сигнала x_i (Вып_i) и соответствующего отклика y_i(Пот_i). Полученная парная зависимость сигнала и отклика может быть записана в табличном виде и обработана с помощью методов регрессионного анализа.

Регрессионный анализ статистических данных

Полученные парные зависимости могут быть изображены на графике:

На полученном изображении явно видно, что может быть проведена некая средняя линия таким образом, что все точки будут лежать от нее на некотором оптимальном расстоянии. Такая линия будет носить название линии регрессии. Построение линий регрессии может быть осуществлено методом наименьших квадратов, либо методом ортогональных полиномов Чебышева.

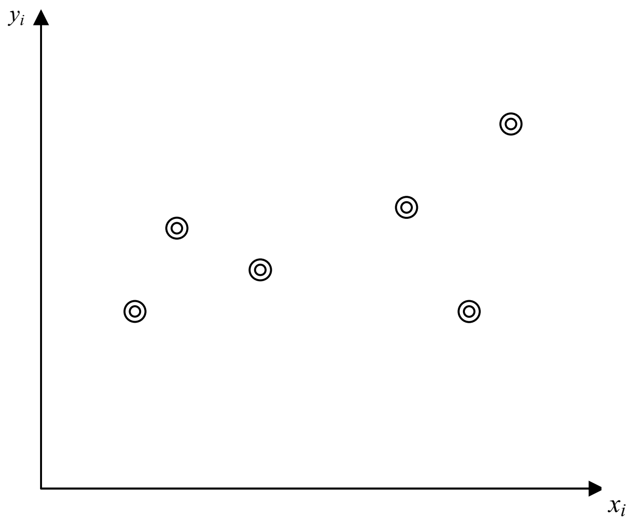

Рис. 1.

Метод наименьших квадратов

Линию регрессии будем определять, как полином заданной степени:

$$y = a + bx + cx^2 + dx^3 + \ldots + ex^m.$$

При этом необходимо получить линию, от которой точки будут отстоять на минимальное расстояние, то есть должно выполняться следующее условие

$$U = \sum_{i=1}^{n} \left[y_i - y(x_i) \right]^2 \to \min.$$

Таким образом, для решения задачи необходимо вычислить коэффициенты $a, b, c, d, \ldots, e$, минимизирующие сумму квадратов отклонений U. Для этого необходимо вычислить частные производные функции U по каждому из коэффициентов $a, b, c, d, \ldots, e$ и приравнять их к нулю:

$$\begin{cases} \dfrac{\partial U}{\partial a} = 0; \\ \dfrac{\partial U}{\partial b} = 0; \\ \dfrac{\partial U}{\partial c} = 0; \\ \vdots \\ \dfrac{\partial U}{\partial e} = 0. \end{cases}$$

В результате получаем систему линейных нормальных уравнений, которою можно решить известными методами, например методом Гаусса.

Для случая полинома третьей степени система будет выглядеть следующим образом:

$$\begin{cases} an + b\sum x + c\sum x^2 + d\sum x^3 = \sum y; \\ a\sum x + b\sum x^2 + c\sum x^3 + d\sum x^4 = \sum xy; \\ a\sum x^2 + b\sum x^3 + c\sum x^4 + d\sum x^5 = \sum x^2 y; \\ a\sum x^3 + b\sum x^4 + c\sum x^5 + d\sum x^6 = \sum x^3 y. \end{cases}$$

Для оценки значимости полученного уравнения регрессии можно использовать F-критерий Фишера. Для этого вычисляют общую дисперсию точек y_i

$$\overline{S_y^2} = \frac{\sum\left(y_i - \overline{y}\right)^2}{n-1}, \text{ где } \overline{y} = \frac{\sum y_i}{n}.$$

Также вычисляется остаточная дисперсия $S_{\text{ост}}^2 = \dfrac{\sum\left(y_i - y(x_i)\right)^2}{n-2}$.

