

Методология построения зависимостей нейтронно-физических сечений для малогрупповых реакторных расчетов на основе регрессионного анализа

В.Г. Зимин и А.А. Семенов

*Лаборатория тренажерных систем,
Московский Инженерно-Физический Институт,
Каширское шоссе 31, Москва, 115409
E-mail: slava@ets.mephi.ru, and@ets.mephi.ru*

Аннотация

В докладе предложена методика представления зависимости малогрупповых нейтронно-физических сечений от выгорания и теплофизических параметров ячейки. Нейтронно-физические сечения представляются в виде суммы двух членов: базовых сечений, которые зависят только от выгорания и рассчитываются при номинальных значениях теплофизических параметров и отклонений, которые зависят как от выгорания, так и от теплофизических параметров ячейки.

Зависимость базовых сечений от выгорания интерполируется кубическим сплайном. Зависимость отклонений аппроксимируется многомерным полиномом. Особенностью данной методики является построение оптимальной полиномиальной зависимости отклонений для заданной точности аппроксимации. Оптимальность понимается в смысле минимизации числа членов полинома, описывающего данную зависимость. Решение этой задачи выполнено в рамках регрессионного анализа, члены многомерного полинома формируются программой шаговой регрессии.

В качестве примера использования предложенной методики строятся зависимости двухгрупповых сечений ячейки реактора ВВЭР от выгорания, плотности и температуры теплоносителя, температуры топлива и концентрации бора. Заданную среднеквадратичную ошибку в 0.05% обеспечивает пятимерный полином, который содержит не более 20 членов для сечений быстрой группы и около 50 членов для сечений тепловой группы. Построенная зависимость тестировалась как на исходных данных, так и на проверочной последовательности точек выбранных случайным образом. Результаты показывают, что исходные точки, использованные для построения модели, можно также применять для контроля точности модели.

1. Введение

Для нейтронно-физического расчета ядерного реактора требуются малогрупповые нейтронные сечения, которые зависят как от материального состава ТВС, так и от теплофизических параметров. Эти характеристики ТВС будем называть переменными состояниями. Например, переменные состояния ТВС реактора ВВЭР включают выгорание топлива, температуру топлива, температуру и плотность теплоносителя, концентрацию бора. Для того чтобы построить зависимости сечений ТВС от переменных состояний, область изменения переменных состояний покрывается расчетной сеткой и сечения рассчитываются в узлах этой сетки. Расчет обычно организуется следующим образом. Для номинальных теплогидравлических параметров ячейки проводится расчет выгорания, называемый базовым расчетом. На каждом шаге по выгоранию выполняются дополнительные расчеты для получения зависимости сечений от остальных параметров

состояния ТВС. Различные программы, например TBC-M, UNK, WIMS могут быть использованы для данных расчетов. Для приемлемой точности описания зависимости сечений от параметров состояния ТВС обычно выполняется расчет около тысячи вариантов. Затем полученный набор сечений интерполируется или аппроксимируется на всю область определения параметров состояния ячейки. В нашей работе рассматриваются две проблемы возникающие в процессе подготовки нейтронных сечений:

- выбор расчетной сетки, покрывающей область определения переменных состояния ТВС;
- процедура интерполяции или аппроксимации полученного набора сечений.

Рассмотрим традиционные подходы к решению этих вопросов. В качестве расчетной сетки для сечений обычно используется регулярная прямоугольная сетка. Применение регулярной сетки для 5-ти мерного пространства переменных состояния ТВС реактора ВВЭР требует более тысячи точек для приемлемой точности, что приводит к значительным затратам вычислительных ресурсов даже для современных компьютеров.

Многомерная табличная интерполяция или *аппроксимация функцией многих переменных* используются на практике для описания зависимостей полученного набора сечений. Например, в работе (Семенов и др., 1998) применяется многомерная линейная интерполяция для расчета сечений ТВС реактора ВВЭР, тогда как в работе (Turski et al., 1997) макросечения каналов РБМК аппроксимируются многомерным полиномом. Оба этих подхода имеют свои достоинства и недостатки.

В случае табличной интерполяции, значение функции получают линейной или кубической интерполяцией значений функции в узлах сетки. Для заданного порядка интерполяции, точность можно увеличить путем увеличения числа точек сетки. К достоинствам данного подхода можно отнести универсальность алгоритма, который практически не зависит от исходных данных. Недостатком этого подхода является тот факт, что используемые алгоритмы требуют применения регулярной сетки при расчете сечений, что приводит к значительному числу точек сетки. Наиболее часто используемая линейная интерполяция обладает низкой точностью. Это особенно заметно при расчете эффектов реактивности, где важна не только точность представления самих сечений, но и их производных.

Аппроксимация сечений функцией многих переменных (обычно полиномом) обладает определенными достоинствами. Во первых, для построения аппроксимации набор точек может быть рассчитан на любой сетке, применение регулярной прямоугольной сетки не требуется. Во вторых, в результате аппроксимации получается набор данных (коэффициентов полинома), который значительно меньше числа точек интерполяционных таблиц. Для практических расчетов это достаточно важный факт, так как размер малогрупповой библиотеки для расчета реактора использующий процедуру табличной интерполяции достигает сотни мегабайт. Основная трудность использования процедуры аппроксимации заключается в выборе вида аппроксимирующего полинома, который обычно подбирается вручную методом проб и ошибок.

В докладе представлен согласованный подход к решению этих двух вопросов. Расчетная сетка для расчета сечений строится в виде комбинации одномерной сетки по выгоранию для базового расчета, которая либо задается пользователем либо определяется самой программой выгорания, и сетки для дополнительных расчетов на каждом шаге по выгоранию, построенной на основе равномерно-распределенной ЛП, последовательности Соболя (Соболев, 1967). Эти вопросы рассматриваются в разделе 2. Комбинированная процедура интерполяции/аппроксимации зависимости сечений от переменных состояния ТВС изложена в разделе 3. Нейтронно-физические сечения представляются в виде суммы двух членов: базовых сечений, которые зависят только от выгорания и отклонений, которые зависят как от выгорания, так и от теплофизических параметров ячейки.

Зависимость базовых сечений от выгорания интерполируется кубическим сплайном. Зависимость отклонений аппроксимируется многомерным полиномом. Особенностью данной методики является автоматическое построение оптимальной полиномиальной зависимости отклонений. Оптимальность понимается в смысле минимизации числа членов полинома для заданной точности аппроксимации. Решение этой задачи выполнено в рамках регрессионного анализа, члены многомерного полинома выбираются программой шаговой регрессии. Основные понятия регрессионного анализа используемые в данной методике изложены в разделе 3.1. В разделе 3.2 описан алгоритм метода шаговой регрессии, использующийся для построения оптимального аппроксимирующего полинома. В качестве примера применения предложенной методики в разделе 4 строятся зависимости двухгрупповых сечений ячейки реактора ВВЭР от выгорания, плотности и температуры теплоносителя, температуры топлива и концентрации бора. Раздел 5 содержащий заключение и выводы завершает статью.

