

УДК 004.89.032.26: 551.577

КОНЦЕПЦИЯ ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ ОЦЕНКИ НЕОПРЕДЕЛЕННОСТИ ПРИ ОБОСНОВАНИИ БЕЗОПАСНОСТИ ПУНКТОВ ЗАХОРОНЕНИЯ РАО

Савельева Е.А., к.ф.-м.н. (esav@ibrae.ac.ru)
(Институт проблем безопасного развития атомной энергетики РАН)

В работе рассматриваются вопросы разработки специального программного комплекса, направленного на учет и анализ неопределенностей, связанных с параметрическим наполнением расчетно-прогностических моделей, использующихся при обосновании безопасности пунктов захоронения радиоактивных отходов, включая особые РАО.

Основные задачи данного комплекса состоят в создании модельно-независимых средств для анализа чувствительности расчетных моделей к изменению их параметров и оценке неопределенности результатов расчетно-прогностического моделирования. Сформулированы основы формирования такого программного комплекса.

► **Ключевые слова:** особые РАО, захоронение РАО, обоснование безопасности, неопределенность, модельные расчеты.

THE CONCEPT OF SOFTWARE SYSTEM TO ASSESS UNCERTAINTIES AT SAFETY DEMONSTRATION OF RW DISPOSAL FACILITIES

SAVEL'EVA E., Ph. D. (esav@ibrae.ac.ru)
(Nuclear Safety Institute of the RAS (IBRAE))

This paper addresses the specialized software system development designed to account and analyse uncertainties related with the parametric content of calculation and predictive models applied in the course of safety demonstration of RW disposal facilities (including special RW).

This system basic purpose is to create the simulative-independent means for calculation models sensitivity analysis to their parameters changing and uncertainty assessment regarding calculation and predictive simulation results. Basis for such system creation is formulated.

► **Key words:** special RW, RW disposal, safety demonstration, uncertainty, model calculation.

Проблема обоснования безопасности объектов атомной отрасли на настоящем этапе ее развития приобрела особенно важное значение, так как она связана с выполнением обязательств Объединенной конвенции о безопасности обращения с отработавшим ядерным топливом и о безопасности обращения с радиоактивными отходами (РАО). В ходе работ в рамках Федеральной целевой программы «Обеспечение ядерной и радиационной безопасности на 2008 год и на период до 2015 года» был вскрыт целый пласт нерешенных проблем, связанных с объектами ядерного наследия (ОЯН) [1]. Для решения проблемы ОЯН разработан специальный системный подход [2], проанализирована возможность характеристики их состояния, например с использованием комплексного показателя риска [3], а также проведена их первичная регистрация, позволившая выделить отдельный класс особых РАО [4], включающих в себя пункты захоронения жидких РАО. Кроме того, в настоящее время ведутся работы по проблеме захоронения накопленных РАО, а именно выполняется проектирование и строительство подземной исследовательской лаборатории в Нижнеканском массиве с перспективой перевода его в пункт захоронения высокоактивных отходов.

Процедура обоснования безопасности объектов атомной отрасли не может обходиться без расчетно-прогностического моделирования, результат которого всегда сопряжен с неопределенностями, имеющими различные источники происхождения [5]. Необходимость анализа и учета неопределенностей также прописана в базовых документах МАГАТЭ, содержащих основные требования по нормам безопасности пунктов захоронения РАО. Появление неопределенностей связано как с невозможностью точно знать внешние условия существования объекта на отдаленный период времени (систематическая неопределенность), так и с неизбежными упрощениями модели реальных процессов, численными неточностями (модельная неопределенность) и ошибками измерений (неопределенность данных).

Разные типы неопределенности требуют различных подходов к своему анализу и учету. Неопределенности, связанные с невозможностью точного прогнозирования эволюции внешних факторов, воздействующих на систему захоронения, рассматриваются в рамках разработки сценариев, где главный акцент делается на полноту охвата возможностей.

Вопросы модельной неопределенности регулируются при выборе математической модели, разработке численных схем, реализации вычислительного алгоритма и верификации расчетного кода на тестовых задачах. При этом устанавливаются допустимые расхождения с аналитическими решениями, которые соответственно указываются в соответствующей документации, в частности, в аттестационных паспортах программных средств.

Гораздо меньше внимания при обосновании достоверности выполненного моделирования исторически уделяется учету неопределенностей данных, с которыми самым непосредственным образом связана неопределенность параметрического наполнения модели [5]. То есть результат расчетного моделирования должен сопровождаться адекватной оценкой неопределенности, вызванной неопределенностью в исходных данных. Выполнение такого рода анализа требует соответствующего программного инструментария. Данная статья посвящена концепции реализации программного комплекса для выполнения анализа неопределенности и чувствительности в контексте развития прогностического моделирования для обоснования безопасности пунктов захоронения РАО.

Прогнозное моделирование для обоснования безопасности пунктов захоронения РАО предполагает:

- моделирование совокупности всех явлений, происходящих с РАО в пределах контейнеров и определяющих характеристики источников излучения;
- моделирование, обеспечивающее оценку состояния и прогноз эволюции внешних инженерных барьеров безопасности, а также процессы миграции радионуклидов в пределах инженерных барьеров безопасности;
- моделирование переноса радионуклидов за пределами барьеров безопасности в геологической и водной среде, а также в атмосфере;
- моделирование воздействия ионизирующего излучения от пункта захоронения на человека и биоту.

Все перечисленные выше расчетные модели требуют параметризации – задания параметров модели, например упругих и прочностных свойств материалов, коэффициентов, характеризующих перенос (фильтрационные, сорбционные, метеорологические и т.п. поля), начальных значений, граничных условий и т.д. Качество параметризации модели, то есть насколько адекватно значения параметров, используемых в модели, соответствуют

реальным свойствам моделируемого объекта, во многом определяет достоверность получаемого результата. Но, с другой стороны, параметризация, как и сама модель, является своего рода упрощенным описанием реальности, что непременно ведет к неточности. Неточность результата моделирования в таких условиях может характеризоваться оценкой его неопределенности на основе учета неопределенности исходных данных и параметризации.

Таким образом, основная цель данной работы состоит в формировании концепции программного средства, ориентированного на обслуживание прогностических модельных расчетов в плане анализа их неопределенности к неточности и неполноте исходных данных.

Подробный обзор большого количества программных средств, направленных на работу с неопределенностями различного происхождения, приведен в [6]. Можно отметить, что программные средства делятся на две части: одна адаптирована под выполнение задач анализа чувствительности модели и оценки неопределенности результата моделирования для набора независимых скалярных параметров, другая предоставляет расширенные

возможности генерации случайных полей. Таким образом, ни одно из рассмотренных программных средств не может по отдельности являться полным аналогом проектируемого расчетного комплекса. Однако часть задач комплекса может быть реализована за счет внедрения готовых модулей, что является важным с точки зрения отработки реального взаимодействия с расчетными моделями, которые предполагается обслуживать.