Остаточная дисперсия представляет собой показатель ошибки предсказания уравнением регрессии значений отклика y_i. Качество предсказания определяют, сравнивая общую и остаточную дисперсии, то есть определяют, насколько уравнение регрессии точнее определяет зависимость сигнала и отклика, чем просто среднее $\overline{y}$. Отношение общей и остаточной дисперсий сравнивают со значением коэффициента Фишера при соответствующем уровне значимости (5 или 1 %). Если отношение дисперсий больше значения коэффициента Фишера

$$\hat{F} = \overline{S_y^2} \Big/ S_{\text{ост}}^2 > F^T_{(n-1,\ n-2,5\,\%)},$$

то говорят, что уравнение регрессии статистически значимо описывает парную зависимость сигнала и отклика при соответствующем уровне значимости.

Выбор наилучшего приближения может быть осуществлен постепенным увеличением степени полинома регрессии до тех пор, пока остаточная дисперсия не перестанет уменьшаться.

Метод ортогональных полиномов Чебышева

Недостатком метода наименьших квадратов является необходимость пересчета всего полинома при выборе наилучшего приближения. Метод ортогональных полиномов Чебышева позволяет существенно упростить этот процесс. Аппроксимирующий полином строится в виде суммы слагаемых с повышающимися степенями, при этом вычисление очередного коэффициента не влечет за собой пересчет ранее определенных коэффициентов.

Искомый полином определятся в следующем виде:

$$y = a_0\phi_0(x) + a_1\phi_1(x) + \ldots + a_m\phi_m(x),$$

где $\phi_0(x) = 1, \quad \phi_1(x) = x + \alpha_1, \ldots, \phi_l(x) = x^l + \alpha_l^{(1)}x^{l-1} + \ldots.$

Если многочлены $\phi_l(x)$ уже построены, то соответствующие коэффициенты a_i могут быть найдены по тому же принципу, что и в методе наименьших квадратов (принцип Лежандра)

$$U = \sum_{i=1}^{n}\left\{y_i - \left[a_0\phi_0(x) + a_1\phi_1(x) + \ldots + a_m\phi_m(x)\right]\right\}^2 \to \min.$$

В результате вычисления частных производных U получаем систему уравнений

$$\begin{cases} a_0\sum\left[\phi_0(x_i)\right]^2 + a_1\sum\phi_0(x_i)\phi_1(x_i) + a_2\sum\phi_0(x_i)\phi_2(x_i) + \ldots + a_m\sum\phi_0(x_i)\phi_m(x_i) = \sum y_i\phi_0(x_i); \\ a_0\sum\phi_0(x_i)\phi_1(x_i) + a_1\sum\left[\phi_1(x_i)\right]^2 + a_2\sum\phi_1(x_i)\phi_2(x_i) + \ldots + a_m\sum\phi_1(x_i)\phi_m(x_i) = \sum y_i\phi_1(x_i); \\ a_0\sum\phi_0(x_i)\phi_m(x_i) + a_1\sum\phi_1(x_i)\phi_m(x_i) + \ldots + a_m\sum\left[\phi_m(x_i)\right]^2 = \sum y_i\phi_m(x_i). \end{cases}$$

Для упрощения системы многочлены $\phi_l(x)$ следует подбирать, исходя из следующих условий:

$$\sum\phi_l(x_i)\phi_k(x_i) = 0, \quad l \neq k;$$

$$\sum\left[\phi_l(x_i)\right]^2 \neq 0, \quad l = \overline{0, m}.$$

Такие многочлены называют ортогональными многочленами Чебышева.

