

На правах рукописи

Кузьмин Вадим Александрович

**ФУНДАМЕНТАЛЬНЫЕ ОСНОВЫ И МЕТОДОЛОГИЯ
АВТОМАТИЧЕСКОЙ КАЛИБРОВКИ МНОГОПАРАМЕТРИЧЕСКИХ
ГИДРОЛОГИЧЕСКИХ МОДЕЛЕЙ**

Специальность 25.00.27 – гидрология суши, водные ресурсы и
гидрохимия

Автореферат
диссертации на соискание ученой степени
доктора технических наук

Санкт-Петербург – 2010

Работа выполнена в Государственном образовательном учреждении высшего профессионального образования «Российский государственный гидрометеорологический университет» на кафедре гидрологии суши

Официальные оппоненты: доктор физико-математических наук
Кондратьев Сергей Алексеевич
(Институт озероведения РАН)

доктор технических наук
Болгов Михаил Васильевич
(Институт водных проблем РАН)

доктор технических наук
Бузин Владимир Александрович
(Государственный гидрологический институт)

Ведущая организация: Научно-исследовательский гидрометеорологический центр Российской Федерации
(г. Москва)

Защита состоится 18 марта 2010 года в 15 часов 30 минут на заседании диссертационного совета Д 212.197.02 при ГОУ ВПО «Российский государственный гидрометеорологический университет» по адресу: 195196, г. Санкт-Петербург, Малоохтинский проспект, 98

С диссертацией можно ознакомиться в библиотеке ГОУ ВПО «Российский государственный гидрометеорологический университет»

Автореферат разослан «18» февраля 2010 г.

Ученый секретарь
диссертационного совета Д 212.197.02
кандидат географических наук,
профессор

В.Н. Воробьев

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Точное и своевременное прогнозирование катастрофических наводнений является одной из наиболее актуальных задач, стоящих перед международным гидрологическим сообществом. В первую очередь, это вызвано стремительными темпами роста ущерба от паводков и половодий, как в Российской Федерации, так и в других странах мира. Например, если в начале XX века среднегодовой ущерб от наводнений в США составил 100 млн. долларов, то в его второй половине он превышал 1 млрд. долларов, а в отдельные годы последнего десятилетия – 10 млрд. долларов; прямой экономический ущерб от наводнений в Китае за первые 8 месяцев 2009 года составил 71,1 миллиарда юаней (около 10,5 миллиарда долларов), пострадали 91,88 миллиона человек и 7,1 миллиона гектар сельхозугодий; тайфун «Кетсана», вызвавший масштабные наводнения во Вьетнаме и на Филиппинах в 2009 г., повлек смерть более 1000 человек, повредил или уничтожил 360 тыс. домов, школ и других учреждений, 73 тыс. га урожая; наводнение на Кубани вызвало гибель 10 человек и материальный ущерб на сумму около 2 млрд. руб.

В условиях изменения климата и возрастающей антропогенной нагрузки на речные водосборы статистические методы прогнозирования теряют свою эффективность, и единственным методом качественного прогнозирования паводков и половодий становится математическое моделирование, эффективность которого во многом зависит от калибровки (параметризации) используемых гидрологических моделей. Однако сравнительно простые модели с небольшим количеством параметров, калибровка которых больших трудностей не вызывает, обычно не в состоянии адекватно описать весь спектр природных процессов, определяющих формирование паводков и половодий. Более сложные модели с большим числом взаимозависимых параметров (такие, как например, модель «Сакраменто») являются достаточно гибкими для моделирования катастрофических паводков и половодий в различных условиях, но их эффективная калибровка связана со значительными техническими трудно-

стями, обусловленными ограниченностью процессорных ресурсов. Кроме того, в результате выполнения диссертационной работы было установлено, что существующие методы оптимизации параметров сложных моделей не в состоянии учесть ряд свойств многомерной поверхности целевой функции, на которой отыскиваются оптимальные наборы параметров, и по этой причине не могут применяться при автоматической калибровке оперативных гидрологических моделей.

Особую теоретическую и практическую значимость проведенным исследованиям придает возможность их непосредственного использования в рамках программы Международной Ассоциации Гидрологических Наук (IAHS) PUB (Prediction in Ungauged Basins), а так же для усовершенствования системы прогнозирования катастрофических паводков и половодий в рамках программы «Модернизация и техническое перевооружение учреждений и организаций Росгидромета», выполнение которой начато в соответствии с Соглашением о займе (№4769 RU) между Российской Федерацией и Международным Банком Реконструкции и Развития.

Цель работы. Разработка комплекса усовершенствованных методов автоматической калибровки многопараметрических гидрологических моделей с взаимозависимыми параметрами (частным случаем которых являются модели с независимыми параметрами), используемых для оперативного прогнозирования опасных гидрологических явлений при различном уровне обеспеченности исходными гидрометеорологическими данными, а также повышение точности прогнозирования дождевых паводков и половодий в условиях изменения климата и переменной антропогенной нагрузки на малые и средние водосборы.

Задачи исследования. Поскольку диссертационная работа направлена на решение практической проблемы автоматизированного прогнозирования дождевых паводков в различных условиях (включая неизученные водосборы),

перечень основных задач включает как фундаментальные, так и технические задачи, решение которых обеспечивает целостность данного исследования:

1. Анализ поведения поверхности многомерной целевой функции гидрологических моделей с взаимозависимыми параметрами (на примере модели «Сакраменто») при изменении внешних условий, включая изменение «входа» модели и длины временных рядов, используемых для калибровки гидрологической модели;
2. Выявление причин низкой эффективности существующих методов автоматической калибровки гидрологических моделей с взаимозависимыми параметрами, связанных с особенностями эволюции поверхности целевой функции;
3. Разработка фундаментального подхода к идентификации оптимальных параметров оперативных гидрологических моделей на основе результатов анализа причин и особенностей эволюции поверхности целевой функции;
4. Разработка практически применимых алгоритмов автоматической калибровки многопараметрических моделей с взаимозависимыми параметрами (на примере модели «Сакраменто») при различном уровне обеспеченности исходными данными, а именно: а) при отсутствии и низкой пространственно-временной дискретности данных об осадках; б) при низкой пространственно-временной дискретности данных о стоке; в) при отсутствии или низкой пространственной дискретности данных о типах и водоудерживающих характеристиках доминирующих почвогрунтов; основным требованием к этим алгоритмам является их универсальность (т.е. возможность использования по умолчанию, в автоматическом режиме) по отношению к любым гидрологическим моделям;
5. Выявление причин низкой эффективности существующих критериев точности гидрологических прогнозов, применяемых в качестве оптимизируемых функционалов при калибровке многопараметрических гидрологических моделей, которые используются для прогнозирования опасных гидрологических явлений (на примере дождевых паводков);

6. Разработка универсальных типов целевой функции, которая может быть использована для автоматической калибровки гидрологических моделей при технической невозможности или экономической нецелесообразности построения матрицы потерь;
7. Создание автоматизированной системы прогнозирования паводков при различном уровне обеспеченности исходной гидрометеорологической информацией на основе разработанной методологии автоматической калибровки многопараметрических моделей;
8. Создание аппаратно-программного комплекса по поиску, передаче и архивированию гидрометеорологических данных, необходимых для прогнозирования опасных гидрологических явлений;
9. Апробация разработанной методологии калибровки моделей для моделирования паводков на реальных водных объектах;
10. Использование результатов исследований в гидрологическом программном обеспечении «SLS+», предназначенном для экспорта в страны Юго-Восточной Азии, Африки, Центральной и Южной Америки.

Объект исследования. Эффективность оперативных многопараметрических моделей, применяемых для прогнозирования дождевых паводков, половодий и других опасных гидрологических явлений в условиях изменения климата и при различном уровне обеспеченности исходными гидрометеорологическими данными.

Предмет исследования. Многомерная невыпуклая поверхность целевой функции, пространственно-временная изменчивость которой является ключевым фактором автоматической калибровки многопараметрических гидрологических моделей, используемых для прогнозирования паводков и половодий.

Методы исследований. Работа выполнена на основе классических и усовершенствованных методов гидрологического моделирования, теории оптимизации, статистического и вероятностного анализа стохастических при-

родных процессов. Анализ и обработка результатов численных экспериментов по определению вероятностных характеристик свойств моделируемых гидрологических процессов проводились как классическими, так и разработанными соискателем статистическими методами с привлечением современной компьютерной техники.