1. Цель и задачи комплекса

Проектируемый программный комплекс предназначен для анализа чувствительности расчетных моделей к изменениям их параметров и оценки неопределенности результата расчетно-прогностического моделирования, выполняющегося при проведении процедуры по обоснованию безопасности ОЯН, пунктов захоронения РАО, в том числе особых. Он предназначен для обслуживания набора расчетно-прогностических моделей, обеспечивающих всестороннюю оценку всех аспектов существования ядерно- и радиационно опасных объектов (ЯРОО) (рис. 1).

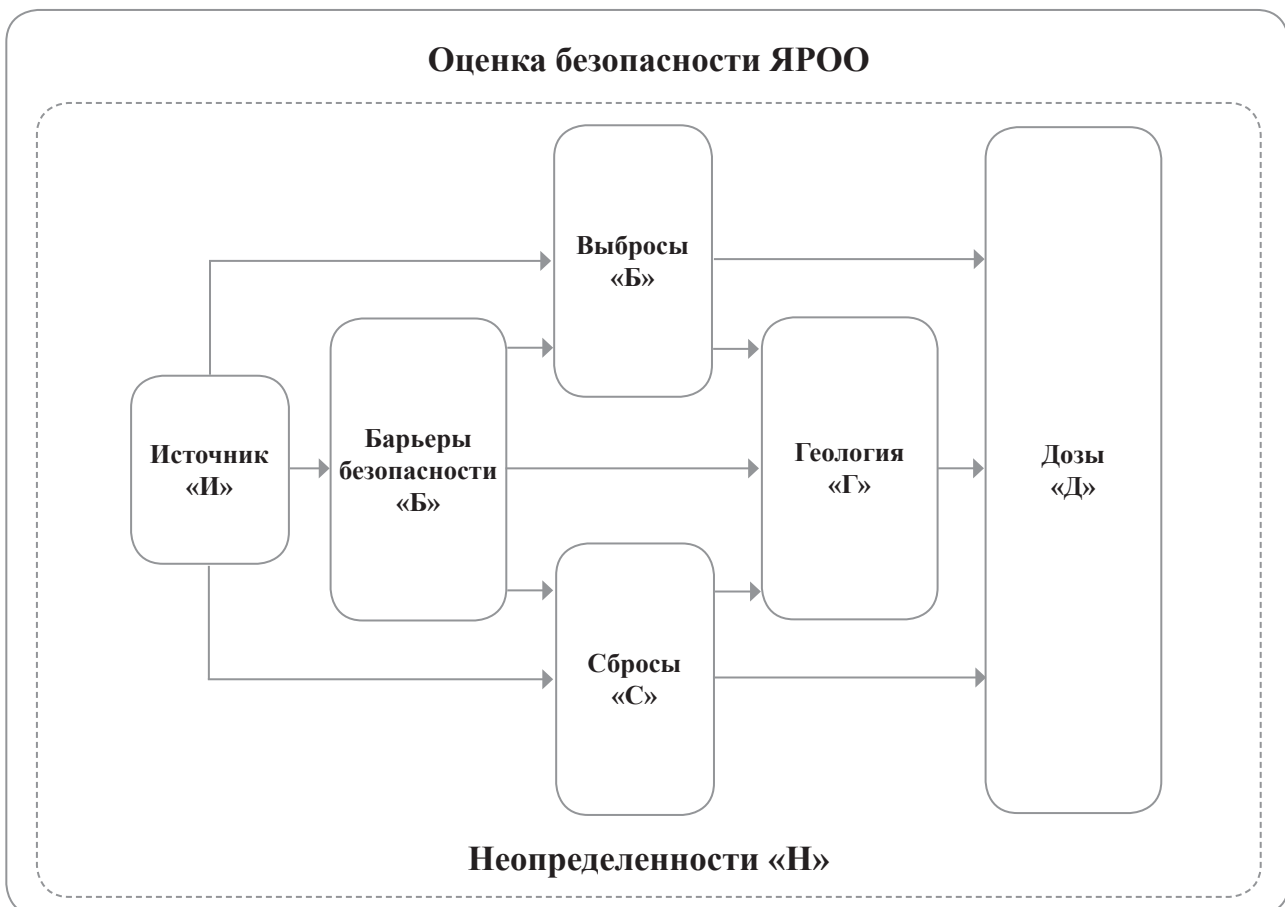


Рис. 1. Схема практической методологии обоснования безопасности ЯРОО

То есть фактически он является модельно-независимым комплексом, обслуживающим различные расчетные модели, а, следовательно, взаимодействие с расчетными моделями может осуществляться только посредством специальной системы управляющих файлов и файлов с входными параметрами.

Основные задачи комплекса следующие:

- осуществлять варьирование параметрического набора расчетной модели; для этого требуется набор стохастических генераторов, обеспечивающий все возможные варианты наборов входных параметров, а именно скалярных и категориальных переменных, временных рядов и пространственных полей;

- выполнять процедуры по анализу результатов многовариантных расчетов; сюда относятся реализация возможности оценивать и визуализировать чувствительности модели к ее неопределенным параметрам, а также выполнять комплексный статистический анализ для оценки неопределенности прогнозов.

Примерный список типов входных параметров, которые могут быть не точно известны и, соответственно, требуют варьирования для анализа чувствительности к их изменению и оценки неопределенности, вызванной ограниченностью знаний, может выглядеть следующим образом:

- скалярный параметр, который изменяется в заданном диапазоне значений или определяется известной функцией распределения;

- категориальный, то есть принимающий ограниченное число значений, параметр, который определяется вероятностью каждого возможного значения;

- скалярный или категориальный временной ряд, который определяется набором известных значений, или моделью корреляционной зависимости, или концептуальной моделью зависимости;

- скалярное или категориальное пространственное поле, для которого известны значения в точках наблюдений, или модель пространственной корреляционной структуры, или концептуальная модель зависимости;

- векторный временной ряд или векторное временное поле, для которого известны значения всего вектора или его части в точках наблюдений, или модели пространственной (временной) корреляционной структуры отдельных компонент вектора и модели пространственной (временной) кросс-корреляционной структуры нескольких компонент вектора, или концептуальная модель зависимости между компонентами вектора.

Таким образом, для решения основной задачи комплекса требуется разработка и реализация набора разнообразных случайных генераторов, начиная с примитивного генератора случайной переменной с известной функцией распределения до генератора многомерных коррелированных полей. Проверенные генераторы случайных чисел со стандартными функциями распределения реализованы практически во всех языках программирования в виде подпрограмм, так что детальная разработка такого рода алгоритмов не требуется, только если некоторый параметр имеет сложную нестандартную форму функции распределения. При генерации многомерной случайной переменной требуется знание о наличии корреляции между ее компонентами. Если корреляция между переменными отсутствует, то каждая из них может генерироваться независимо с использованием собственного генератора. Для оценки корреляции необходимо наличие такого количества наборов значений переменной, которое дает возможность получать достоверные статистические оценки, в противном случае модель корреляции может быть задана только эмпирически. Также для генерации временных рядов и пространственных полей, которые заданы набором известных значений, необходимо иметь возможность оценивать наличие пространственной (временной) корреляции и моделировать ее. При отсутствии зависимости поле может генерироваться как набор независимых случайных переменных, поэтому программный комплекс должен иметь набор модулей для оценки и моделирования корреляции в данных.