2. Построение расчетной сетки для области определения переменных состояния ячейки

В данном разделе рассматривается вопрос выбора точек в пространстве переменных состояния ячейки или кассеты реактора используемых при расчете сечений.

Исходная область изменения переменных состояния обычно описывается неравенствами

$$x_m^{\min} \leq x_m \leq x_m^{\max}, \quad m = 1, \dots, M$$

формируя многомерный параллелепипед. Путем масштабирования независимых переменных этот параллелепипед отображается на единичный куб и задача выбора пробных точек в пространстве параметров сводится к задаче построения расчетной сетки в многомерном единичном кубе. Обычно полагают, что наиболее равномерное покрытие многомерного куба обеспечивает кубическая решетка состоящая из N^M точек с координатами

$$\left\{ \frac{m_1 + 1/2}{N}, \frac{m_2 + 1/2}{N}, \dots, \frac{m_M + 1/2}{N} \right\},$$

где m_1, \dots, m_M независимо принимают все значения $0, 1, \dots, N - 1$. Пример такой решетки для двумерного куба с 16 точками показан на Рис. 1а). Однако такая решетка оптимальна только в одномерном случае, а с увеличением размерности куба ее равномерность быстро ухудшается. Доказательство этого утверждения дано еще в 1957 году в работе И. М. Соболя (1957). Приведем пример, взятый из книги (Соболев и Статников, 1981), иллюстрирующий это утверждение. Сравним кубическую решетку, изображенную на Рис.1а) с сеткой показанной на Рис. 1б). В обоих случаях, каждой из 16 ячеек куба принадлежит одна расчетная точка, так что казалось бы, что равномерность расположения этих точек примерно одинакова.

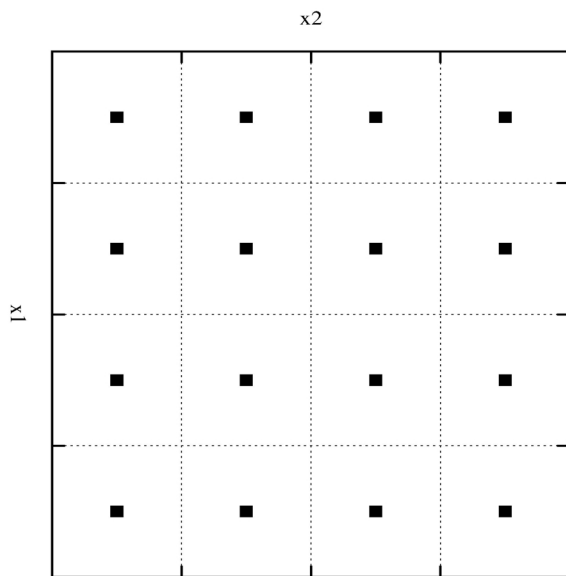


Рис. 1а) Кубическая решетка для двумерного куба с 16 точками

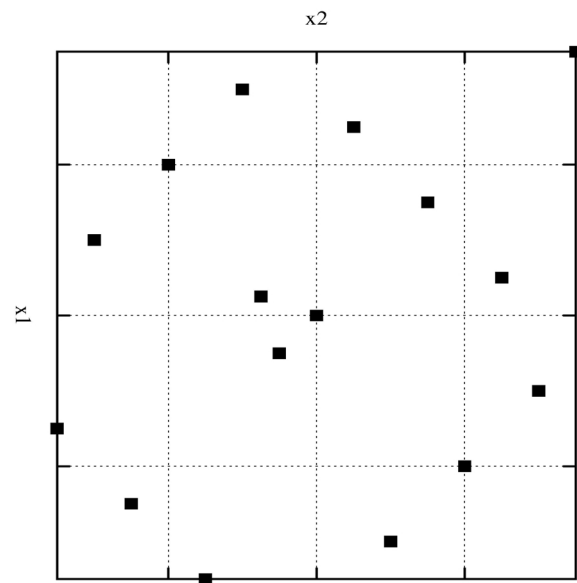


Рис. 1б) Равномерно-распределенная сетка для двумерного куба с 16 точками

Предположим, что требуется аппроксимировать методом наименьших квадратов функцию $f(x_1, x_2)$, которая сильно зависит только от одной переменной $f(x_1, x_2) \approx f(x_1)$. В этом случае, если используется кубическая решетка получаются только 4 различных значения функции, которые повторяются 4 раза. При использовании расчетной сетки, показанной на Рис. 1б), мы имеем 16 различных значений функции, дающих гораздо лучшее представление о диапазоне изменения функции и позволяющих построить лучшую аппроксимацию. В многомерном случае кубическая решетка еще хуже, так как потеря информации может быть еще больше: вычислили N^M очек, а получили всего $N^{1/M}$ различных значений функции. Этот пример, хотя и представляет собой самый неблагоприятный случай для кубической решетки, имеет существенное практическое значение, так как нейтронно-физические сечения обычно сильно зависят от одних переменных состояния и слабо от других. Более того для различных нейтронных сечений характер зависимости различный, что не позволяет оптимизировать сетку с учетом функциональной зависимости.

В данной работе для равномерного покрытия области изменения параметров ячейки используются пространственные сетки построенные на основе равномерно-распределенных последовательностей (Соболев и Статников, 1981). Пусть $P_1, P_2, \dots, P_n, \dots$ последовательность точек принадлежащих многомерному кубу C . Рассмотрим в C произвольный M -мерный параллелепипед Π со сторонами параллельными координатным осям. Обозначим через $S_N(\Pi)$ последовательность точек P_n с номерами $1 \leq n \leq N$ принадлежащих параллелепипеду Π . Последовательность точек $P_1, P_2, \dots, P_n, \dots$ называется *равномерно-распределенной*, если для любого Π

$$\lim_{N \rightarrow \infty} \frac{S_N(\Pi)}{N} = V_{\Pi} \quad (1)$$

где V_{Π} объем параллелепипеда. Применение таких последовательностей в вычислительной математике началось в начале 60-х годов, когда удалось построить последовательности, для которых скорость сходимости определяемая по формуле (1) близка к наилучшей, а

равномерность расположения точек начинается с небольших N . Наиболее известны и широко используются ЛП_τ-последовательности Соболя (Соболь, 1967).