Научная новизна. Научная новизна выполненных исследований заключается в решении ряда фундаментальных и научно-технических задач, открывающем возможность моделирования и прогнозирования стока с малоизученных и неизученных малых и средних водотоков:

1. В разработке методологии и программного обеспечения по определению трансферабельных параметров гидрологических моделей, которые могут быть использованы для моделирования стока с малоизученных и неизученных водосборов;
2. В разработке новых мультимасштабных целевых функций MSOF, ASOF и MISOF, обеспечивающих естественное сглаживание невыпуклых поверхностей многомерных целевых функций;
3. В выполнении детального анализа поведения поверхности многомерной целевой функции при незначительных изменениях обучающих выборок и установлении особенностей ее эволюции во времени;
4. В разработке семейства фундаментальных методов параметризации многопараметрических прогностических моделей с взаимозависимыми параметрами, основанных на квазилокальной оптимизации в физически предопределенном районе области определения параметров (SLS, SLS-F, SLS-2L, SLS-E и др.);
5. В разработке концепции естественной трансформации целевой функции, позволяющей повысить стабильность и надежность определяемых параметров, основанной на вычислении показателей F-робастности (нового статистического подхода, также разработанного соискателем);
6. В создании автоматизированной системы прогнозирования (АСП) «SLS+», позволяющей выполнять фоновое прогнозирование паводков в полностью ав-

томатическом режиме (при различном уровне обеспеченности исходными данными, а также при их различной пространственно-временной дискретности);

7. В создании методологии постобработки прогнозов на основе одновременного выполнения калибровки модели и генерирования случайных ансамблей «входа» модели путем использования алгоритма SLS-E;

8. В создании аппаратно-программного комплекса (АПК) «INWADA», предназначенного для автоматического поиска, копирования, усвоения и архивирования гидрометеорологических данных;

9. В создании концепции и программного обеспечения динамической базы данных (в рамках АПК «INWADA»).

Приведенные фундаментальные и прикладные разработки реализованы в виде учебно-оперативной автоматизированной системы прогнозирования паводков «SLS+» и соответствующего программного обеспечения.

Достоверность полученных результатов обеспечивается корректностью постановки задач исследования, большим объемом использованных данных гидрометеорологических наблюдений, численными экспериментами по моделированию стока 88 водотоков Российской Федерации, Австралии, США, Турции и Японии, применением апробированных аналитических и численных методов анализа и расчета, реализацией алгоритмов и процедур расчета на современной вычислительной технике, объективным анализом полученных результатов и внедрением разработанных методов в оперативное прогнозирование дождевых паводков в США и Австралии.

На защиту выносятся:

Комплекс новых методов и процедур, позволяющих выполнять моделирование и прогнозирование паводочного стока с малых и средних водосборов с разной степенью покрытия гидрометеорологическими наблюдениями (включая малоизученные и неизученные водосборы), а именно:

1. Концепция и методология калибровки гидрологических моделей на основе сглаженных естественным образом поверхностей мультимасштабных целевых функций MSOF, ASOF и MISOF;
2. Метод анализа временной изменчивости многомерных поверхностей многопараметрических целевых функций;
3. Концепция F-робастности – нового статистического метода учета временной устойчивости и стабильности локальных и глобальных оптимумов;
4. Метод текущей трансформации многомерной поверхности целевой функции путем расчета локальных индексов F-робастности в районе выполнения оптимизации;
5. Принцип повышения эффективности автоматической калибровки многопараметрических гидрологических моделей за счет квазилокальной оптимизации в физически предопределенном районе области определения параметров;
6. Методы автоматической калибровки многопараметрических моделей группы SLS (базовый алгоритм SLS и его модификации, применяемые при недостаточности исходных данных – алгоритмы SLS-F, SLS-2L, SLS-E, SLS-2LF, SLS-2LE, SLS-EF и SLS-2LEF);
7. Метод текущей постобработки прогнозов паводков на основе алгоритма SLS-E;
8. Метод идентификации трансферабельных параметров (то есть параметров, которые можно использовать для моделирования стока с неизученных водосборов);
9. Концепция аппаратно-программного комплекса по автоматическому поиску, передаче, усвоению и архивированию гидрометеорологических данных («INWADA»);
10. Концепция и методология оптимизации гидрологических моделей с полураспределенными параметрами;
11. Результаты апробации предложенных методов автоматической калибровки гидрологических моделей (выполненных на примере модели «Сакраменто»), используемых для прогнозирования дождевых паводков.

Практическая значимость проведенных в диссертационной работе исследований заключается в следующем:

1. Разработанные оптимизационные процедуры группы SLS позволяют выполнять прогнозирование паводков в регионах с недостаточно развитой сетью или с низкой пространственно-временной дискретностью гидрометеорологических наблюдений (т.е. на малоизученных и неизученных водосборах);
2. Разработанный метод анализа поведения многомерной поверхности целевой функции позволяет установить причину низкой эффективности существующих оптимизационных процедур, применяемых для калибровки многопараметрических гидрологических моделей с взаимозависимыми параметрами;
3. Разработанные алгоритмы квазилокальной оптимизации параметров в физически предопределенном районе многомерной области определения позволяют выполнять автоматическую калибровку и рекалибровку сложных оперативных гидрологических моделей для большого числа гидрологических объектов (например, водотоков целой страны или крупного региона);
4. Разработанные методы автоматической калибровки многопараметрических гидрологических моделей позволяют прогнозировать паводочный сток с малых и средних водосборов при отсутствии или низкой точности геологических и морфологических данных;
5. Разработанная автоматизированная система прогнозирования паводков «SLS+» может быть использована для эффективного фонового и уточненного прогнозирования паводков при практически любом уровне обеспеченности гидрометеорологической информацией (включая неизученные и малоизученные малые и средние водосборы);
6. Разработанный метод рекалибровки модели и постобработки прогнозов позволяет повысить точность прогнозирования (включая катастрофические);
7. Созданные методы, методики, алгоритмы и процедуры могут быть использованы в качестве методической основы для совершенствования системы прогнозирования паводков в Российской Федерации и других странах; сделанные выводы и полученные рекомендации могут быть использованы при

планировании и выполнении мероприятий по модернизации и техническому перевооружению учреждений и организаций Росгидромета.

Реализация результатов работы. Представленный в работе метод автоматической калибровки гидрологической модели «Сакраменто» (SLS) внедрен в оперативное прогнозирование дождевых паводков Национальной Службой Погоды США (National Weather Service) и включен в план развития системы гидрологического прогнозирования Австралии (в рамках проекта eWater). Разработанная соискателем автоматизированная система прогнозирования паводков «SLS+» используется в ГОУ ВПО «Российский государственный гидрометеорологический университет» при подготовке студентов (бакалавров, инженеров и магистров) по специальности «гидрометеорология». Программное обеспечение «SLS+» представляет собой коммерческий продукт, предназначенный для экспорта в страны Юго-Восточной Азии, Африки, Центральной и Южной Америки. Аппаратно-программный комплекс «INWADA» используется в ГОУ ВПО «Российский государственный гидрометеорологический университет» для пополнения и обслуживания динамической базы гидрометеорологических данных. Методические разработки, алгоритмы и процедуры, предложенные соискателем, могут быть использованы при модернизации и техническом перевооружении учреждений и организаций Росгидромета

Апробация работы. Основные результаты работы были представлены на следующих российских и международных научно-технических конференциях, совещаниях и симпозиумах: Международном совещании «Stochastic models of hydrological processes and their applications to problems of environmental preservations» (Москва, РФ, 1998 г.), Европейской конференции по инновациям в исследовании наводнений (Потсдам, Германия, 2000), Международном симпозиуме «River flood defence» (Кассель, Германия, 2000 г.), совещаниях в Техническом Университете г. Дельфт (Нидерланды, 2000 и 2001 гг.), Международной конференции «Современные проблемы стохастической гидрологии»

(Москва, 2001 г.), Международном симпозиуме «Hydrological challenges in transboundary water resources management» (Кобленц, Германия, 2001 г.), весенних конференциях Американского Геофизического Союза (Бостон, США, 2001 г., Вашингтон, США, 2002 г., Ницца, Франция, 2003 г., Монреаль, Канада, 2004 г., Балтимор, США, 2006 г.), Международном совещании «River runoff: minima and maxima» (Санкт-Петербург, РФ, 2001 г.), Международном семинаре «Extreme value theory and applications» (Гетеборг, Швеция, 2001 г.), 5-ом Международном совещании «Application of remote sensing in hydrology» (Монпелье, Франция, 2001 г.), Международном симпозиуме «The extremes of the extremes» (Рейкьявик, Исландия, 2003 г.), конференции Международного союза геодезии и геофизики (Саппоро, Япония, 2003 г.), Международной конференции «Hydrology, water resources and environment» (Нанкин, КНР, 2003 г.), Международном совещании по проекту «The distributed models inter-comparison project (DMIP)» (Силвер-Спринг, США, 2004 г.), совещаниях и ежегодных конференциях проекта «eWater» (Голд-Кост, Австралия, 2007 и 2008 гг.), Международном совещании «The catchment-scale hydrological modeling & data assimilation (CANMDA III)» (Мельбурн, Австралия, 2008 г.), 31-ом Международном симпозиуме по водным ресурсам и 4-ой Международной конференции по водным ресурсам и экологическим исследованиям «Water down under 2008» (Аделаида, Австралия, 2008 г.), 2-ой Международной конференции «HydroEco 2009» (Вена, Австрия, 2009 г.), Генеральной Ассамблее Европейского геофизического союза (Вена, Австрия, 2009 г.), семинарах в Институте водных проблем РАН и Гидрометцентре РФ (2010 г.).