Неопределенность результата прогностического расчета связана не только с неточностью значений параметров, но и с их концептуальной моделью [5]. Концептуальные модели одного и того же процесса могут различаться структурной моделью параметров, например за счет различий в геологической модели среды (размещение разлома, разные варианты толщ осадочных пород), или подходом к параметризации, например размерностью моделирования, подходом к заданию свойств (трещины как каналы или модель двойной пористости). Для сравнения различных концептуальных моделей используются средства автоматической калибровки параметров. Калибровка параметров модели состоит в минимизации целевой функции, определяющей качество соответствия результатов расчета и реальных измерений некоторой характеристике. Таким образом, для реализации возможности оценки концептуальной неопределенности,

связанной с неоднозначностью способа моделирования процесса, в комплекс должны включаться оптимизационные модули.

Можно сформулировать, что для решения основных задач программный комплекс по анализу чувствительности и оценке неопределенности должен включать в себя модули для:

- анализа чувствительности расчетной модели к изменению ее параметров различного типа;
- оценки неопределенности исходных данных и результатов многовариантного моделирования;
- генерации наборов входных параметров расчетных моделей;
- оценки и моделирования корреляции по набору исходных измерений;
- минимизации функций, в том числе со сложной структурой и обладающих большим числом локальных минимумов.

2. Технологическая платформа

Так как проектируемый программный комплекс будет использоваться для анализа чувствительности и оценки неопределенности внешних расчетно-прогностических моделей, основным требованием для него является совместимость с этими моделями. В свою очередь, базовой операционной системой для большинства программных средств расчетно-прогностического моделирования является Microsoft Windows, а в силу потенциальной потребности в высокопроизводительных параллельных вычислениях на кластерных системах также обязательна поддержка ОС Linux. В связи с этим важным требованием является кроссплатформенность – прототип программного комплекса должен разрабатываться для запуска как в операционных системах семейства Microsoft Windows, так и в операционных системах Linux.

Этим требованиям удовлетворяет язык программирования Python, который является системно-независимым языком: разработанные на нем программы могут быть легко перенесены с одной операционной системы на другую. Кроме того, язык Python позволяет включать готовые расчетные модули, написанные на других языках, таких как FORTRAN, C++, Java, OCaml, Prolog, а также при необходимости работать с различными системами управления базами данных (СУБД).

В процессе реализации комплекса будут востребованы такие библиотеки, как:

- NumPy – библиотека, поддерживающая высокоуровневые математические функции, предназначенные для работы с многомерными массивами;
- SciPy – библиотека, предназначенная для выполнения научных и инженерных расчетов, в которой, в том числе, реализованы: поиск минимумов и максимумов функций, вычисление интегралов функций, поддержка специальных функций, обработка сигналов, обработка изображений, работа с генетическими алгоритмами, решение обыкновенных дифференциальных уравнений и др.;
- Matplotlib – библиотека для визуализации результатов вычислений, поддерживает такие типы графиков, как линейные графики (line plot), диаграммы разброса (scatter plot), столбчатые диаграммы (bar chart) и гистограммы (histogram), круговые диаграммы (pie chart), ствол-лист диаграммы (stem plot), контурные графики (contour plot), поля градиентов (quiver), спектральные диаграммы (spectrogram) и другие. Полученные изображения могут быть использованы в интерфейсе или сохранены в распространенных форматах (.eps, .emf, .jpeg, .pdf, .png, .svg, .tiff и других);
- python-docx, xlrd, xlwt, Xlsxwriter – библиотеки для ввода-вывода в файлы форматов MS Office;
- pandas – библиотека, предоставляющая специальные структуры данных и операции для манипулирования числовыми таблицами и временными рядами;
- Mserp, runc – библиотеки, реализующие стохастические выборки;
- SQLAlchemy – библиотека для работы с реляционными СУБД.

3. Структура организации программного комплекса

Для решения каждой из перечисленных в разделе 1 задач могут быть предложены различные алгоритмы. Каждый конкретный алгоритм задает предписания, основанные на определенном методе решения задачи, который также имеет теоретическую основу, математическую формализацию исходной задачи в виде математических понятий и формул, набор предположений (ограничений), связанных с теоретической основой и т.п.

Реализация алгоритма в виде программного кода представляет собой программный модуль. Но при этом в различных алгоритмах могут присутствовать идентичные блоки операций (процедуры), например генерация случайного числа с заданной

функцией распределения или вычисление статистических характеристик (среднего, вариации и т.п.), или поиск ближайших соседей в соответствии с заданной метрикой пространства, или другие. Программные коды таких процедур будут формироваться в библиотеку процедур, доступную всем модулям на этапе их сборки.

Иллюстрация иерархической структуры перехода от основной цели к конкретным модулям представлена на рис. 2. Модульная структура является очень гибкой и удобной, она позволяет добавлять новые модули, реализующие улучшенные версии старых алгоритмов, или новые, разработанные на

считывания и записи файлов (рис. 3). Вся входная информация для работы модуля будет сосредоточена в системе файлов, в которых задаются управляющие параметры расчетных модулей и исходные данные (измерения, результаты расчетов и т.п.). Файлы с управляющими параметрами можно будет также готовить через графический интерфейс комплекса. Файлы с исходными данными будут передаваться извне так же, как и другая информация, необходимая для работы с внешними расчетными моделями. Результаты расчетов будут собираться в файлах, возможна реализация сохранения в базе данных.