Тогда коэффициенты уравнения регрессии могут найдены по следующей формуле

$$a_l = \frac{\sum y_i\phi_l(x_i)}{\sum\left[\phi_l(x_i)\right]^2}, \text{ где } l = \overline{0, m}.$$

Многочлены Чебышева могут быть найдены по следующим рекуррентным формулам:

$$\phi_{r+1}(x) = (x + \beta_{r+1})\phi_r(x) + \gamma_{r+1}\phi_{r-1}(x);$$

$$\beta_{r+1} = -\frac{\sum x_i\left[\phi_r(x_i)\right]^2}{\sum\left[\phi_r(x_i)\right]^2}, \quad \gamma_{r+1} = -\frac{\sum x_i\phi_{r-1}(x_i)\phi_r(x_i)}{\sum\left[\phi_{r-1}(x_i)\right]^2}.$$

Постепенно увеличивая степень регрессионного полинома до тех пор, пока остаточная дисперсия не перестанет существенно уменьшаться, получаем оптимальное приближение статистических данных линией регрессии.

Определение норм безвозвратных потерь

Найденное уравнение регрессии описывает зависимость безвозвратных потерь от выпуска готовой продукции. Но каждое из значений отклика (потери) является случайной величиной, полученной из ряда измерений с определенными погрешностями. Следовательно, действительные значения безвозвратных потерь будут находиться не на линии регрессии, а в некоторой области вокруг найденной линии. Определив верхнюю границу этой области, можно будет принять ее за норму безвозвратных потерь, так как превышение этой границы будет говорить о том, что имеет место пропажа (недостача) материала. Определение верхней границы может быть осуществлено следующим образом.

Величина безвозвратных потерь при наборе статистики определялась по формуле: Пот = КК – ФК = НК + УВ – УМ – Пест – ФК.

Каждое из слагаемых является результатом измерений, произведенных с определенной погрешностью σ. Для простоты рассуждений будем считать, что каждое из слагаемых определено с одинаковой погрешностью (по одной методике). Тогда можно определить погрешность для каждого значения отклика y_i : σ_i.

Общую дисперсию для уравнения регрессии можно определить по формуле:

$$S^2_{\text{рег}} = S^2_{\text{ост}} + \frac{D\left(\sum\limits_{i=1}^{n} \text{Пот}_i\right)}{n-1}.$$

Очевидно, что набор статистических данных по безвозвратным потерям не является в чистом виде набором независимых случайных величин. Каждая пара последовательных значений Пот связаны величиной НК для вторых Пот из пары:

$$\text{Пот}_i = \text{КК}_i - \text{НК}_{i+1} = \text{НК}_i + \text{УВ}_i - \text{УМ}_i - \text{НК}_{i+1}.$$

Таким образом, учитывая, что все слагаемые НК, УВ, УМ являются результатами независимых измерений, а следовательно независимыми случайными величинами, можно определить ковариацию каждой пары последовательных значений Пот:

$$\begin{gathered}
\text{cov}(\text{Пот}_i, \text{Пот}_{i+1}) = \text{cov}(\text{НК}_i + \text{УВ}_i - \text{УМ}_i - \text{НК}_{i+1}, \text{НК}_{i+1} + \text{УВ}_{i+1} - \text{УМ}_{i+1} - \\
- \text{НК}_{i+2}) = \text{cov}(\text{НК}_i - \text{НК}_{i+1}, \text{НК}_{i+1} - \text{НК}_{i+2}) + \text{cov}(\text{НК}_i - \text{НК}_{i+1}, \text{УВ}_{i+1} - \\
- \text{УМ}_{i+1}) + \text{cov}(\text{УВ}_i - \text{УМ}_i, \text{НК}_{i+1} - \text{НК}_{i+2}) + \text{cov}(\text{УВ}_i - \text{УМ}_i, \text{УВ}_{i+1} - \text{УМ}_{i+1}) = \\
= \text{cov}(\text{НК}_i - \text{НК}_{i+1}, \text{НК}_{i+1} - \text{НК}_{i+2}) = \text{cov}(\text{НК}_i, \text{НК}_{i+1}) + \text{cov}(\text{НК}_i, -\text{НК}_{i+2}) + \\
+ \text{cov}(-\text{НК}_{i+1}, \text{НК}_{i+1}) + \text{cov}(-\text{НК}_{i+1}, -\text{НК}_{i+2}) = \\
= -\text{cov}(\text{НК}_{i+1}, \text{НК}_{i+1}) = -\text{D}(\text{НК}_{i+1}).
\end{gathered}$$

Ковариационная связь непоследовательных величин Пот будет более слабой, поэтому ей можно пренебречь.