Публикации. По теме диссертации опубликованы 46 работ (в том числе 8 в изданиях, определенных ВАК РФ, и 26 работ в зарубежных изданиях на английском языке).

Структура и объем работы. Диссертация состоит из введения, пяти глав, основных результатов и выводов, списка используемой литературы из

198 наименований и приложений. Работа содержит 223 страницы текста, включая 32 рисунка и 25 таблиц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснована актуальность темы исследования, сформулированы цель и задачи диссертационной работы, обоснована научная новизна и дается краткое описание содержания разделов диссертации.

Глава 1 диссертации посвящена основным принципам автоматизированного прогнозирования опасных гидрологических явлений (ОГЯ) и базирующейся на них последовательности действий при разработке прогностических методов, выполнение которых обязательно для достижения максимальной экономической эффективности прогнозирования данной гидрологической переменной в данном месте при данных условиях. Этот раздел необходим для обоснования необходимости систем автоматизированного фонового прогнозирования стока на больших территориях (в частности, как альтернативы по отношению к частным прогностическим методикам, разработанных для отдельных водосборов).

При разработке прогностического метода, во-первых, необходимо четко определить цель прогноза (то есть пару «прогнозируемая переменная — водный объект»). Чем географически крупнее и физически разнообразнее объект, тем универсальнее должна быть модель. Если объектом является один небольшой водосбор или водоток, модель может быть весьма простой (но при этом она должна адекватно описывать все те процессы, в результате которых формируется ОГЯ) и обладать сравнительно небольшим числом независимых параметров, которые легко идентифицировать. Именно поэтому разработка физико-математической модели конкретного водосбора или водотока является технически идеальным путем решения проблемы прогнозирования, однако с экономической точки зрения этот путь представляется малоэффективным и неприменимым для фонового прогнозирования стока на больших territori-

ях, когда корректное описание всех стокообразующих процессов и факторов в рамках одной модели с технической точки зрения крайне затруднительно или даже невозможно, поэтому часть процессов, оказывающих влияние на формирование стока, может быть смоделирована лишь абстрактно (концептуально). На рисунке 1 показана диаграмма типов гидрологических моделей, позволяющая правильно выбрать модель для решаемой задачи в зависимости от степени понимания процесса формирования стока и степени неопределенности, которая, в первую очередь, зависит уровня обеспеченности исходными данными. В частности, в соответствии с этой диаграммой:

- физико-математические (детерминистические) модели следует использовать для моделирования стока в тех случаях, когда процесс формирования стока полностью понятен и достаточно хорошо освещен данными наблюдений;
- стохастические (динамико-стохастические) модели следует использовать в тех случаях, когда процесс формирования стока в целом понятен, но не полностью хорошо освещен данными наблюдений; интегральная неопределенность отражается при помощи разнообразных вероятностных подходов;
- статистические модели следует использовать в тех случаях, когда основные причинно-следственные связи известны лишь в первом приближении и не могут быть отражены физически корректным дифференциальным уравнением, в то же время имеются достаточные данные для построения статистических связей между предикторами и предиктантом;
- концептуальные модели следует использовать в тех случаях, когда возможно лишь абстрактное выражение части стокообразующих процессов; интегральная неопределенность отражается при помощи разнообразных вероятностных подходов, а также принципами концептуализации и степенью абстрагирования, что является весьма гибким инструментом моделирования процессов с высоким уровнем неопределенности.

Исходя из вышесказанного, был определен круг моделей, наилучшим образом подходящих для прогнозирования паводочного стока и потенциально подходящих для использования в качестве главного функционального элемента АСП «SLS+»: модель «Сакраменто» (The Sacramento Soil Moisture Accounting Model), модель «VIC» (Variable Infiltration Capacity macro-scale hydrologic model), позволяющая учитывать и водный, и энергетический баланс, австралийская воднобалансовая модель «AWBM» (Australian Water Balance Model), модель «SIMHYD» (от англ. «SIMplified HYDrograph» – упрощенный гидрограф), модель PDM (probability distribution model), а также модель «SMAR» (Soil Moisture Accounting Runoff model). Анализ этих моделей позволил выбрать в качестве основной модели модель «Сакраменто» (о чем подробно рассказано в Главе 2).

При разработке следующего поколения АСП будут также рассмотрены модели формирования талого стока, причем наиболее перспективными будут считаться те модели, которые позволяют выполнять эффективное моделирование, не только половодий, но и паводков.

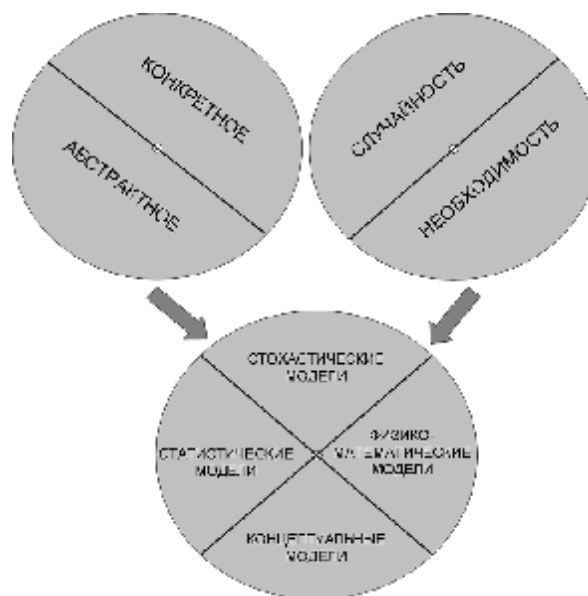


Рисунок 1 — Диаграмма типов моделей с точки зрения их отношения к категориям конкретного и абстрактного, случайности и необходимости

Во-вторых, необходимо определить целевую функцию, являющуюся формализованным критерием точности выпускаемых прогнозов и эффективности применяемого метода прогнозирования. Почти идеальным видом целевой функции при оценивании выпущенных прогнозов является матрица потерь, отражающая размер материального ущерба L_f как функции от погрешности прогноза ΔQ при различных фактических значениях прогнозируемой переменной Q : $L_f = f(\Delta Q|Q)$. Интегральные потери за весь период прогнозирования, и просуммированные с общей стоимостью прогнозирования P , выражают общие затраты P_Σ :

$$P_\Sigma = P + L_f. \quad (1)$$

Таким образом, даже при абсолютно точном прогнозировании затраты на выпуск прогнозов равны стоимости прогнозирования P , которая складывается из стоимости получения, передачи и обработки данных P_D , стоимости моделирования P_M (включая стоимость самой модели, расходы на ее калибровку, стоимость постобработки прогнозов и заработную плату персонала) и стоимости представления результатов прогнозирования. Экономическая эффективность прогнозирования E выражается через размер уменьшения потерь в результате использования прогнозов:

$$E = L_0 - L_f, \quad (2)$$

где L_0 – ущерб, нанесенный ОГЯ в случае, если никакого прогнозирования не было, и превентивные меры не принимались. Построение матрицы потерь требует определенных затрат, поэтому при прогнозировании ОГЯ на больших территориях этот шаг значительно повышает расходы на прогнозирование. Кроме того, матрица потерь является неэффективной при моделировании низких уровней или расходов воды, при которых потери, вызванные почти любой

погрешностью, малы. (По этой причине модель может достаточно точно описать пиковые расходы или уровни, но, например, кривая спада будет рассчитана неточно; в свою очередь, это приведет к заданию неверных начальных условий перед очередным катастрофическим паводком, который будет спрогнозирован также неверно). В этом случае в качестве целевой функции нужно выбрать некое математическое выражение J , отражающее уровень неопределенности прогнозов $\overline{\Delta Q}$ (т.е. $J = f(\overline{\Delta Q})$), минимизация которого будет означать повышение точности прогнозов. (В Главе 3 рассмотрена мультимасштабная целевая функция MSOF и ее модификации, позволяющие добиваться высокой эффективности прогнозирования в тех случаях, когда определение матрицы потерь технически невозможно или экономически нецелесообразно).