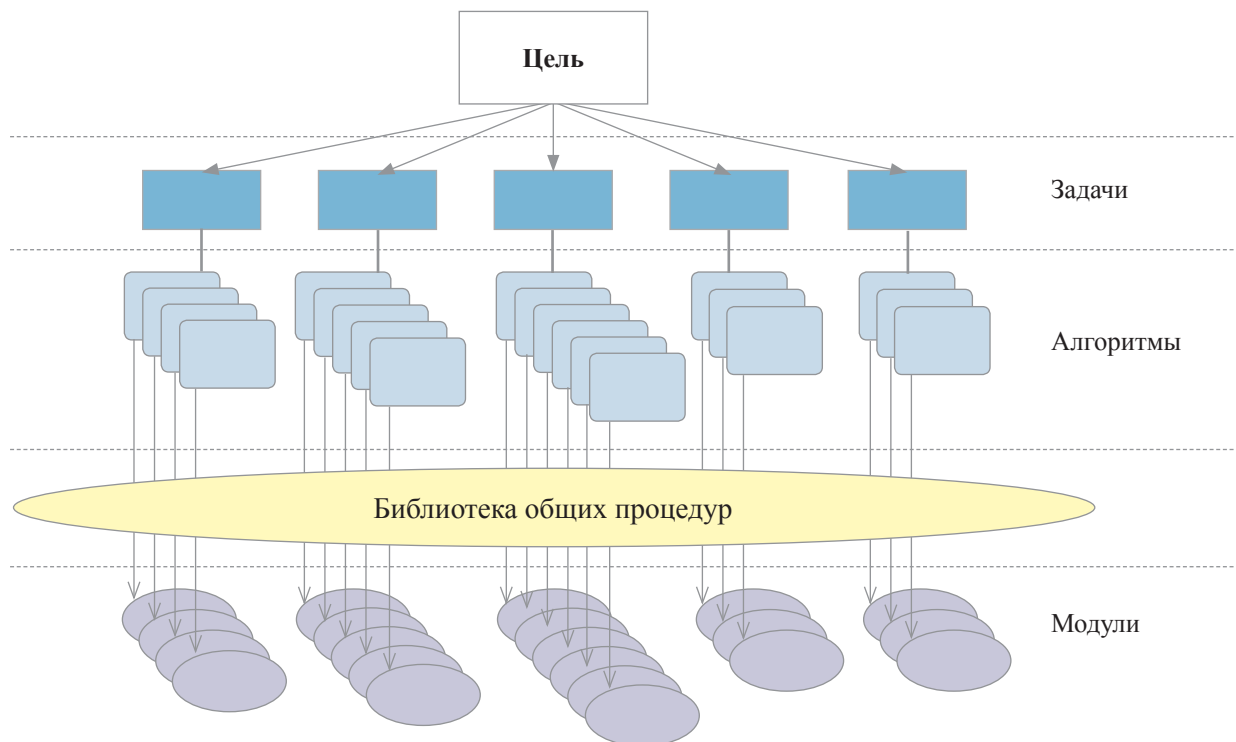


Рис. 2. Иерархическая структура формирования программных модулей для осуществления основной цели комплекса

основе опыта эксплуатации программного комплекса, алгоритмы. В процессе развития может пополняться и библиотека общих процедур.

Программный комплекс может быть организован как система независимых модулей, каждый из которых реализует некоторый метод решения конкретной задачи, связанной с оценкой чувствительности расчетных моделей к изменению их параметров и анализом неопределенности расчетно-прогностического моделирования. Модули можно для удобства формального описания архитектуры объединять в блоки в соответствии с классом решаемых задач (рис. 3).

Взаимодействие между модулями может осуществляться управляющим модулем посредством

Форматы файлов для передачи результатов расчетов внутри комплекса удобно организовать по формату данных GeoEAS [7], который характерен для наиболее распространенных геостатистических приложений. Традиционно это текстовый файл, содержащий только цифровую информацию и имеющий простую и удобную для произвольно распределенных пространственно-временных данных организацию: данные записываются в колонки, а над колонками формируется описание данных, позволяющее в дальнейшем организовывать действия с определенными колонками цифр. Файл такой структуры легко конвертируется в формат таблиц программы Excel, а также может записываться в виде цифровой информации в базы данных.

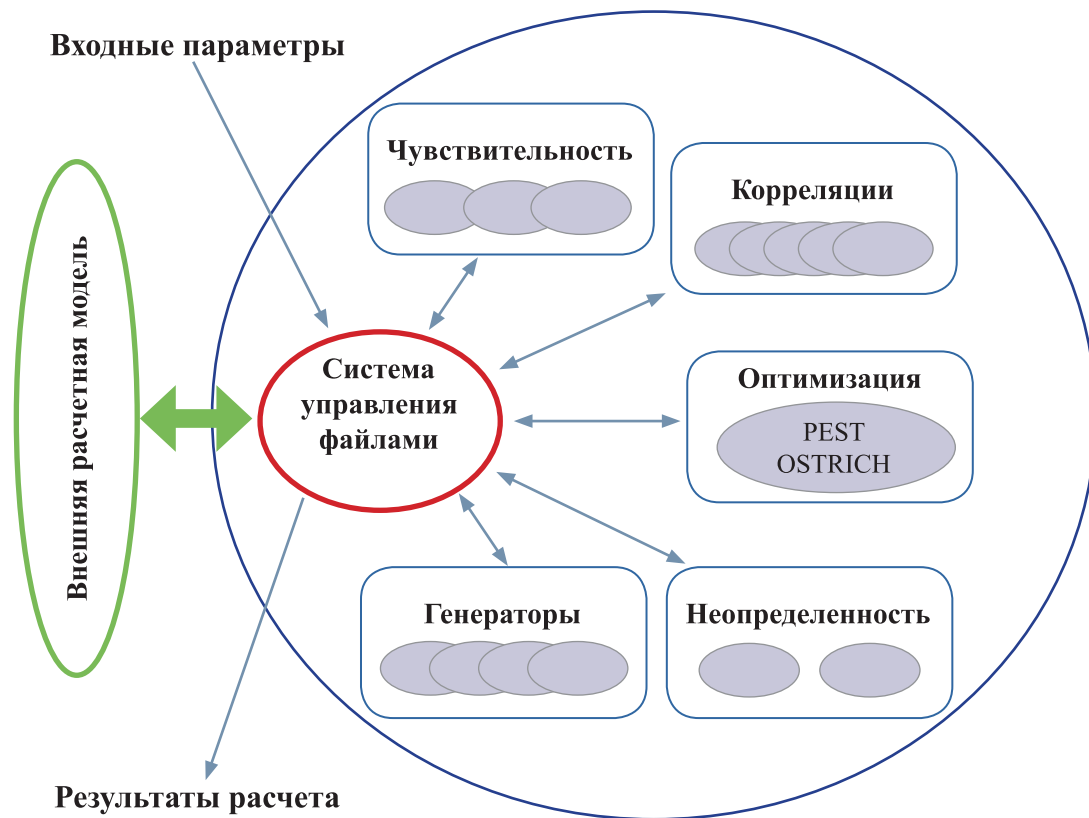


Рис. 3. Архитектура программного комплекса

Основная концепция взаимодействия с внешними расчетно-прогностическими моделями состоит в том, что совершенно неважно, что конкретно с помощью данной модели оценивается; важно, чтобы ее можно было запускать в режиме командной строки, а все параметры, необходимые для ее функционирования, считывались из файла. В состав входной информации для комплекса должны включаться указания на конкретные параметры, требующие варьирования в рамках проведения анализа чувствительности или оценки неопределенности, и исходные данные для генератора. В рамках управления функционированием комплекса в файловом пространстве должен быть реализован специальный управляющий модуль для формирования на основе полученной информации цепочек проведения расчетов внутри комплекса, файлов с входными параметрами для внутренних и внешних расчетных моделей, файлов с результатами расчетов (рис. 3).