В данной работе применяется расчетная сетка построенная на основе ЛП_τ-последовательности Соболя. Алгоритм расчета ЛП_τ-последовательности разработан Антоновым и Салеевым (1979), программа расчета на языке FORTRAN приведена в работе (Bratley and Fox, 1988). Определение ЛП_τ-последовательностей и анализ их свойств выходит за рамки данной статьи, заинтересованному читателю рекомендуется обратиться к оригинальным работам (Соболь, 1967; 1969; Соболь и Статников, 1981). Мы лишь отметим, что для практических расчетов целесообразно выбирать первые 2^k точек ЛП_τ-последовательности, так как в этом случае равномерность полученной сетки максимальна.

При построении расчетной сетки для расчета сечений возникает проблема комбинирования одномерной сетки по выгоранию, которая выбирается пользователем или определяется самой программой исходя из точности расчета и многомерной сетки для остальных переменных состояния ячейки. В нашем случае для расчета отклонений используется равномерно-распределенная последовательность Соболя. Алгоритм построения расчетной сетки приведен ниже:

1. Оценивается число точек для расчета выгорания N_{burnup} ;
2. Выбирается число точек дополнительных расчетов на один шаг по выгоранию J_{dev} ; желательно чтобы число точек сетки отклонений $N_{dev} = J_{dev} \times N_{burnup}$ было равно 2^k , где k - произвольное целое число;
3. Строится равномерно-распределенная последовательность Соболя $\{x_m^n\}$, $m = 1, \dots, M$; $n = 1, \dots, N_{dev}$, где M -размерность пространства переменных состояния ячейки без учета выгорания;
4. Проводится расчет выгорания с равномерной сеткой заданной пользователем, а для расчета отклонений на каждом шаге по выгоранию используется J_{dev} точек равномерно-распределенной последовательности Соболя.

3. Комбинированная процедура интерполяции/аппроксимации нейтронных сечений

В результате расчета по "спектральной" программе получается два набора сечений:

- базовый набор сечений $\Sigma^{base}(B_n)$, $n = 1, \dots, N_{burnup}$, которые зависят только от выгорания
- набор сечений $\Sigma^{dev}(B_n, T_f^{nj}, \rho_c^{nj}, T_c^{nj}, c_{bor}^{nj})$, $n = 1, \dots, N_{burnup}$; $j = 1, \dots, J_{dev}$, рассчитанных в точках отклонений, которые зависят как от выгорания, так и от остальных переменных состояния ячейки.

Зависимость базового набора сечений от выгорания плохо описывается полиномами и для представления этой зависимости используется интерполяция кубическим сплайном. Для расчета кубического сплайна используется пакет PCNIP (Fritsch and Carlson, 1980) из библиотеки SLATEC (Buzbee, 1984). Программа рассчитывает сплайн в эрмитовом кубическом представлении. Для расчета сечений это представление переводится в кусочно-полиномиальное представление. Программа перевода и расчета значений сплайн функции написана авторами. Подробное изложение сплайн интерполяции и оценка точности приведена в отчете (Зимин и Семенов, 2002), здесь мы только отметим что для проведенных расчетов ошибка сплайн интерполяции не вносит существенного вклада в ошибку представления сечений.

Для уменьшения функциональной зависимости от выгорания для аппроксимации вместо сечений используются отклонения сечений от базового набора сечений:

$$\Delta\Sigma^{dev} = \begin{cases} \Sigma^{dev}(B_n, T_f^{nj}, \rho_c^{nj}, T_c^{nj}, c_{bor}^{nj}) - \Sigma^{base}(B_n), \\ n = 1, \dots, N_{burnup}; j = 1, \dots, J_{dev} \\ 0 = \Sigma^{base}(B_n) - \Sigma^{base}(B_n), \\ n = 1, \dots, N_{burnup}; \end{cases}$$

Набор отклонений $\Delta\Sigma^{dev}$ аппроксимируется многомерными полиномами методом наименьших квадратов. Для выбора членов аппроксимирующего полинома используется метод шаговой регрессии. Перед тем как приступить к описанию процедуры аппроксимации введем основные понятия метода наименьших квадратов и регрессионного анализа.

3.1 Основные понятия линейного регрессионного анализа

Рассмотрим задачу построения функциональной зависимости переменной y , называемой также откликом, от независимых переменных x_1, x_2, \dots, x_M , называемых регрессорами. В рамках линейного регрессионного анализа рассматриваются только линейная зависимость, то есть

$$y = b_0 + b_1 x_1 + \dots + b_M x_M + error$$

где b_m - неизвестные коэффициенты модели, $error$ - ошибка регрессионной модели.

Пусть задан набор из N точек, которые содержат значения независимых переменных и отклика

$$x_{n1}, x_{n2}, \dots, x_{nM}, y_n, \quad n = 1, \dots, N$$

Задача нахождения коэффициентов модели по эти точкам формулируется в рамках *метода наименьших квадратов*: найти набор коэффициентов который минимизирует квадратичную норму вектора ошибки модели, т.е. найти $b \in \mathbb{R}^{M+1}$, такой что $\|Xb - y\|_2$ минимальна,

$$\text{где } X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1M} \\ 1 & x_{21} & x_{22} & \dots & x_{2M} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{N2} & \dots & x_{NM} \end{bmatrix} \in \mathbb{R}^{(M+1) \times N},$$

$$y = col\{y_1, y_2, \dots, y_N\} \in \mathbb{R}^N - \text{вектор отклика};$$

$$b = col\{b_0, b_1, b_2, \dots, b_M\} \in \mathbb{R}^{M+1} - \text{вектор неизвестных коэффициентов}.$$

Для решения задачи метода наименьших квадратов разработано несколько алгоритмов (Trefethen and Bau, 1997). В данной работе применяется алгоритм на основе QR разложения, устойчивый к ошибкам округления (Miller, 1992), реализованный в модуле LSQ Аланом Миллером на языке FORTRAN 90 (Miller, 2000).

Полученный вектор коэффициентов минимизирует сумму квадратов ошибки:

$$SS_E = \|y - \hat{y}\|_2^2,$$

где $\hat{y} = Xb$ - предсказанное значение отклика

Число степеней свободы связанных с ошибкой равно числу точек N минус число коэффициентов модели $M+1$. Средний квадрат ошибки определяется как сумма квадратов ошибки деленная на соответствующее число степеней свободы:

$$MS_E = SS_E / (N - M - 1)$$

Средний квадрат ошибки является несмещенной оценкой дисперсии ошибки σ^2 . Квадратный корень из оценки дисперсии определяет оценку стандартного отклонения, называемой также *стандартной ошибкой регрессионной модели*.