В-третьих, нужно сформулировать критерии целесообразности прогнозирования данной гидрологической переменной с различной заблаговременностью в данном месте. Очевидно, что чем больше заблаговременность точного прогноза T_f , тем меньше потери L_f . Но при увеличении заблаговременности неминуемо увеличивается уровень неопределенности прогнозов $\overline{\Delta Q}$, и в какой-то момент затраты на прогнозирование становятся больше, чем размер потерь, которых можно избежать в результате использования прогнозов. Поэтому каждое значение заблаговременности прогноза можно сопоставить с его экономической эффективностью:

$$T_{f,i} \leftrightarrow E_i. \quad (3)$$

Прогнозирование целесообразно, если расходы на прогнозирование заведомо меньше потерь, понесенных в случае, если результаты прогнозирования не использовались. Именно этот факт определяет необходимость выпуска оценочных фоновых прогнозов стока для больших территорий: в самом деле, разработка собственной методики прогнозирования стока для экономически незначимого водосбора или водосбора, на котором ОГЯ наблюдаются сравнительно редко, в большинстве случаев экономически нецелесообразна.

В-четвертых, необходимо определить оптимальную заблаговременность прогнозирования T_f^{λ} , при которой достигается максимальная экономическая эффективность прогнозирования $E: T_f^{\lambda} = T_f | E \rightarrow \max$. Реальная заблаговременность прогноза T_f складывается из:

1. Заблаговременности получения прогнозистом или моделью информации о предикторах T_{pr} , которая состоит из естественной заблаговременности (то есть инерционности самого быстроразвивающегося предиктора из n) \bar{T}_{pr} и времени передачи информации обо всем наборе предикторов $T_{pr}^T: T_{pr} = \bar{T}_{pr} - T_{pr}^T$ (например, если решение о паводочном водосбросе с гидроэлектростанции было принято за 3 суток, а информация о его объеме дошла до прогнозиста через 2 суток после принятия этого решения, заблаговременность получения информации о предикторах составляет 1 сутки; если в качестве предиктора используются данные наблюдений ($\bar{T}_{pr} = 0$), то величина T_{pr} может быть отрицательной, поскольку время передачи информации T_{pr}^T всегда неотрицательно);
2. Времени обработки информации о предикторах T_{pr}^{da} (например, процедуры усвоения полученных данных – data assimilation), взятого со знаком «минус»;
3. Времени формирования прогнозируемого процесса T_{pf} ;
4. Времени расчета T_M по какой-либо модели или методике M (включая их текущую настройку или калибровку), взятого со знаком «минус»;
5. Времени постобработки выпущенного прогноза (post-processing time) T_{pp} , взятого со знаком «минус»;
6. Естественной инерционности прогнозируемого процесса T_{in} , которая определяет верхнюю и нижнюю границу заблаговременности (например, если паводок начинается через 1 час после выпадения осадков на водосборе и полностью прекращается через 10 часов, то инерционность процесса составляет 9

часов, а время формирования $T_{pf} = 1$ ч). Таким образом, общее уравнение для расчета реальной заблаговременности прогноза T_f выглядит так:

$$T_f = (\bar{T}_{pr} - T_{pr}^T) - T_{pr}^{da} + T_{pf} - T_M - T_{pp} + T_{in}. \quad (4)$$

Это уравнение, в частности, объясняет необходимость обработки поступающей информации и расчета стока с использованием достаточно мощных и быстродействующих процессоров и максимально возможной унификации этих процедур для больших территорий.

Наконец, в-пятых, выбирается оптимальная модель \hat{M} , обеспечивающая максимальную экономическую эффективность прогнозирования E (при заданной целевой функции J и оптимальной заблаговременности \hat{T}_f , оцененной для некой обучающей выборки S_{t_1, t_2} , которая описывает моделируемый процесс с момента времени t_1 до момента времени t_2):

$$\hat{M} = M | (T_{f, M}; J; S_{t_1, t_2}) : E \rightarrow \max. \quad (5)$$

Таким образом, уравнение (5) выражает общий принцип выбора модели, целевой функции и обучающих временных рядов при прогнозировании ОГЯ, который учитывался при выполнении исследований, описанных в последующих разделах диссертационной работы. Кроме этого, уравнение (5) частично объясняет причину снижения эффективности статистических методов прогнозирования, основанных на предположении неизменности \hat{M} от интервала времени $[t_1; t_2]$.

Завершается Глава 1 рассмотрением общих принципов параметризации прогностической модели и вопросов представления результатов прогнозирования, как фактора, влияющего на фундаментальную часть прогностических методов.

В Главе 2 приводится подробное описание концептуальной гидрологической модели «Сакраменто», которая является наиболее широко используемой оперативной моделью в мире и которая была выбрана в качестве основного функционального элемента АСП «SLS+» (в частности, из-за ее экономической эффективности, отраженной в уравнении (5)), а также причины этого выбора. Основанием для выбора модели «Сакраменто» является, во-первых, обзор литературы, посвященной моделированию дождевых паводков, во-вторых, консультации с ведущими экспертами в области моделирования стока из США, Австралии, Великобритании, Японии, Турции, Нидерландов и Российской Федерации, а в-третьих – численные эксперименты, выполненные соискателем в рамках нескольких российских и иностранных проектов (перечисленных в разделе «Основные результаты и выводы»), которые целиком подтверждают выводы предварительного анализа.

Глава 3 посвящена мультимасштабным целевым функции, которые наиболее робастны по отношению к обучающим выборкам и, следовательно, наиболее эффективны для прогнозирования катастрофических паводков и половодий. В Главе 1 уже отмечалось, что при параметризации прогностической модели необходимо обратить внимание на символ S_{t_1, t_2} , отражающий в уравнении (5) обучающую выборку – временные ряды предикторов и предиктанта за период времени с t_1 до t_2 . Дело в том, что модель, оказавшаяся оптимальной (то есть, наиболее эффективной) за период времени $[t_1; t_2]$ может отличаться от модели, наилучшим образом описывающей прогнозируемый процесс за период времени $[t_2; t_3]$. Поэтому уравнение (5) можно использовать для выбора лучшей модели за некий промежуток времени в прошлом (при условии, что используемые временные ряды достаточно репрезентативны по отношению к описываемому процессу), но при намерении моделировать экстремальные значения, например, расхода воды Q , использование этого уравнение не станет наилучшим способом определения параметров. В частности, идентификация параметров модели, основанная на максимизации экономической эффективности E может привести к тому, что периоды спада высоких

расходов и уровней и периоды низкого стока, мало влияющие на величину E , будут смоделированы плохо (в отличие от нескольких паводков или половодий, определяющих размер потерь L_f и, следовательно, эффективность прогнозирования). Это означает, что начальные условия (например, степень увлажненности рассматриваемой территории перед началом очередного паводка) будут заданы неверно, поэтому этот паводок, который вполне может оказаться катастрофическим, также будет смоделирован неверно. В строгом смысле слова, говорить об оптимальности модели для прогнозирования данного катастрофического явления можно лишь постериорно: даже если исследуемая модель блестяще описала все предыдущие паводки, она может подвести при прогнозировании следующего. Поэтому при калибровке модели необходимо помнить, что метод калибровки, вид целевой функции и выбранные обучающие выборки должны обеспечивать максимально точное отражение разных фаз гидрологического режима (включая подъем уровней, пик паводка или половодья, спад уровней и низкий сток) при различном порядке их чередования (например, единичный паводок после засухи, серия из нескольких паводков и т.д.). Эта точка зрения всецело подтверждается численными экспериментами, проведенными соискателем в NOAA. Полученные результаты позволяют сделать вывод о том, что наиболее перспективным видом целевой функции является мультимасштабные целевые функции группы MSOF (от англ. «Multi-Scale Objective Function»). Концепция целевых функций MSOF была предложена В. И. Корнем с целью симитировать «мультимасштабный» характер «ручной» калибровки и отразить разные частоты стока. В проведенных экспериментах была использована целевая функция, отражающая дисперсию погрешностей относительно разных периодов осреднения:

$$J = MSOF = \sqrt{\sum_{k=1}^n \left(\frac{S_1}{S_k} \right)^2 \sum_{i=1}^{m_k} (q_{o,k,i} - q_{s,k,i}(X))^2}, \quad (6)$$