Формализация процесса взаимодействия будет аналогична системе модельно-независимого программного пакета PEST [8], предназначенного для калибровки параметров модели (рис. 4). Для реализации серии расчетов с варьированием параметров будут требоваться входные файлы трех типов:

- шаблоны, которые задают структуру входного файла параметров для данной расчетной модели (в шаблонах также обозначаются те параме-

тры, которые подлежат варьированию);

- шаблоны, определяющие структуру выходного файла расчетной модели с указанием характеристик, для которых выполняется анализ чувствительности или оценка неопределенности;

- контрольный файл, который объединяет все вместе, то есть содержит информацию о том, какие шаблоны относятся к какой расчетной модели (если запускается последовательность расчетных моделей), задает правила варьирования параметров (диапазоны или функции распределения, модели корреляции) и их типы, определяет решаемую задачу, данные, используемые для калибровки, и прочее.

4. Состав базовой версии комплекса

При формировании набора модулей для базовой версии комплекса следует принимать во внимание, что он:

- должен покрывать полный спектр основных и дополнительных задач, которые были сформулированы в разделе 1;

- учитывать международный опыт и современные тенденции в использовании алгоритмов для решения задач, аналогичных поставленным;

- использовать алгоритмы, адекватность которых обосновывается фундаментальной математической теорией.

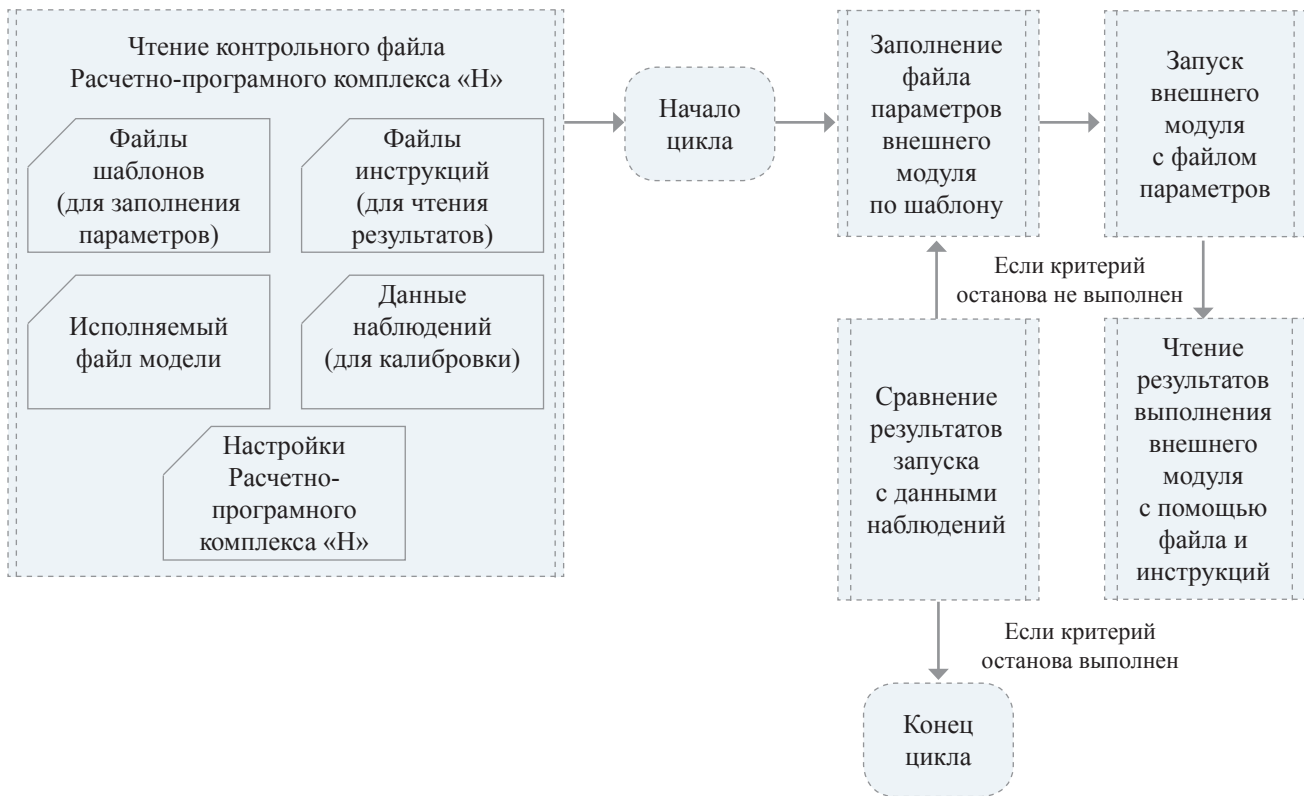


Рис. 4. Цикл запуска калибровки внешнего модуля

4.1. Модули анализа чувствительности

Модуль оценки локальной чувствительности модели к изменению параметра предназначен для оценки зависимости результата модельного расчета от изменения параметра. При этом зависимость рассматривается на момент завершения расчета или на некоторый другой, важный для модели, момент времени.

Оценка локальной чувствительности осуществляется по результатам многократного запуска программного кода, реализующего модель. Модельные расчеты выполняются для основного набора значений параметров, а также для серии наборов параметров, где каждый получен посредством работы генератора равномерного распределения в заданном для него диапазоне значений. Использование именно равномерного генератора объясняется тем, что в данном случае нас интересует не реальное поведение конкретного параметра, а связь между изменением его значения и изменением результата модели. Вычисляется локальная чувствительность по формуле:

$$S_{ij} = \frac{\partial y_i}{\partial x_j} = \frac{1}{N} \sum_{k=1}^N \frac{(\Delta y_i)_k}{(\Delta x_j)_k}, \quad (1)$$

где:

N – число выполненных расчетов;

k – номер конкретного расчета;

Δx_j – разница между основным значением параметра x_j и его k -ой реализацией;

Δy_i – разница в расчете результата y_i с основным набором параметров и их k -ой реализацией.

Серию оценок локальной чувствительности, выполненных для различных локальных точек для набора переменных, удобно анализировать используя подход ранговой корреляции, ориентированной на выявление синхронности в проявлении роста значений. Такой анализ позволит выявить наличие корреляций в тенденциях изменения локальной чувствительности для различных переменных. Для выполнения такого анализа полезно включить модули для оценки одного или нескольких коэффициентов ранговой корреляции из серии наиболее известных, таких как критерии Кендалла, Спирмена, Гедфинга, Ширахатэ и другие.

Визуализация результатов анализа локальной чувствительности может быть выполнена средствами графической библиотеки Matplotlib.

Локальная чувствительность оценивается для скалярных параметров модели, для которых может быть вычислено значение изменения. Для того, чтобы использовать локальный анализ чувствительности в контексте полей параметров требуется ввести меру, характеризующую изменение поля от реализации к реализации.