Аналогично определяется *сумма квадратов регрессии*:

$$SS_R = \|\hat{y} - \bar{y}\|_2^2,$$

где \bar{y} - среднее значение отклика;

и *средний квадрат регрессии*:

$$MS_R = SS_R / (N - 1),$$

где $N - 1$ - число степеней свободы регрессионной модели.

3.2 Значимость регрессионной модели и ее членов

В рамках регрессионного анализа часто возникает вопрос о значимости модели и отдельных членов регрессионного уравнения. Для оценки значимости статистики вводят понятие *нуль гипотезы*, которую можно выразить следующим образом: "нуль гипотеза истинна, если нет линейной зависимости между откликом и какой-либо независимой переменной". Нуль гипотезу можно выразить уравнением:

$$H_0 : b_1 = b_2 = \dots = b_M = 0$$

Отрицание нуль гипотезы H_1 можно сформулировать следующим образом

$$H_1 : b_m \neq 0 \text{ по крайней мере для одного } j$$

Если нуль гипотеза не верна, по крайней мере одна из независимых переменных вносит значимый вклад в модель и можно сделать вывод о функциональной зависимости между откликом и по крайней мере одной независимой переменной. Для оценки значимости модели и проверки нуль гипотезы используется значение *F-критерия*, который рассчитывается следующим образом

$$F = MS_R / MS_E$$

Нуль гипотеза может быть отвергнута, если рассчитанное значение *F*-критерия больше чем критическое значение *F*-распределения $F_{crit} = F_{\alpha, M, N-M-1}$ для уровня значимости α с M и $N - M - 1$ числом степеней свободы. Уровень значимости α определяет вероятность того, что нуль-гипотеза верна, то есть модель не значима. В практике статистического анализа обычно используются уровни значимости 10%, 5% и 1%, при этом, чем меньше уровень значимости, тем больше должно быть значение *F*-критерия и тем больше уверенности в значимости модели.

Аналогично определяется нуль-гипотеза H_0 и отрицание нуль-гипотезы H_{1m} для отдельных членов регрессионного уравнения:

$$H_{0m} : b_m = 0$$

$$H_{1m} : b_m \neq 0$$

Если нуль-гипотеза H_{0m} отвергается, соответствующий член вносит значимый вклад в модель, в противном случае эту переменную можно исключить из регрессионного уравнения. Для оценки значимости коэффициента b_m рассматриваются две модели: модель которая включает этот член с коэффициентами $\{b_0, b_1, \dots, b_m, \dots, b_M\}$ и модель без этого члена с коэффициентами $\{b_0, b_1, \dots, b_{m-1}, b_{m+1}, \dots, b_M\}$. Для проверки нуль гипотезы H_{0m} рассчитывается частный $F(b_m)$ -критерий

$$F(b_m) = [SS_R(\text{with } b_m) - SS_R(\text{without } b_m)] / MS_E(\text{with } b_m), \quad (2)$$

где

$SS_R(\text{with } b_m)$ - сумма квадратов регрессии для модели с коэффициентом b_m ;

$SS_R(\text{without } b_m)$ - сумма квадратов регрессии для модели без коэффициента b_m ;

$MS_E(\text{with } b_m)$ - средний квадрат ошибки для модели с коэффициентом b_m .

Для проверки нуль гипотезы расчетное значение $F(b_m)$ сравнивается с теоретическим. Если значение превосходит теоретическое значение, нуль-гипотеза отбрасывается и член с коэффициентом b_m остается в модели, в противном случае этот член можно из модели исключить.

3.3 Формирование оптимального аппроксимирующего полинома

Одной из основных проблем построения функциональной модели зависимости сечений от параметров состояния ячейки является выбор членов полинома, которые имеют значимый вклад. При анализе одномерных зависимостей достаточно просто методом проб и ошибок определить степень полинома, который описывает зависимость с достаточной для практики точностью. В реальных задачах сечения зависят от нескольких параметров и с увеличением размерности пространства переменных задача построения оптимального полинома значительно усложняется. Например для пространства из 5 параметров, полином 4-ой степени имеет 126 членов. Очевидно, что различные члены вносят различный вклад в построенную зависимость и необходимо оставить только те члены, которые значимы. Решение задачи выбора наилучшего подмножества членов полиномиальной зависимости и рассматривается в данном разделе.

Исходя из анализа одномерных зависимостей достаточно просто оценить порядок полиномиальной зависимости по каждой из исходных переменных. Зная этот порядок, исходная модель строится как тензор-произведение одномерных полиномов по всем направлениям. Допустим, что есть три переменных $\{z_1, z_2, z_3\}$, степень которых не превышает $\{5, 2, 3\}$, тогда члены полиномиального базиса строятся следующим образом

$$x_m = P_i(z_1) \times P_j(z_2) \times P_k(z_3),$$

где $P_n(z)$ - полином Лежандра степени n , $\{0 \leq i \leq 5, 0 \leq j \leq 2, 0 \leq k \leq 3\}$, $i + j + k \leq 5$

Использование ортогональных полиномов Лежандра рекомендуется, так как это уменьшает число обусловленности матрицы в методе наименьших квадратов, и улучшает работу алгоритма выбора членов полиномиальной зависимости.

Выбор тех членов, которые оптимально описывают зависимость сечений осуществляется методами регрессионного анализа. В первую очередь сформируем критерий оптимальности. Для использования модели желательно, чтобы она описывала зависимость сечений с некоторой заданной пользователем точностью. В качестве критерия точности можно использовать относительную стандартную ошибку регрессионной модели, определяемую следующим образом

$$\tilde{\sigma}_E = \sqrt{MS_E} \times 100 / \bar{y} \quad (3)$$

Стандартная ошибка модели обычно уменьшается при включении в модель новых членов. С другой стороны для минимизации времени счета как самой модели, так и времени предварительных расчетов для получения набора сечений используемых в построении модели, желательно минимизировать число членов полиномиальной зависимости. Выбор оптимального решения, который с одной стороны обеспечивает приемлемую точность модели, а с другой минимизирует число членов модели и является нашей задачей.

Оптимальное решение этой задачи требует перебора всех возможных комбинаций членов полиномиальной зависимости. В реальных задачах, где число возможных членов зависимости несколько сотен, этот подход не реализуем из-за ограничения на время расчета. Поэтому рассматривается применение алгоритмов, которые хоть и не гарантируют оптимального решения, но позволяют построить модель близкую к оптимальной. К таким алгоритмам, разработанными в рамках регрессионного анализа

относятся: *метод включения, метод исключения, метод шаговой регрессии* и ряд других, которые здесь не рассматриваются. Подробное изложение этих методов представлено в работах (Дрейпер и Смит, 1987; Miller, 2002).