где $q_{o,k,i}$ и $q_{s,k,i}$ — измеренные и смоделированные расходы воды, осредненные за интервал времени k , s_k — среднеквадратические отклонения расхода воды масштаба k , n — общее число масштабов, m_k — число элементов каждого масштаба k . Для калибровки тестовых водосборов, расположенных в США, использовался критерий MSOF с часовым, суточным, недельным и месячным масштабами ($k = 1, 2, 3, 4$). При калибровке российских, австралийских и турецких водосборов часовой масштаб был исключен ($k = 1, 2, 3$). Заметим, что каждый из элементов суммы в уравнении (6) имеет весовой коэффициент, равный среднеквадратическому отклонению измеренных расходов соответствующих масштабов. Подобная схема «взвешивания» основана на предположении, что неопределенность моделируемого стока в пределах каждого из масштабов пропорциональна естественной изменчивости стока для этих масштабов. Другая причина использования критериев группы MSOF — это дополнительное сглаживание поверхности целевой функции, которое уменьшает вероятность «застревания» процесса поиска оптимума в незначительных «депрессиях». Необходимо подчеркнуть, что одной из разновидностей целевой функции MSOF является «всемасштабная» целевая функция ASOF (от англ. «All-Scale Objective Function»), которая отражает все возможные частоты от, например 1 часа (при часовой дискретности) до количества часов, составляющих интервал $[t_1, t_2]$ или его часть:

$$J = ASOF = \sqrt{\sum_{k_1=1}^{k_n=t_2} \left(\frac{s_1}{s_k} \right)^2 \sum_{i=1}^{m_k} (q_{o,k,i} - q_{s,k,i}(X))^2}. \quad (7)$$

Например, при проведении численных экспериментов по моделированию стока австралийского ручья Куямба в створе Леди Смит (Kyeamba Creek – Lady Smith) использовалась ASOF с масштабами от 1 до 720 часов: 1, 2, 3, ..., 719, 720 (рисунок 2). В результате были зафиксированы следующие значения целевых функций:

- при прогнозе с использованием параметров, найденных в соответствии с рекомендациями Корня (то есть без калибровки) – 2,12 м³/с,
- при прогнозе с калибровкой методом SCE (Shuffled Complex Evolution) и целевой функцией MSOF – 1,23 м³/с,
- при прогнозе с калибровкой методом SLS (Stepwise Line Search) и целевой функцией MSOF – 1,21 м³/с,
- при прогнозе с калибровкой методом SCE и целевой функцией ASOF – 0,89 м³/с,
- при прогнозе с калибровкой методом SLS и целевой функцией ASOF – 0,80 м³/с.

В качестве еще одного частного случая ASOF можно также рассмотреть целевую функцию MISOF (от англ. «Most Informative Scales Objective Function»), которая отражает только наиболее информативные частоты моделируемого процесса и выражается уравнением, общим для всех MSOF:

$$J = MISOF = \sqrt{\sum_{i=1}^{i=m} \left(\frac{s_1}{s_k} \right)^2 \sum_{i=1}^{m_k} (q_{o,k,i} - q_{s,k,i}(X))^2} . \quad (8)$$

В целевой функции MISOF масштабы отражают только наиболее информативные (с какой-либо точки зрения) частоты, которые могут быть назначены субъективно (например, 1 час, 24 часа, 240 часов и 720 часов, как в примере с MSOF) или идентифицированы с использованием какого-либо формального признака (например, можно выбрать фиксированное число масштабов, обладающих наибольшими значениями отношения дисперсии к коэффициенту автокорреляции с шагом длины масштаба k или радиусу корреляции $I_k = s_k^2 / r(k)$). Нетрудно заметить, что в предельных случаях целевая функция MISOF может вырождаться либо в среднеквадратическое отклонение 1-часовых ошибок (только один 1-часовой или 1-суточный масштаб), либо в разность средних значений моделируемого и фактического гидрографов (один масштаб, равный длине периода калибровки $[t_1, t_2]$), либо в ASOF (все масшта-

бы) и т.д., что делает целевую функцию MISOF исключительно удобной для использования при автоматической калибровке моделей в различных условиях формирования стока и при различных стохастических особенностях моделируемого гидрографа.

Однако необходимо подчеркнуть, что разработка критериев ASOF и MISOF не входила в первоначальный план исследований, поэтому они в данной диссертационной работе для калибровки модели не использовались, а лишь были протестированы в ходе небольшого числа экспериментов, подобных эксперименту по моделированию стока ручья Куямба, выполненных по данным о 12 малых водотоках США и Австралии. Дальнейшее применение этих видов целевой функции требует тщательного экспериментального подтверждения и нуждается в апробации.

Выбор целевой функции – это лишь первая ключевая составляющая оптимизации параметров. Второй ключевой составляющей является выбор эффективной оптимизационной процедуры, этот важнейший аспект рассмотрен в Главе 4. При моделировании опасных гидрологических явлений крайне важно, чтобы используемые параметры имели физически реалистичные значения. Наличие нереальных или значительно искаженных параметров, имеющих физическую интерпретацию, во-первых, является свидетельством неадекватности модели по отношению к прогнозируемому процессу и, во-вторых, как следствие, снижает практическую уверенность в прогнозе. Поэтому калибровку модели целесообразно производить на основе квазилокальной оптимизации в физически predetermined районе области определения параметров. Для этой цели можно использовать, например, алгоритм SLS (от англ. «Stepwise Line Search»), формально относящийся к так называемым шаблонным процедурам нелинейной оптимизации. Необходимо подчеркнуть, что если модель (имеется в виду ее базовая версия с сосредоточенными параметрами) сравнительно проста и содержит не более 3–4 независимых параметров, для ее калибровки можно использовать такие классические алгоритмы, как метод координатного спуска, градиентный метод, метод Розенброка и т.п.

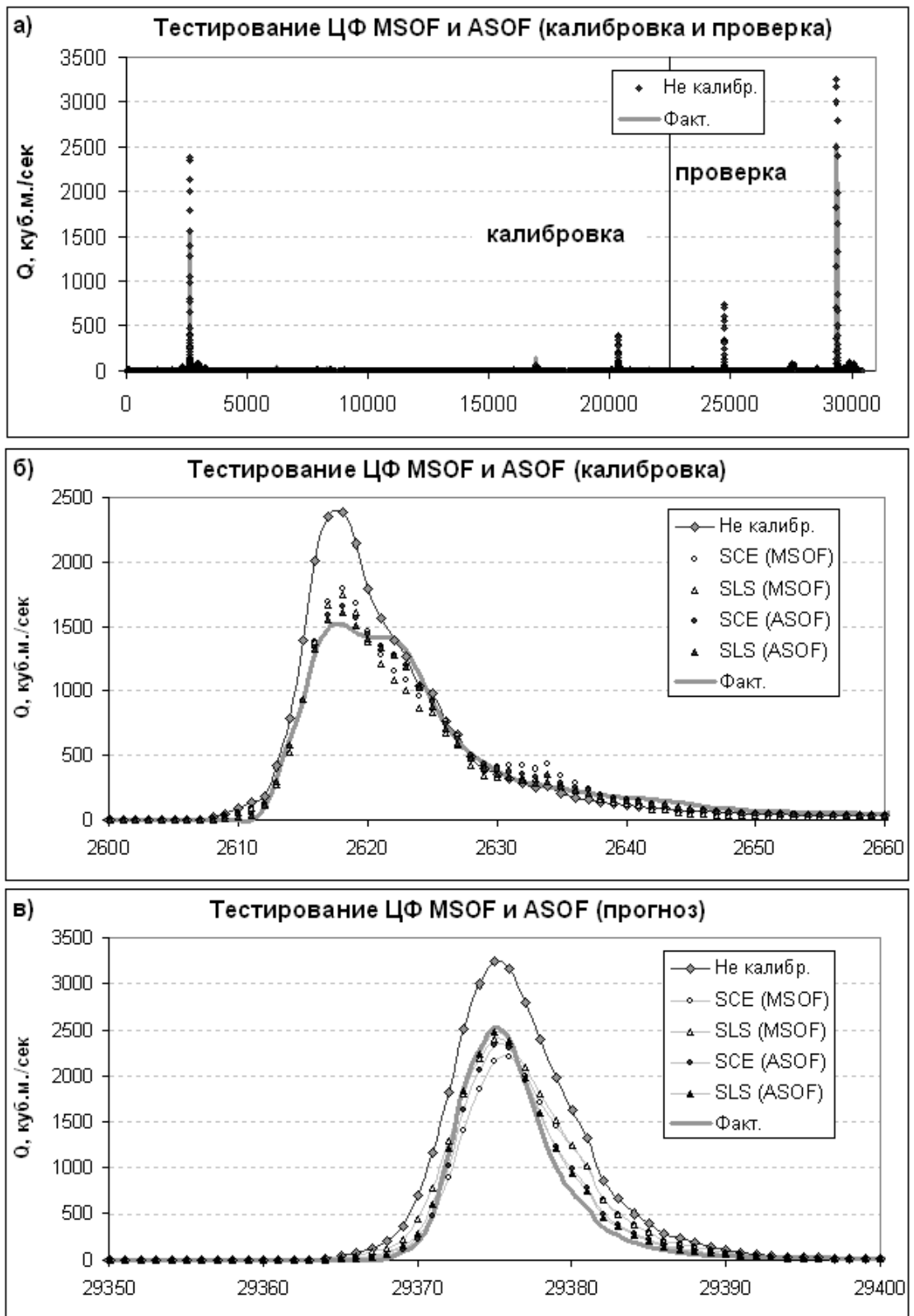


Рисунок 2 — Тестирование целевых функций MSOF и ASOF (руч. Куямба (Куеамба Creek) – Леди Смит (Lady Smith), Австралия); время указано в часах, начиная с 01.01.2000.