Глобальная чувствительность позволяет определить, как погрешность результата модельного расчета распределилась между входными параметрами. Она может характеризоваться различными индексами. Наиболее простым индексом глобальной чувствительности является нормализованный коэффициент регрессии, который может быть использован в предположении о линейной зависимости между параметрами модели и выходным результатом:

$$Y = \beta_0 + \sum_{i=1}^n \beta_i X_i, \quad (2)$$

где:

$X = (x_1, \dots, x_n)$ – входные параметры;

$Y = (y_1, \dots, y_m)$ – выходные параметры, являющиеся результатом расчета.

Модуль для расчета нормализованного коэффициента регрессии предназначен для вычисления соответствующего индекса:

$$SRC_i = \frac{\beta_i^2 \sigma(x_i)}{\sum_{i=1}^n \beta_i^2 \sigma(x_i)}, \quad (3)$$

где $\sigma(x_i)$ – вариация параметров x_i .

Дополнительной опцией этого модуля является возможность выполнения проверки корректности гипотезы о линейном приближении (2).

Модуль для расчета индекса чувствительности Соболя предназначен для вычисления индекса глобальной чувствительности, который корректно использовать для любых, в том числе нелинейных, моделей. Теоретическая основа этого подхода состоит в разложении функции f на разноразмерные слагаемые, которое удобнее выписать, используя сумму по различным группам индексов (i_1, \dots, i_s , где $1 \leq i_1 < \dots < i_s \leq n$ и $s = 1, \dots, n$) [9]:

$$f(x) = f_0 + \sum f_{i_1 \dots i_s}(x_{i_1}, \dots, x_{i_s}) \quad (4)$$

и оценке соответствующих величин дисперсий:

$$D = \int f^2(x) dx - f_0^2, \dots, \\ D_{i_1 \dots i_s} = \int \dots \int f_{i_1 \dots i_s}^2 dx_{i_1} \dots dx_{i_s}. \quad (5)$$

На основе дисперсий (5) вычисляется индекс глобальной чувствительности Соболя:

$$S_{i_1 \dots i_s} = \frac{D_{i_1 \dots i_s}}{D}. \quad (6)$$

Для вычисления интегралов (5) может использоваться метод Монте-Карло [9], который требует значительного числа запусков модели. Для уменьшения вычислительной нагрузки при оценке индексов Соболя можно использовать версию с аппроксимацией модели на основе расширенного полиномиального хаоса [9] как специальную опцию данного модуля. В этой версии на основе результатов расчета модели будет реализовано построение ее аппроксимации:

$$f(X) \approx f_{PC}(X) = \sum_{i=1}^{p-1} F_i \Psi_i(X), \quad (7)$$

где:

X – вектор, содержащий все компоненты входного вектора, преобразованные к интервалу $[-1, 1]$;

p – ограничение бесконечного суммирования, которое в теории дает точную аппроксимацию;

F_i – коэффициенты разложения, а $\Psi_i(X)$ – полиномы Лежандра.

Вычисление коэффициентов разложения (7) может быть осуществлено, например, посредством минимизации среднеквадратичного отклонения [9]. Аппроксимация индекса чувствительности Соболя может быть оценена как:

$$SU_{i_1 \dots i_s} = \frac{\sum_{\alpha \in L_{i_1 \dots i_s}} F_{\alpha}^2 E[\Psi_{\alpha}^2]}{D_{PC}}. \quad (8)$$

4.2. Модули анализа и моделирования корреляционной структуры

Как уже отмечалось выше, для генерации случайных реализаций временных рядов или пространственных полей параметров задание модели корреляционной зависимости, то есть модели связи между значениями и их взаимным пространственным (временным) расположением, является необходимым. Если модель пространственной корреляции не известна, но имеется набор значений измерений, то она при определенных предположениях [10] может быть построена.

Для описания корреляционной структуры будут использованы следующие характеристики:

- корреляционная матрица, описывающая степень зависимости между компонентами многомерной переменной (вектора);
- пространственная ковариация для полей и авто-ковариация для временного ряда, вычисляемая по формуле:

$$C(h) = \frac{1}{N_h} \sum (Z(x_i) - \bar{Z})(Z(x_j) - \bar{Z}), \quad (9)$$

где суммирование ведется по всем x_i и x_j , разделенным вектором h ($h = x_i - x_j$);

$Z(x)$ – значение функции в точке с координатой x .

\bar{Z} – локальное среднее, которое в силу условия стационарности не зависит от времени или точки в пространстве, а постоянно для всего поля или ряда;

N_h – количество пар точек, разделенных вектором h ;

- спектральная мощность временного ряда, которая позволяет выделить для него наиболее значимые периоды;

- вариограмма, двухточечный момент второго порядка, характеризующий пространственную коррелированность данных в предположении о стационарности приращений (внутренней гипотезе) [10].

Для использования в генераторе не всегда достаточно оценить поле ковариации или вариограммы. В методах генерации случайных полей на основе геостатистической теории требуется модель пространственной корреляции (вариограммы). Поэтому в комплекс включается модуль, предназначенный для подбора параметров модели вариограммы и кросс-вариограммы. Модуль работает в предположении, что пространственная корреляция либо не обладает анизотропией, либо анизотропия может быть сведена к геометрической [10]. Кроме того, не делается подбор модели вариограммы по поведению вблизи нуля и на бесконечности, а используется классическая сферическая модель [10]:

$$\gamma(h) = \begin{cases} c_0 + c \left[\frac{3h}{2a} - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right] & \text{для } h \leq a \\ c_0 + c & \text{для } h > a \end{cases}, \quad (10)$$

где:

a – радиус корреляции, который характеризует расстояние, после которого данные уже не коррелированы;

c_0 – эффект самородка или кратко самородок, который характеризует значение вариограммы в малой окрестности нуля;

c – плато, которое характеризует значение вариограммы на бесконечности минус значение самородка.

Геометрическая анизотропия характеризуется тем, что изолинии вариограммы имеют форму

эллипсов (в случае 2D) или эллипсоидов (в случае 3D), которые и являются зоной корреляции. В этом случае существует положительно определенная матрица B , такая, что после преобразования пространства ($h' = h^T B h$) вариограмма становится изотропной.

Процедура подбора параметров сводится к калибровке модели на основе целевой функции [10]:

$$I = \sum_{i=1}^k \frac{N(h_i)}{h_i^2} \left[\gamma(h_i, \lambda) - \gamma(h_i) \right]^2, \quad (11)$$

где:

$N(h_i)$ – количество пар оцененной по данным вариограммы (экспериментальной) для вектора h_i ;

$\gamma(h_i)$ – значение экспериментальной вариограммы для вектора h_i ;

$\gamma(h_i, \lambda)$ – значение модели вариограммы для вектора h_i и набора параметров модели λ (c_0, c, a).

Процедура подбора параметров модели вариограммы является задачей оптимизации. Для ее автоматического решения предполагается использовать внешнюю программу OSTRICH [11], информация о которой более подробно изложена в подразделе 4.4.