В *методе включения* процесс построения модели начинается с модели которая содержит только константу. Новые переменные добавляются в модель исходя из максимальной величины частного $F(b)$ -критерия, определяемого по формуле (2). На первом шаге выбирается член с максимальным значением $F(b)$ -критерия, рассчитанного для модели с одной переменной и константой. Метод включения обычно использует некоторое критическое значение F_{in} для того чтобы определить что переменная достаточно значима для включения в модель. На следующем шаге определяется следующая переменная, которая дает максимальный $F(b)$ -критерий модели с двумя предикторными переменными. Процедура продолжается до тех пор, пока либо модель достигает заданное значение точности, т.е. относительная стандартная ошибка меньше заданной, либо если среди оставшихся переменных не включенных в модель нет переменной с $F(b)$ -критерием превышающим заданный критерий F_{in} .

Метод исключения работает строго в противоположном направлении. Построение модели начинается с уравнения, которое включает в себя все возможные переменные. Переменные удаляются из модели исходя из заданного критерия F_{out} . Если рассчитанный $F(b)$ -критерий для переменной меньше заданного, то переменная исключается из уравнения.

Метод шаговой регрессии представляет собой комбинацию этих двух подходов. Как и в методе включения построение модели начинается с уравнения содержащего только константу, в которое добавляется одна предикторная переменная x_1 . Далее находится переменная x_2 с максимальным значением $F(b_m)$ -критерия и также добавляется в модель. Для модели с двумя предикторными переменными находятся частные $F(b_m)$ -критерии для обеих переменных. Величины частного $F(b_m)$ -критерия этих переменных сравниваются с критерием для исключения F_{out} . Соответствующая предикторная переменная сохраняется в уравнении или удаляется из него исходя из результатов проверки. Такая проверка "наименее полезного предиктора", в уравнении проводится на каждом шаге метода. Если переменная исключается, то регрессионное уравнение пересчитывается с учетом всех оставшихся в нем предикторных переменных. Переменные, которые не входят на данном шаге в модель проверяются на включение. Если переменная добавляется, то опять возвращаются к проверке критерия на исключение. Процесс прекращается, если достигнута заданная точность модели или если нет больше предикторов не вошедших в модель, которые бы удовлетворяли критерию на включение F_{in} .

Описанные методы выбора оптимального набора предикторов реализованы во многих статистических пакетах обработки данных, например в SPSS, Statistica, статистическом пакете системы Matlab. Для своей программы мы использовали модуль SUBSET разработанный Аланом Миллером (Allan F. Miller) на языке FORTRAN-90 (Miller, 2001).

4. Аппроксимация двухгрупповых сечений ячейки реактора ВВЭР

Применение предложенной методики аппроксимации демонстрируется на примере расчета двухгрупповых сечений трехзонной цилиндризованной ячейки реактора ВВЭР с урановым топливом. Радиусы зон топлива, оболочки, теплоносителя равны 0.386, 0.4582,

0.6694 см, соответственно. Концентрации изотопов в материалах ячейки приведены в Таблице 1. Топливо содержит равновесные концентрации ксенона и йода. Плотность энерговыделения ячейки - 108 Вт/см³. Для расчетов сечений использовалась программа UNK (Цибульский и Давиденко, 2002). Рассматривались зависимости сечений от параметров ячейки на следующих интервалах:

- *burnup* - выгорание [0,60] МВт×сут/кгU;
- ρ - плотность теплоносителя [0.25, 1] г/см³;
- T_c - температура теплоносителя [293, 603] К;
- T_f температура топлива [293, 1793] К;
- c_B концентрация борной кислоты [0, 1500] ppm.

В процедурах аппроксимации использовались масштабированные переменные, рассчитанные из исходных переменных x следующим образом

$$\xi = (2x - x_{\max} - x_{\min}) / (x_{\max} - x_{\min}),$$

где x_{\min}, x_{\max} - минимальное и максимальное значение переменной x .

Рассматривалась зависимость следующих двухгрупповых макросечений:

$$\{\Sigma_{tr1}, \Sigma_{tr2}, \Sigma_{a1}, \Sigma_{a2}, \nu\Sigma_{f1}, \nu\Sigma_{f2}, \Sigma_{f1}, \Sigma_{f2}\},$$

коэффициента размножения ячейки k_{inf} и микросечений поглощения тепловой группы ксенона σ_{a2}^{Xe-135} и самария σ_{a2}^{Sm-149} . Шаг расчета по выгоранию равен 0.5 МВт×сут/кгU. Для расчета базовых сечений использовались следующие значения параметров: $\rho = 0.72$ г/см³, $T_c = 575$ К, $T_f = 950$ К, $c_B = 500$ ppm. Для расчета отклонений использовалась ЛП_т сетка, на каждом шаге по выгоранию рассчитывалось 4 точки сетки. Две двумерные проекции используемой расчетной сетки показаны на Рис. 2.

Зависимость базового набора сечений от выгорания аппроксимировалась кубическим сплайном. На Рис. 3 показаны зависимости от выгорания сечения поглощения тепловой и быстрой группы. Для построения сплайн-интерполяции использовались 64 точки: каждая вторая точка расчета начиная с первой. Оставшиеся точки составляют контрольный набор данных, применявшийся для контроля точности сплайн-интерполяции. Погрешности сплайн интерполяции, рассчитанные на контрольном наборе представлены в Таблице 2. Представленные данные показывают что для шага сетки по выгоранию 1 МВт×сут/кгU кубический сплайн описывает зависимость базовых сечений с достаточной для практики точностью. В случае необходимости точность интерполяции можно легко повысить уменьшив шаг сетки.

Полиномиальная зависимость сечений подбиралась программой шаговой регрессии исходя из условия, что относительная стандартная ошибка регрессионной модели, определенная уравнением (3) не превышает 0.05 %.

Таблица 1. Исходные концентрации изотопов ячейки реактора в номинальных условиях. ($burnup = 0$ МВт×сут/кгU, $\rho = 0.72$ г/см³, $T_c = 575$ К, $T_f = 950$ К, $c_B = 500$ ppm).