Если число параметров, некоторые из которых являются взаимозависимыми, равняется 5–6, то тестирование целесообразно начать с метода SCE, а при большем числе взаимозависимых параметров следует использовать методы группы SLS, однако для их эффективного применения необходимо также наличие информации для физического предопределения района поиска оптимальных параметров, поскольку этот подход позволяет ограничить район поиска и избежать вынужденного увеличения шага оптимизации, а значит и округления поиска.

При калибровке моделей с распределенными и полураспределенными параметрами необходимо исходить из имеющихся в распоряжении прогнозиста ресурсов процессора, которые определяют предельно допустимое число запусков модели (и, следовательно, сравниваемых наборов параметров) v_{lim} и предельную длительность калибровки $T_{cal\ lim} = v_{lim} t$ (где t – длительность одного запуска модели), которая влияет на экономическую эффективность прогнозирования. Далее определяется дискретность рассредоточения параметров. Число частей водосбора или элементарных площадок N_s с индивидуальным набором параметров в первом приближении оценивается по формуле

$$N_s \approx T_{cal\ lim} / \bar{T}_{cal}, \quad (9)$$

где \bar{T}_{cal} – среднее время оптимизации одного набора параметров. Необходимо заметить, что при калибровке моделей с распределенными и полураспределенными параметрами выполняется последовательный «шаблонный» перебор каждого вида параметров по всем ячейкам (элементарным площадкам). В случае, если отдельные части водосбора оснащены гидрологическими постами, то их параметризацию следует производить независимо от других частей.

Для того, чтобы добиться большей устойчивости и стабильности топографии поверхности целевой функции во времени, была разработана и использована для естественного сглаживания процедура, основанная на так на-

зывается F-преобразованием, которое заключается в замене исходной многомерной поверхности $J(\vec{P})$ одномерной поверхностью $F(\vec{P})$. Остановимся на этом подходе более подробно, поскольку он представляет собой фундаментальную основу для идентификации трансферабельных параметров, используемых для моделирования стока с неизученных водосборов.

Временная стабильность (или нестабильность) может быть отражена при помощи F-робастной целевой функции (буква «F» может обозначать «forecast», «feasibility», «fluctuations» – «прогноз», «обоснованность», «флуктуации» и т.д.). В общем случае, термин «F-робастность» отражает способность поверхности минимизируемого функционала приводить к достаточно хорошему результату прогнозирования даже после некоторого смещения этой поверхности в области определения параметров (обычно на 1–2 шага в обоих направлениях). Следовательно, это означает и временную стабильность, робастность по отношению к небольшим изменениям длины временного ряда, что автоматически ведет к прогностической эффективности модели и ее параметров. Модель или набор параметров считаются наиболее F-робастными, если n-параметрический относительный индекс F^n является наименьшим из сравниваемых:

$$F^n = \frac{\int_{a_1}^{A_1} \int_{a_n}^{A_n} D(P) dP_1 \dots dP_n + D_{\lim}(P) \left(\prod_{i=1}^{i=n} (P_{i,\max} - P_{i,\min}) - \prod_{i=1}^{i=n} (A_i - a_i) \right)}{D_{\lim}(P) \prod_{i=1}^{i=n} (P_{i,\max} - P_{i,\min})} \quad (10)$$

где $D(P)$ — минимизируемый функционал (например, среднеквадратическая погрешность прогноза, или мультимасштабная целевая функция MSOF), или любой другой критерий оценивания качества прогнозов в зависимости от параметров P , A_i и a_i — верхняя и нижняя граница параметра P_i , $P_{i,\max}$ и $P_{i,\min}$ — наибольшее и наименьшее значения параметра P_i , при которых значение критерия $D_i(P)$ остается приемлемым, $D_{\lim}(P)$ — наибольшее приемлемое значение

ние критерия $D(P)$ (например, в России широко используется отношение $D = S/s$, представляющее собой отношение среднеквадратической погрешности прогнозов к среднеквадратической погрешности «природного»(или инерционного) прогноза, $D_{\text{lim}}(P) = 0.8$).

Этот индекс показывает, насколько точность модели зависит от погрешности определения параметров. Чем меньше F^n , тем шире пределы допустимых значений параметров P , тем робастнее модель. В случае, если величина предельно допустимой неточности $D_{\text{lim}}(P)$ не определена или же если стоит задача сравнения нескольких оптимумов из одного и того же пространства J (например, при автоматической калибровке гидрологических моделей), индекс F^n может быть найден по упрощенному уравнению:

$$F^{n,r} = \int_{P-r}^{P+r} J dP, \quad (11)$$

где r — это радиус осреднения значений целевой функции (F-радиус). В практических задачах, когда целевая функция $J(P)$ дискретна, индекс F^n определяется именно по уравнению (2а), как среднее значение $J(P)$ в пределах определенного количества шагов s по всем направлениям от исследуемого вектора параметров (например, 1, 2, 3 или более):

$$F^{n,s} = \left(\sum_{i=1}^{i=n} \left[\sum_{j=1}^{j=s+1} J_{i,j} / (2s+1) \right] \right) / n, \quad (12)$$

где $J_{i,j}$ — целевая функция для n параметров и s -шагового радиуса вокруг каждого из них.

Оптимизируя трансформированную целевую функцию $J^F = F^n$ или $J^F = F^{n,s}$, можно найти наиболее стабильные параметры. Этот способ практи-

чески применим при сравнительно небольшом числе параметров. Впрочем, это зависит от имеющихся ресурсов процессора. Например, если уравнение (12) используется для оценивания индексов F^n или $F^{n,s}$ в m точках по обе стороны и в центре исследуемого вектора параметров, то число расчетов по модели (число выполненных оценок J) равно $(m+1)^n$. Пусть v – число сравниваемых наборов параметров. Тогда общее число запусков модели равно $v(m+1)^n$. Если время одного запуска обозначить через t , то общее время T расчета F-индекса равно

$$T = vt(m+1)^n. \quad (13)$$

Это уравнение может быть использовано для определения подходящего количества точек m , необходимых для осреднения, исходя из имеющегося времени:

$$m = \sqrt[n]{T/vt} - 1. \quad (14)$$

Отсюда легко найти величину приращения по каждому из параметров r/m . Независимые параметры можно оценивать отдельно, поэтому общее число параметров при отдельном оценивании F-индекса может быть уменьшено. Сэкономленное время можно использовать для сравнения большего числа наборов параметров, уменьшения шага и т.д. В заключение заметим, что в простейшем случае (т.е. если F-радиус равен 0), уравнения (11) и (12) вырождаются до $F^n = J_i$. Таким образом, F-индексы представляют собой расширение обычной целевой функции (и наоборот, значение целевой функции J_i в точке равно F-index с радиусом $r = 0$). В ходе проведения численных экспериментов по идентификации параметров модели «Сакраменто», было установлено, что многомерные поверхности таких популярных в гидрологии целевых функций как среднеквадратическое отклонение или средняя абсолютная погрешность являются существенно невыпуклыми и весьма чувствительными по отношению даже к незначительному удлинению ($\Delta \sim 10^{-5}$) обучающих выборок.

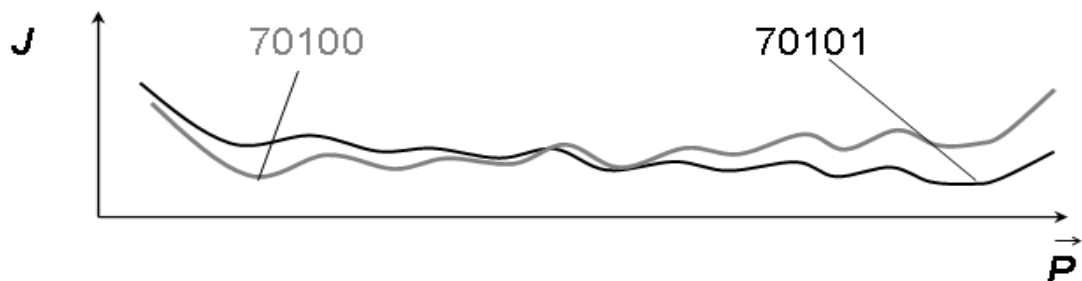


Рисунок 3 — Схема «дна» поверхности целевой функции. Изменение «критической массы» приводит к значительному изменению местоположения «глобального» оптимума.