4.3. Модули генераторов

Количество и состав генераторов в комплексе несомненно будет меняться по мере развития всей методологии расчетно-прогностического моделирования для обоснования безопасности ОЯН и объектов захоронения РАО. Это связано с тем, что по мере развития расчетных моделей и углубления в их неопределенности несомненно будут появляться заявки на новые типы генераторов. На данном этапе можно сформулировать стандартный набор модулей генерации случайных полей размерности 1D, 2D и 3D.

Можно выделить основные классы генераторов случайных полей.

- Во-первых, генераторы могут быть условными и безусловными. Условный генератор определяется набором исходных данных, которые должны быть воспроизведены в каждой реализации. Безусловный генератор не опирается на исходные данные. Он воспроизводит только заданные функционалы, например статистические моменты первого и второго порядка (среднее, вариация, корреляционная структура), пропорциональные отношения объектов различных классов, эмпирический тренировочный образ и т.п.

▪ Во-вторых, существуют пиксельные и объектные генераторы. Пиксельные генераторы базируются на распределении точечных значений функции по пикселям (ячейкам). Это реализуется посредством заполнения значениями функции узлов ячеек сетки, которые ассоциируются с пикселями. Объектные генераторы базируются на случайном пространственном распределении объектов заданных геометрических форм и размеров.

▪ В-третьих, можно отдельно рассмотреть генераторы непрерывных и категориальных переменных.

В базовом наборе модулей генераторов случайных полей будут представлены условные и безусловные генераторы на основе пиксельного подхода, которые хорошо зарекомендовали себя при решении различных задач, связанных с оценкой неопределенности. Среди них:

▪ модуль безусловной генерации случайного Гауссова поля или временного ряда с заданной ковариационной структурой;

▪ условный генератор временного ряда с заданными свойствами тренда и периодичности [12];

▪ модуль генерации условного случайного поля на основе набора произвольно распределенных по области N известных значений $(Z(x_i), i = 1, \dots, N)$ с подготовленной заранее моделью пространственной корреляционной структуры (вариограммы – $\gamma(h)$) [13];

▪ модуль условной генерации случайного поля категориальной переменной на основе индикаторного подхода [10].

4.4. Модули оптимизации

Модули оптимизации в программном комплексе предназначены, главным образом, для калибровки параметров моделей, но также на основе стохастических методов оптимизации могут строиться генераторы случайных полей.

Традиционно при автоматической калибровке используются детерминистические градиентные методы, но в работе [14] на двух реальных примерах моделей нефтяных месторождений было показано, что различные градиентные методы не обеспечивают требуемого уровня решения оптимизационной задачи для сложной целевой функции, имеющей множество локальных минимумов. В связи с этим требуется опробовать и использовать более современные подходы.

На настоящий момент разработано настолько большое количество самых разнообразных методов оптимизации – детерминистических, эвристических и гибридных, что выбор тех, которые были бы одновременно пригодны и для задач калибровки расчетных моделей, и для целей генерации случайных полей затруднен. Поэтому на первом этапе разработки комплекса для целей оптимизации будет использована свободно распространяющаяся программа OSTRICH [11].

В программе OSTRICH реализован широкий спектр детерминистических, эвристических и гибридных оптимизационных подходов [11]. Среди детерминистических методов можно выделить классические методы (метод бисекции, алгоритм Флетчера-Ривза, градиентные методы – метод наискорейшего спуска, различные модификации алгоритма Левенберга-Марквардта и алгоритм Пауэлла), а также более современный подход на основе сеточного исчерпывающего поиска. Группа эвристических методов представлена спектром подходов на базе динамически заданного поиска (Dynamically Dimensioned Search): бинарно и вещественно кодированные версии генетических алгоритмов, имитация отжига, эволюционный алгоритм, несколько модификаций роевого подхода, выборочный алгоритм на основе «Большого взрыва». Кроме того, имеется возможность использовать гибридный алгоритм, построенный на основе оптимизации роем частиц с дополнительным уточнением методом Левенберга-Марквардта.

Столь широкий и разнообразный спектр методов должен позволить покрыть текущие проблемы по задачам оптимизации. Кроме того, в процессе апробирования комплекса будут выявлены недостатки имеющихся методов и сформулированы требования, основанные на реальных потребностях и особенностях внешних расчетных моделей.

4.5. Модули оценки неопределенности

Оценка неопределенности результатов моделирования должна выполняться на основе использования самых современных подходов к методам стохастического моделирования. Для анализа результатов многовариантных расчетов в программном комплексе должны быть доступны алгоритмы, реализующие следующие подходы к оценке статистических характеристик: среднее значение, вариация, симметричность, эксцесс, медиана, квантили $1/4$ и $3/4$, доверительные интервалы, вероятности

быть выше (или ниже) заданного уровня и другие. Это позволит обеспечить максимально широкий круг характеристик неопределенности результата и, соответственно, повысит уровень аргументации на основе результатов расчетно-прогностического моделирования. Также должна быть реализована подготовка для визуализации пространственного результата в виде «толстых изолиний», что даст дополнительные возможности визуального представления пространственных зон, где результат оценки расчетной характеристики обладает более высокой неопределенностью.

Другим аспектом данного блока модулей является задача оценки неопределенности, заложенной непосредственно в исходных данных. Используемый подход основан на оценке вариации стохастической составляющей (r) при аппроксимации исходных данных гладкой функцией (f):

$$y = f(X) + r, \quad (12)$$

где:

вектор $X = (x_1, \dots, x_n)$ является входными данными для расчета;

y – выходной результат.

Для оценки $Var(r)$ используется метод Гамма-теста [15], который гарантирует, что $Var(r)$ определяется исключительно неопределенностью, заложенной в исходных данных, и никак не зависит от параметров аппроксимирующей модели. Такая неопределенность очень важна, так как позволяет адекватно оценивать возможности параметризации модели на основе этих данных. Качество калибровки модели и, соответственно, неопределенность результата моделирования во многом определяются используемыми для параметризации исходными данными.

Заключение

В данной работе представлена концепция программного комплекса для анализа чувствительности и оценки неопределенности расчетно-прогностического моделирования при обосновании безопасности ОЯН и пунктов захоронения РАО, основанная на детальном анализе его целей. Сформированная концепция может быть представлена в виде тезисов:

- Базовым языком для разработки программного комплекса будет язык программирования Python, обладающий в числе прочего и графическими средствами для визуализации результатов.

- Программный комплекс будет представлять собой систему независимых модулей, каждый из которых является программной реализацией определенного алгоритма, направленного на решение основной или дополнительной задач комплекса. Для взаимодействия между расчетными модулями внутри программного комплекса и с внешними расчетными модулями будет использовано специальное файловое пространство со средством управления.