Материал	Изотоп	Концентрация
Урановое Оксидное Топливо	U-235	8.73700e-04
	U-238	1.87440E-02
	O-16	3.92350E-02
	Xe-135	9.74049E-09
	I-135	2.07109E-08
Оболочка	Zr-nat	4.23E-03
Теплоноситель	H-1	4.81354E-02
	O-16	2.40677E-02
	B-10	3.96154E-06
	B-11	1.61071E-05

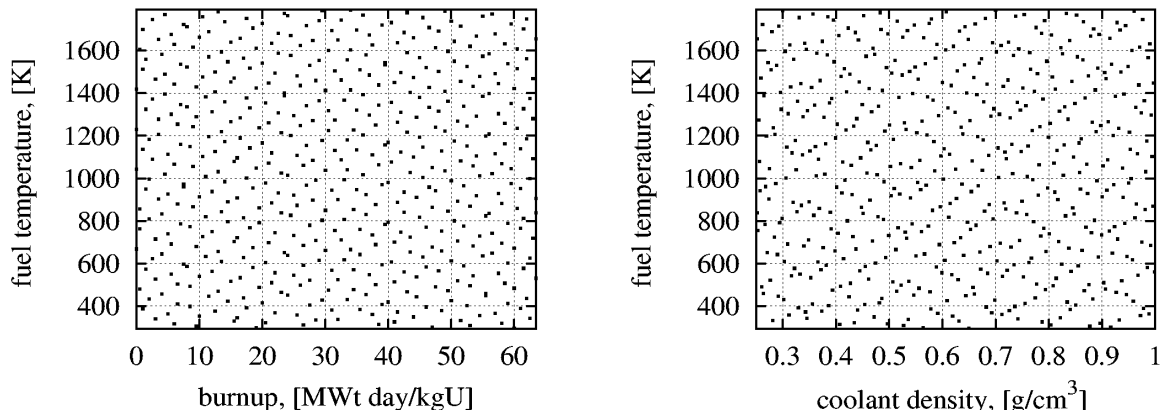


Рис. 2. Двумерные проекции расчетной сетки

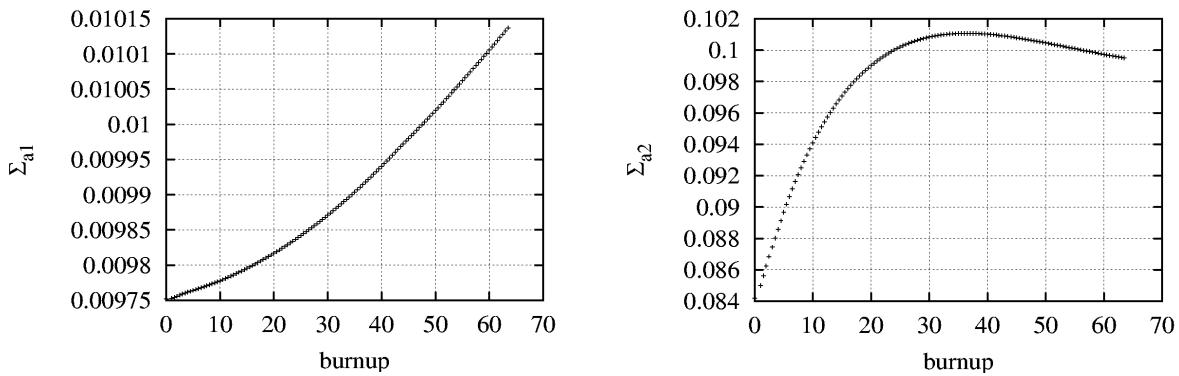


Рис. 3. Примеры зависимости базовых сечений от выгорания

Таблица 2. Среднеквадратичная Δ^{sqrt} и максимальная Δ^{max} ошибки кубической сплайн интерполяции зависимости базовых сечений от выгорания

Ошибка, (%)	Σ_{tr1}	Σ_{tr2}	Σ_{a1}	Σ_{a2}	$\nu\Sigma_{f1}$	$\nu\Sigma_{f2}$	Σ_{f1}	Σ_{f2}	$\Sigma_s^{1->2}$	k_{inf}	σ_{a2}^{Xe}	σ_{a2}^{Sm}
Δ^{sqrt}	0.000	0.000	0.001	0.010	0.006	0.018	0.006	0.016	0.000	0.006	0.003	0.003
Δ^{max}	0.001	0.001	0.002	0.093	0.039	0.133	0.034	0.109	0.002	0.038	0.025	0.021

Исходный многомерный полиномиальный базис строился из полиномов Лежандра степени $\{6,6,5,4,4\}$ для соответствующих переменных: $\{burnip, \rho, T_c, T_f, c_B\}$. Ограничение на степень полинома по каждой переменной определялось заранее из результатов одномерных расчетов. Число исходных членов базиса составляет 448. Число расчетных точек 640: 128 точек базового расчета и 512 точек расчета отклонений. Для выбора оптимального набора аппроксимирующих членов использовался метод шаговой регрессии.

Напомним, что многомерный аппроксимирующий полином представляется в виде тензор-произведения одномерных полиномов. В Таблицах 3 и 4 представлены степени одномерных полиномов полученные для сечений поглощения быстрой и тепловой группы. Аппроксимирующий полином для сечения поглощения содержит 17 членов для быстрой группы и 51 член для тепловой. Число членов многомерного полинома для различных сечений показано на Рис. 4. В среднем около 20 членов используется для аппроксимации сечений быстрой группы и около 50 членов для сечений тепловой группы. Для повышения быстродействия расчета набора сечений члены полиномов используемые для аппроксимации группируются в единый полином для данного набора. Стоит отметить что общее число членов в полиноме менее 80. На Рис. 5 крестом помечены члены полинома используемые в аппроксимации каждого сечения. При вызове библиотеки сечений этот полином рассчитывается только один раз для каждой точки переменных состояния. В результате значение каждого сечения рассчитывается путем скалярного произведения вектора коэффициентов на вектор значений членов аппроксимирующего полинома.

Точность аппроксимации тестировалась как на исходной последовательности точек так и на специальном тестовом наборе сечений рассчитанных по программе UNK на расчетной сетке построенной в виде тензор произведения исходной одномерной расчетной сетки по выгоранию и случайном наборе данных по остальным параметрам. Также как и в исходном расчете использовалось 4 точки расчета отклонений на одну точку по выгоранию. В Таблицах 5 и 6 представлены погрешности аппроксимации сечений полученные на исходной и проверочной последовательности точек. Кроме непосредственно сечений в столбце k_{inf}^{Σ} представлена оценка погрешности коэффициента размножения ячейки, полученного по аппроксимированным константам.

Для всех сечений среднеквадратичная ошибка не превышает 0.065%. Сравнение результатов представленных в таблицах 5 и 6 показывают, что ошибки аппроксимации полученные на исходной последовательности точек вполне представительны и могут применяться как мера точности построенной аппроксимации. Ошибки аппроксимации сечения перевода $\Sigma^{1 \rightarrow 2}$ и сечения поглощения тепловой группы Σ_{a2} в зависимости от плотности теплоносителя показаны на Рис 6. Погрешность аппроксимации этих сечений значительно превышающая среднеквадратичную погрешность наблюдается только для низких значений плотности теплоносителя, характерных для аварийных режимов. Для моделирования аварийных режимов полученная точность аппроксимации констант является вполне удовлетворительной.