Поверхности мультимасштабных целевых функций MSOF, ASOF и MISOF являются более сглаженными, но и они чувствительны по отношению к длине обучающих выборок. Поскольку модель «Сакраменто» имеет как минимум 11 калибруемых параметров (еще 5 параметров обычно принимаются постоянными, но при необходимости их идентифицируют различными способами, включая непосредственный расчет и оптимизацию). Некоторые из этих параметров взаимозависимы и, кроме того, поверхность целевой функции является 12-мерной (при 11 оптимизируемых параметрах), поэтому наглядно представить вид такой поверхности и изобразить ее графически технически невозможно. По этим причинам эволюция топографии этой многомерной поверхности при незначительном удлинении обучающих выборок s , например, 70100 элементов до 70101 элемента проиллюстрирована схематически на рисунке 3. Как показали многочисленные эксперименты, добавление даже небольшого числа новых значений приводит к смещению «критической массы» (целевой функции $J = f(\vec{P})$); общая топография локальных оптимумов при этом почти не меняется, а вот местоположение глобального оптимума может существенно сместиться. Более того, во многих случаях «глобальные» оптимумы не имеют никакого физического смысла, и их использование для прогнозирования паводков ведет к неудовлетворительному результату, поскольку физически некорректные параметры не позволяют точно рассчитать началь-

ные условия (например, содержание влаги в почве) перед началом очередного паводка.

На рисунке 4 показана чувствительность квазилокальных (т.е. найденных в физически предопределенном районе области определения параметров \hat{P} методом SLS) и глобальных оптимумов, найденных методом SCE. Нетрудно заметить, что «квазилокальные» параметры значительно стабильнее «глобальных». На рисунке 5 приведен пример ситуации, когда «глобальный» алгоритм SCE «не заметил» очевидного и весьма глубокого оптимума, находящегося рядом с предопределенным регионом поиска, поскольку алгоритм SCE при технически целесообразных настройках поиска не может исследовать точки в «углах» $(N + 1)$ -мерного параллелепипеда. Действительно, при оптимизации, например, 11 параметров и при лимитирующем числе запусков модели в 30 тысяч раз, средний шаг оптимизации чуть больше 1,5: $N_{FE} = \sqrt[m+1]{(m+1)} = 30\,000$, $(m+1) = 2,55$ и $m = 1,55$. Это говорит о том, что «глобальность» алгоритма SCE, подтверждаемая более широкой полосой поиска, показанной на рисунке 6, достигается за счет существенного увеличения шага оптимизации и, следовательно, огрубления поиска.

Выполненные численные эксперименты по моделированию стока с водосборов, расположенных в штатах Техас (США) и Новый Южный Уэльс (Австралия) показали, что при калибровке модели «Сакраменто» метод глобальной оптимизации в большинстве случаев позволяет получить меньшие (т.е. лучшие) значения целевой функции MSOF, однако проверка найденных параметров на независимом материале (т.е. в режиме прогнозирования стока) показывает бóльшую (более 60% случаев) эффективность параметров, определенных в физически предопределенном районе 12-мерной области определения параметров методом SLS. Также было установлено, что в отдельных случаях метод глобальной оптимизации не позволяет обнаружить более глубокие и, что самое главное, физически корректные оптимумы. Учитывая тот факт, что калибровка модели при помощи метода SCE занимает от 70 до 500 раз больше времени, чем при использовании метода SLS и более высокую эф-

фективность последнего при прогнозировании паводков, этот алгоритм был признан более подходящим для автоматической калибровки оперативных гидрологических моделей.

В Главе 4 приведен анализ существующих методов линейной и нелинейной оптимизации, которые используются для калибровки гидрологических моделей со сравнительно большим числом взаимозависимых параметров, частным случаем которых являются модели с независимыми параметрами или с небольшим числом взаимозависимых параметров. В частности, проанализированы причины низкой эффективности таких наиболее популярных в гидрологии методов линейной и нелинейной оптимизации как метод Розенброка, SCE, метод градиентной оптимизации, метод координатного спуска и некоторых других. Как уже отмечалось выше, многомерные поверхности целевых функций для таких моделей являются крайне нерегулярными и невыпуклыми. Частичное сглаживание достигается за счет перехода к мультимасштабным целевым функциям группы MSOF, но этого может оказаться недостаточно, поскольку поверхность целевой функции обладает определенной нестабильностью во времени.

Дополнительное сглаживание, достигаемое в результате применения F-преобразования, позволяет получать более стабильные параметры (следствием применения которых является повышение практической уверенности в прогнозах), но не решает проблему невыпуклости многомерных поверхностей кардинальным образом, поэтому существующие оптимизационные процедуры, разработанные для поиска оптимумов выпуклых поверхностей (типа «nut-shell» – от англ. «ореховая скорлупа»). Что же касается метода глобальной оптимизации (SCE), то он физически не может «справиться» со значительно невыпуклыми и нерегулярными поверхностями; это утверждение подтверждается в диссертации как теоретическими выкладками, так и многочисленными экспериментами (пример результатов этих экспериментов приведен в Таблице 1 и на рисунке 5).

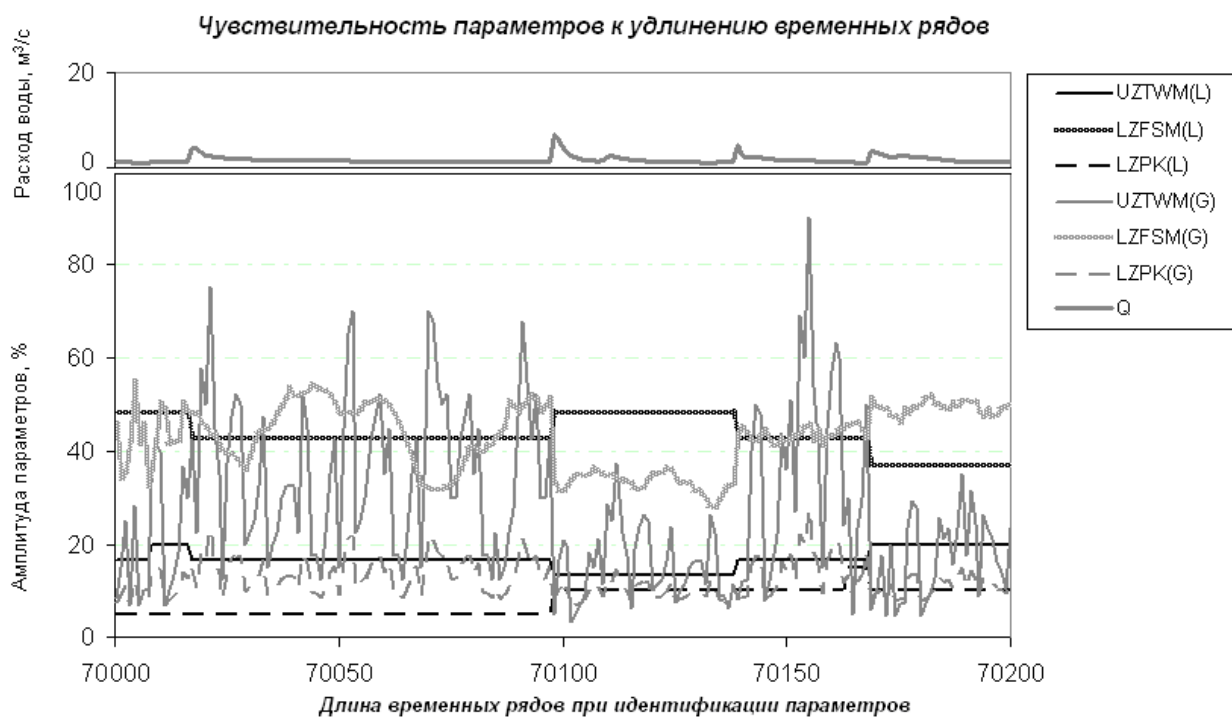


Рисунок 4 — Устойчивость во времени параметров модели «Сакраменто», определенных методами «глобальной» и квазилокальной оптимизации (буква «L» обозначает параметры, оптимизированные квазилокально, а «G» – глобально).