Тогда:

$$D\left(\sum_{i=1}^{n}\text{Пот}_i\right)=D\left(\text{Пот}_1\right)+D\left(\sum_{i=2}^{n}\text{Пот}_i\right)+2\operatorname{cov}\left(\text{Пот}_1,\sum_{i=2}^{n}\text{Пот}_i\right)=D\left(\text{Пот}_1\right)+$$

$$+D\left(\sum_{i=2}^{n}\text{Пот}_i\right)+2\operatorname{cov}\left(\text{Пот}_1,\text{Пот}_2\right)=D\left(\text{Пот}_1\right)+D\left(\sum_{i=2}^{n}\text{Пот}_i\right)-2D\left(HK_2\right)=$$

$$=\ldots=\sum_{i=1}^{n}D\left(\text{Пот}_i\right)-2\sum_{i=2}^{n-1}D\left(HK_i\right).$$

Уравнение норм безвозвратных потерь может быть определено из уравнения регрессии следующим образом

$$y_{\text{норм}}\left(x\right)=y\left(x\right)+\gamma S_{\text{рег}},$$

где γ – квантиль, равный либо 1,96 для доверительной вероятности 0,95, либо 2,58 для доверительной вероятности 0,99.

Для простоты формулы норм безвозвратных потерь сдвиг на величину $\gamma S_{\text{рег}}$ может быть применен к каждому отклику из набора статистики, и по вновь полученным точка может быть построена регрессионная прямая, уравнение которой и будет являться формулой для норм безвозвратных потерь.

Отбраковка грубых ошибок

При наборе статистики ряд значений отклика может получиться в результате влияния посторонних факторов, таких как человеческая ошибка, грубая ошибка прибора, погрешность пробоотбора и т. д. Такие значения заставляют линию регрессии отклоняться от настоящего его положения, значит, подобные грубые ошибки необходимо находить и исключать из рассмотрения.

Для проведения анализа на грубые ошибки найдем отклонения значений отклика из набора статистики от линии регрессии

$$e_i=y_i-y\left(x\right).$$

Для найденных отклонений определим среднее $\bar{e}=\dfrac{\sum e_i}{n}$ и дисперсию $S_e=\sqrt{\dfrac{\sum\left[e_i-\bar{e}\right]^2}{n-1}}$.

Для крайнего e_i (наибольшего или наименьшего) вычисляют

$$\tau=\left|e_i-\bar{e}\right|/S_e,$$

которое затем сравнивают с граничным значением τ_q при соответствующей доверительной вероятности q (0,95 или 0,99).

Граничное значение τ_q может быть выражено через критическое значение распределения Стьюдента $t_{1-q,\,n-2}$:

$$\tau_q = \frac{t_{(1-q,\ n-2)}\sqrt{n-1}}{\sqrt{n-2+\left[t_{(1-q,\ n-2)}\right]^2}}.$$

Если условие $\tau \le \tau_й$ выполняется, но наблюдение не отбрасывается, если нет, то наблюдение исключают, после чего заново находят уравнение регрессии.

Выводы

Предлагаемый метод оценки безвозвратных потерь может быть применен для производств с устоявшимся технологическим процессом. Метод не требует фундаментальных исследований свойств учитываемого материала и оценки возможных потерь за счет особенностей технологии и конструкций установок.

Список литературы

1. Львовский Е. Н. Статистические методы построения эмпирических формул. М.: Высшая школа, 1988.
2. Гмурман В.Е. Введение в теорию вероятностей и математическую статистику. М.: Высшая школа, 1963.
3. Кассандрова О. Н., Лебедев В. В. Обработка результатов наблюдений. М.: Наука, 1970.