- Разработаны алгоритмы базовой версии программного комплекса и сформирован список модулей, включающий 3 модуля для анализа чувствительности, 5 модулей для оценки и моделирования корреляционных свойств исходных данных, 4 модуля генерации пространственных полей различной размерности, 2 модуля для анализа неопределенности исходных данных и результатов расчетного моделирования. В качестве модулей для оптимизации при выполнении калибровки параметров моделей будут использованы алгоритмы, реализованные в программах PEST [8] и OSTRICH [11] (более 10 методов).

Обновление состава модулей и процедур, а также модификация, адаптация и новые подходы для их реализации будут уточняться на этапе практического опытного использования программного комплекса.

Список литературы

1. Большов Л.А., Крюков О.В., Лаверов Н.П., Линге И.И. Проблемы ядерного наследия и пути их решения. Развитие системы обращения с радиоактивными отходами в России. – М., ОАО «Энегпроманалитика», Т. 2, 2013.
2. Абрамов А.А., Дорофеев А.Н., Комаров Е.А., Кудрявцев Е.Г., Большов Л.А., Линге И.И., Абалкина И.Л., Бирюков Д.В., Ведерникова М.В., Хамаза А.А., Шарафутдинов Р.Б., Бочкарев В.В. К вопросу оценки объема ядерного наследия в атомной промышленности и на иных объектах мирного использования атомной энергии в России. Ядерная и радиационная безопасность, № 3 (73), 2014.

3. Хамаза А.А., Бочкарев В.В., Курьиндин А.В., Абакумова А.С., Тихомиров Г.В. О возможности применения комплексного показателя потенциальной опасности в качестве основы дифференцированного подхода к регулированию безопасности объектов «ядерного наследия». *Ядерная и радиационная безопасность*, № 1 (79), 2016.
4. Абрамов А.А. и др. Особые радиоактивные отходы. Под общ. ред. Линге И.И. – М., ООО «САМ полиграфист», 2015.
5. Caers J. *Modeling Uncertainty in the Earth Sciences*. – New Delhi: Wiley-Blackwell, 2011.
6. Matott L.S., Babendreier J.E., Purucker S.T. Evaluating uncertainty in integrated environmental models: a review of concepts and tools. *Water Resources Research*, Vol. 45, № 6, 2009.
7. Englund E., Sparks A. *Geo-EAS 1.2.1. Geostatistical Environmental Assessment Software, User's Guide*. – Las Vegas: U.S. EPA, 1991.
8. Doherty J. *PEST – Model-independent parameter estimation. Version 12*. J. Doherty. *Watermark Computing*. – Australia, 2010. <http://www.pesthomepage.org>.
9. Sudret B. Global sensitivity analysis using polynomial chaos expansion. B. Sudret. *Reliability Engineering and System Safety*, V. 93, 2008.
10. Демьянов В.В., Савельева Е.А. *Геостатистика. Теория и практика*. – М., Наука, 2010.
11. Matott L.S. *OSTRICH: An optimization software tool: documentation and user's guide*. – Buffalo, NY: University at Buffalo, 2005.
12. Савельева Е.А., Уткин С.С. Прогноз изменения уровня воды в водоеме В-11 ТКВ на основе методов геостатистического моделирования. *Вопросы радиационной безопасности*, № 1, 2012.
13. Гончаров С.М., Попов С.Б., Поляков А.В., Савельева Е.А., Mazzola S., Bonano A., Patti B. Определение точности оценки запаса гидробионтов при геостатистической интерполяции исходных данных. *Рыбное хозяйство*, № 4, 2010.
14. Bissel R., Sharma Y., Kilongh J. History matching using the method of gradients. Two case studies. In *Proceedings of the SPE Annual Technical Conference and Exhibition*. – New Orleans, Louisiana, USA, SPE 28590, 1994.
15. A. Stefansson, N. Koncar, A.J. Jones. Note on the Gamma Test. *Neural Computing and Applications*, V. 5, 1997.

References

1. Bolshov L.A., Kryukov O.V., Laverov N.P., Linge I.I. Solving Nuclear Legacy Challenges. Development of Radioactive Waste Management System in Russia. – Moscow, JSC “Energopromanalitica”, Vol. 2, 2013.
2. Abramov A.A., Dorofeyev A.N., Komarov E.A., Kudryavtsev E.G., Bolshov L.A., Linge I.I., Abalkina I.L., Biryukov D.V., Vedernikova M.V., Khamaza A.A., Charafoutdinov R.B., Bochkarev V.V. On the issue of evaluation of nuclear legacy volumes in the nuclear industry and other facilities of peaceful use of nuclear energy in Russia. *Nuclear and Radiation Safety*, No. 3 (73), 2014.
3. Khamaza A.A., Bochkarev V.V., Kuryndin A.V., Abakumova A.S., Tikhomirov G.V. Concerning the possibility to use the complex factor of potential hazard as the basis of graded approach to regulation of safety of nuclear legacy facilities. *Nuclear and Radiation Safety*, No. 1 (79), 2016.
4. Abramov A.A. and others. *Special Radioactive Wastes*. Under the general editorship of Linge I.I. – Moscow, LLC “SAM printer”, 2015.
5. Caers J. *Modeling Uncertainty in the Earth Sciences*. – New Delhi: Wiley-Blackwell, 2011.
6. Matott L.S., Babendreier J.E., Purucker S.T. Evaluating uncertainty in integrated environmental models: a review of concepts and tools. *Water Resources Research*, Vol. 45, No. 6, 2009.
7. Englund E., Sparks A. *Geo-EAS 1.2.1. Geostatistical Environmental Assessment Software, User's Guide*. – Las Vegas: U.S. EPA, 1991.
8. Doherty J. *PEST – Model-independent parameter estimation. Version 12*. J. Doherty. *Watermark Computing*. – Australia, 2010. <http://www.pesthomepage.org>.
9. Sudret B. Global sensitivity analysis using polynomial chaos expansion. B. Sudret. *Reliability Engineering and System Safety*, V. 93, 2008.

10. Demyanov V.V., Savelyeva E.A. Geostatistics. Theory and Practice. – Moscow, Science, 2010.

11. Matott L.S. OSTRICH: An optimization software tool: documentation and user's guide. – Buffalo, NY: University at Buffalo, 2005.

12. Savelyeva E.A., Utkin S.S. Forecast for change of the water level in the reservoir V-11 of the Techa Cascade of water reservoirs based on the methods of geostatistical modeling. Radiation Safety Issues, No. 1, 2012.

13. Goncharov S.M., Popov S.B., Polyakov A.V., Savelyeva E.A., Mazzola S., Bonano A., Patti B. Investigation of accuracy of estimated reserves of hydrobionts in the course of geostatistical interpolation of input data. Fish farm, No. 4, 2010.

14. Bissel R., Sharma Y., Kilongh J. History matching using the method of gradients. Two case studies. In Proceedings of the SPE Annual Technical Conference and Exhibition. – New Orleans, Louisiana, USA, SPE 28590, 1994.

15. A. Stefansson, N. Koncar, A.J. Jones. Note on the Gamma Test. Neural Computing and Applications, V. 5, 1997.