В заключении, чтобы продемонстрировать сложность аппроксимируемых зависимостей, на Рис. 7 представлены несколько двумерных проекций зависимостей сечений ячейки. По этим рисункам можно сделать вывод, что аппроксимируемые зависимости действительно сложные функции переменных состояния, зависимости различны для разных сечений и выбор формы аппроксимирующего полинома - далеко нетривиальная задача.

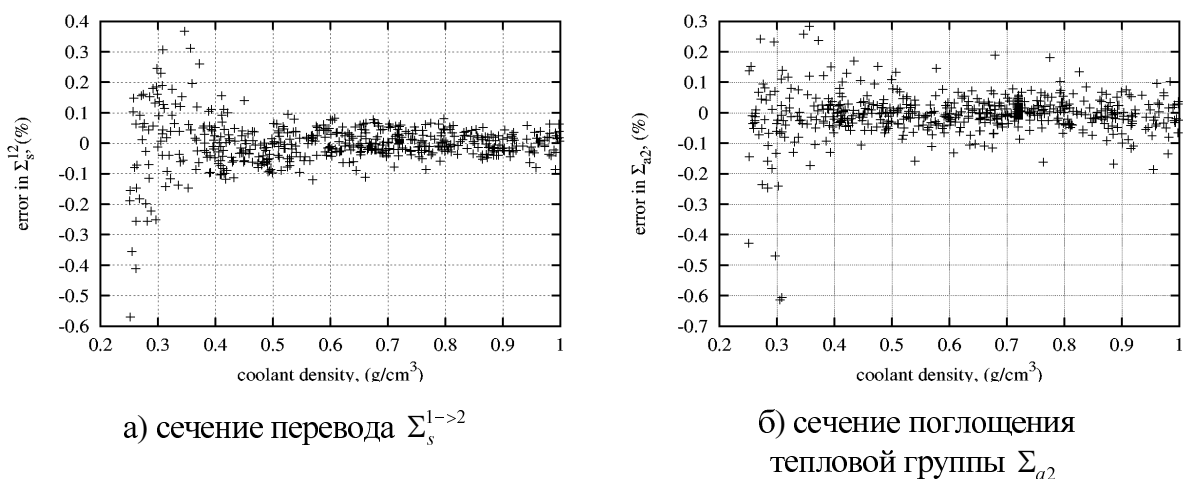


Рис. 6. Ошибка аппроксимации зависимости сечений от плотности теплоносителя

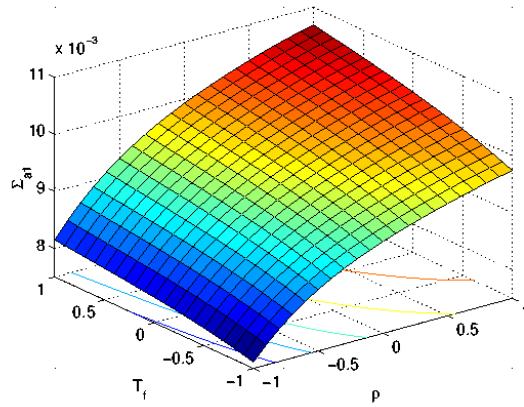
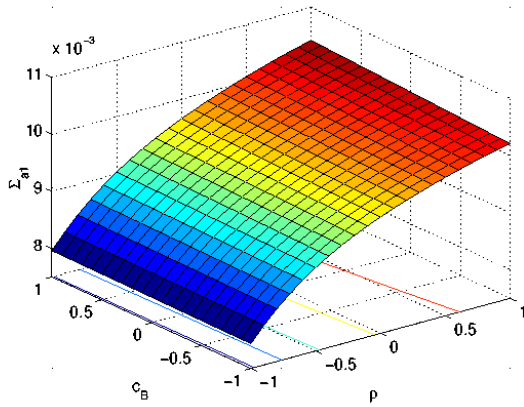
5. Заключение

В статье предложена методика представления зависимости малогрупповых нейтронно-физических сечений от выгорания и теплофизических параметров ячейки. Нейтронно-физические сечения представляются в виде суммы двух членов: базовых сечений, которые зависят только от выгорания и рассчитываются при номинальных значениях теплофизических параметров и отклонений, которые зависят как от выгорания, так и от теплофизических параметров ячейки. Зависимость базовых сечений от выгорания аппроксимируется кубическим сплайном. Зависимость отклонений аппроксимируется многомерным полиномом. Особенностью данной методики является построение оптимальной полиномиальной зависимости отклонений для каждого сечения для заданной точности аппроксимации. Оптимальность понимается в смысле минимизации числа членов аппроксимирующего полинома. Решение этой задачи выполнено в рамках регрессионного анализа, члены полинома выбираются из исходного набора базисных функций программой шаговой регрессии.

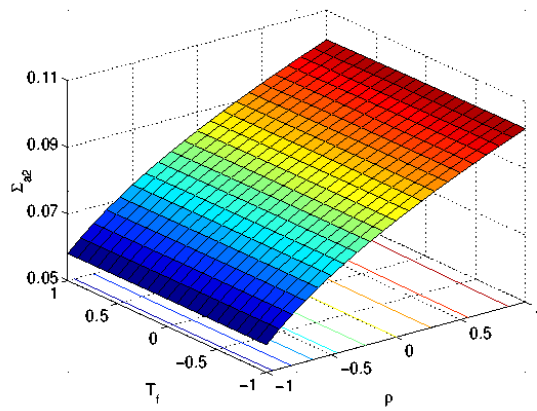
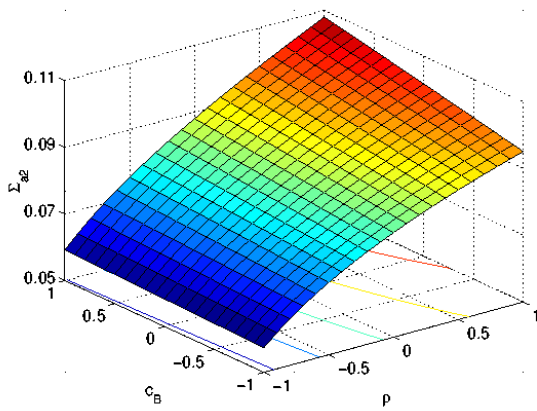
Предложенная методика демонстрируется на примере построения зависимости двухгрупповых сечений ячейки реактора ВВЭР от выгорания, плотности и температуры теплоносителя, температуры топлива и концентрации борной кислоты. Исходные данные были получены по программе UNK. Среднеквадратичная погрешность аппроксимации зависимости базовых сечений от выгорания кубическим сплайном не превышает 0.02 %. Полиномиальная модель отклонений строилась методом шаговой регрессии для заданного значения стандартной ошибки 0.05%. Заданную точность модели обеспечивает пятимерный полином, который содержит не больше 20 членов для сечений быстрой группы и около 50 членов для сечений тепловой группы. Построенная полиномиальная зависимость отклонений тестировалась как на исходных данных, так и на проверочной последовательности точек выбранных случайным образом. Среднеквадратичная ошибка лежит в пределах заданной точности на обеих последовательностях точек. Результаты показывают, что исходные точки, использованные для построения модели, можно также применять для контроля точности модели.

Благодарности

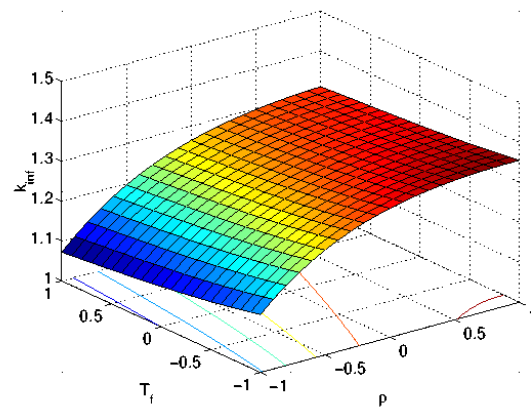
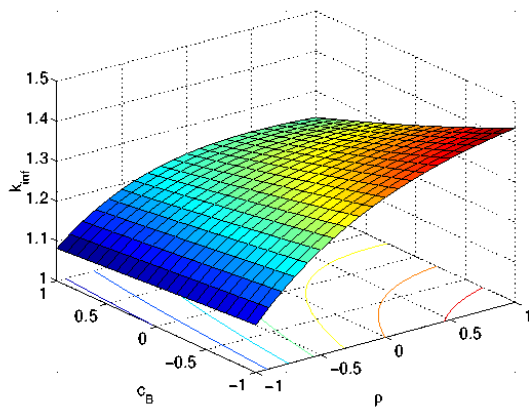
Авторы выражают глубокую признательность сотрудникам ИЯР РНЦ Курчатовский Институт к.т.н. В. Д. Давиденко и к.т.н. В. Ф. Цыбульскому за любезно предоставленную программу UNK и многочисленные консультации, а также сотруднику нашей лаборатории Е. В. Чернову за помощь в проведении расчетов.



а) сечение поглощения быстрой группы Σ_{a1}



б) сечение поглощения тепловой группы Σ_{a2}



в) коэффициент размножения k_{inf}

Рис. 7. Двумерные проекции зависимости некоторых сечений ячейки: слева - от плотности теплоносителя и концентрации борной кислоты; справа - от плотности теплоносителя и температуры топлива.

Литература

- Антонов, И.А. и В.М. Салеев, (1979). Экономичный способ вычисления ЛП₁-последовательностей, *Ж. Вычисл. Матем. и Матем. Физ.*, **19**, 243-245.
- Дрейпер, Н. и Г. Смит, (1987). *Прикладной регрессионный анализ*. Финансы и Статистика. Москва, 1987.
- Зимин В.Г. и А.А. Семенов (2002). Методология построения зависимостей нейтронно-физических сечений для малогрупповых реакторных расчетов на основе регрессионного анализа. Отчет МИФИ.
- Семенов, А.А., С.Б. Выговский, Н.В. Шукин, и О.А. Будникова, (1998). Особенности программ константного сопровождения комплексных динамических кодов, *Интегрированные математические модели и программные комплексы в ядерной энергетике. Тезисы лекций и докладов школы-семинара секции динамики*, Москва, МИФИ, СОЛ "Волга", 6-12 сентября 1998 г., МИФИ, 310-317.
- Соболь, И.М., (1957). Многомерные интегралы и метод Монте-Карло. *Докл. АН СССР*, **114**, 706-709.
- Соболь, И.М., (1967). О распределении точек в кубе и приближенном вычислении интегралов, *Ж. Вычисл. Матем. и Матем. Физ.*, **7**, 784--802
- Соболь, И.М., (1969). *Многомерные квадратурные формулы и функции Хаара*. Наука, Москва.
- Соболь, И.М. и Р.Б. Статников, (1981). *Выбор оптимальных параметров в задачах со многими критериями*. Наука, Москва.
- Цибульский, В.Ф. и В.Д. Давиденко (2002). Пакет программ UNK для расчета ядерных реакторов. Описание применения. Отчет No. 35-410-4/10, ИЯР РНЦ «Курчатовский Институт», февраль 2002.
- Bratley, P. and B. L. Fox, (1988). Algorithm 659: Implementing Sobol's quasirandom sequence generator. *ACM Transactions on Mathematical Software*, **14**, no. 1, 88-100.
- Buzbee, B. L., (1984). The SLATEC common mathematical library. *Sources and Development of Mathematical Software*, Ed. by W. R. Cowell, Prentice-Hall, Englewood Cliffs, New Jersey, 1984.
- Fritsch, F. N. and R. E. Carlson, (1980). Monotone piecewise cubic interpolation. *SIAM J. Num. Anal.*, **17**, 238-246.
- Miller, A. J. (1992) AS274: Least square routines to supplement those of Gentleman. *Appl. Statist.*, **41**, 458-478.
- Miller, A. J., (2000). LSQ: Module for unconstrained linear least-squares calculations. Available at Allan Miller website <http://users.bigpond.net.au/amiller>.
- Miller, A. J., (2001). SUBSET: Interactive program for finding best-fitting subsets of variables, Available at Allan Miller website <http://users.bigpond.net.au/amiller>.
- Miller, A. J., (2002). *Subset selection in regression*. 2nd Ed., Chapman and Hall, New York.
- Trefethen, L. N. and D. Bau, (1997). *Numerical linear algebra*. SIAM, Philadelphia.
- Turski R. B., E. E. Morris, T. A. Taiwo, J. E. Cahalan, (1997). Macroscopic cross section generation and application for coupled spatial kinetics and thermal hydraulics analysis with SAS-DIF3DK. *Proc. of the Joint Int. Conf. on Mathematical Methods and Supercomputing for Nuclear Applications, Saratoga Springs, New York, Oct. 5-9, 1997*, vol. 2, American Nuclear Society, 1997, pp. 1072-1081.