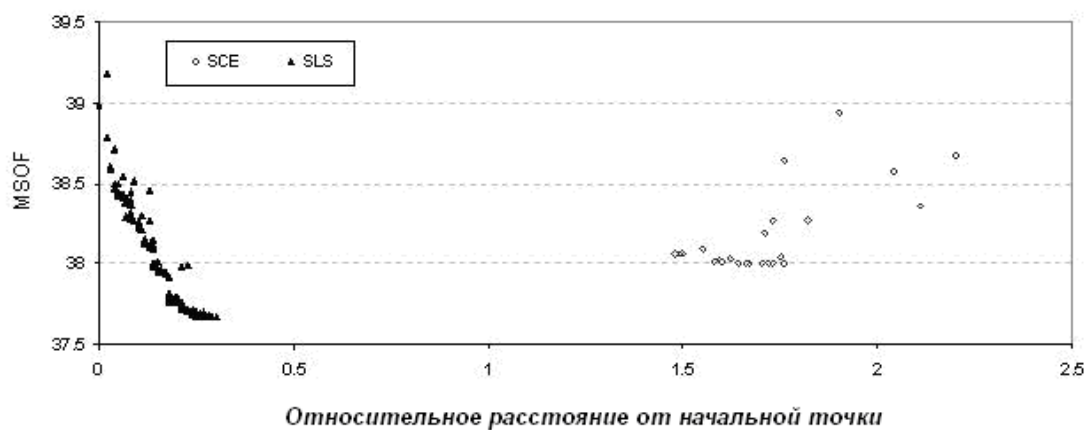


Рисунок 5 — Пример, иллюстрирующий тот факт, что «глобальность» оптимизации достигается в ущерб детальности поиска.

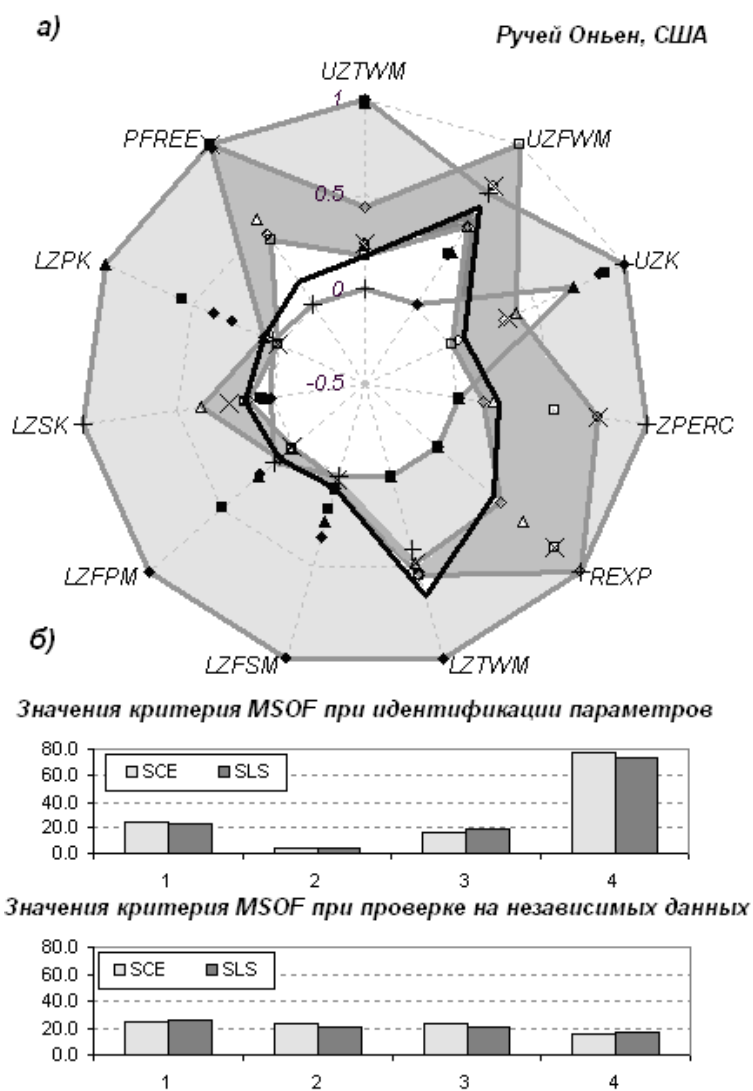


Рисунок 6 — (а) «Роза параметров», (б) результаты параметризации и верификации.

В качестве альтернативы существующим оптимизационным процедурам предлагаются уже упомянутые методы нелинейной оптимизации группы SLS, которые были разработаны таким образом, чтобы свойства многомерных поверхностей целевых функций были учтены наилучшим образом. Базовый оптимизационный алгоритм SLS, представляющий собой алгоритм «шаблонной» оптимизации, производит отыскание оптимумов в направлении уменьшения значения целевой функции, последовательно продвигаясь вдоль направления каждого параметра на определенное количество шагов (в нашем примере — на 1 шаг). Процедура SLS состоит из следующих операций:

1. Старт заключается в оценивании (вычислении) значения целевой функции $J(P_a)$, соответствующего predetermined набору (вектору) параметров P_a (здесь и в тексте диссертации символ вектора опускается);
2. Значения всех параметров, кроме первого, фиксируются. Значение первого — увеличивается или уменьшается на один шаг, при этом определяется направление понижения значения целевой функции J ;
3. Первому параметру присваивается новое значение, соответствующее уменьшенной величине J (или сохраняется старое, если целевая функция не изменилась), затем это значение фиксируется, и операция, описанная в п.1, выполняется в отношении второго параметра;
4. Операция из п. 3 повторяется в отношении всех остальных параметров, пока все параметры не будут испытаны и, если нужно, исправлены;
5. Операции из пп. 2–4 выполняются до тех пор, пока уменьшение значения J не прекратится. Следует подчеркнуть, что многочисленные эксперименты с моделью «Сакраменто», проведенные соискателем в 2002–2006 гг. позволили упростить базовый алгоритм следующим образом: если значение параметра не меняется на протяжении трех последовательных итераций (то есть трех последовательных корректировок всех оставшихся параметров), этот параметр исключается из дальнейшего рассмотрения (это объясняется локальной ортогональностью параметров вблизи оптимума, вызванной дискретностью поиска). Это упрощение позволяет сократить число запусков модели примерно в 4 раза. Набор наименее чувствительных параметров может изменяться от водосбора к водосбору и во времени и отражает взаимозависимость параметров, которая также изменяется во времени и пространстве.

Если шаг изменения параметров чересчур велик, алгоритм SLS может пропустить минимум (который мог бы быть обнаружен при более мелком шаге). В ходе выполнения экспериментов было установлено, что размер шага, равный 1÷10% амплитуды параметра, является вполне универсальным, хотя в некоторых случаях было бы интересно оптимизировать и его. Условие

остановки оптимизации — исключение всех параметров. Результаты параметризации модели «Сакраменто» на основе метода SLS с постоянным шагом оптимизации равным 5% показаны в работе [8].

Таблица 1 – Сравнение значений целевой функции MSOF при калибровке модели «Сакраменто» (за 4 года) и проверке найденных параметров на независимом материале (за 1 год)

Водоток*	Метод SCE					Метод SLS				
	Годы, использованные для оценивания MSOF									
	1	2	3	4	5	1	2	3	4	5
<i>Калибровка с использованием данных за 4 года (кроме указанного в соответствующей колонке)</i>										
GBHT2	14.1**	14.0	13.8	14.3	13.8	14.3	14.4	14.0	14.6	14.1
GETT2	18.3	18.6	18.9	12.6	18.3	18.8	18.9	19.5	12.9	18.8
HBMT2	31.9	33.5	32.9	33.3	32.1	33.4	35.4	34.8	35.0	33.0
HNTT2	36.8	38.8	28.3	34.2	36.5	36.9	38.8	28.6	34.5	36.7
JTBT2	11.7	12.6	7.21	15.7	13.2	12.6	12.2	7.15	15.6	13.8
KNLT2	15.0	18.7	18.0	18.7	10.9	17.3	20.1	19.8	19.7	11.2
LYNT2	12.7	12.8	12.3	12.2	8.54	12.7	13.0	12.4	12.3	8.69
MTPT2	38.0	41.7	41.3	40.0	38.0	37.9	41.5	41.3	40.1	37.9
<i>Проверка на независимом материале (за год, указанный в соответствующей колонке)</i>										
GBHT2	13.0	14.6	15.4	10.3	15.0	14.8	14.3	15.7	11.4	15.4
GETT2	14.2	9.73	3.57	27.7	13.9	14.1	8.71	3.50	26.2	13.0
HBMT2	29.9	27.9	25.1	21.5	35.9	27.0	34.6	27.6	25.2	47.1
HNTT2	33.3	4.81	66.1	47.4	32.0	32.0	4.51	66.3	44.1	32.0
JTBT2	12.4	4.32	25.9	9.59	26.2	4.96	3.79	24.7	6.47	17.6
KNLT2	31.9	4.38	13.7	18.0	47.9	28.5	11.1	10.8	15.3	43.1
LYNT2	11.4	5.89	11.0	11.4	36.9	11.8	4.92	10.3	11.1	37.3
MTPT2	45.1	16.2	20.9	34.2	52.4	45.4	14.5	19.6	33.7	52.0

* Полное название водотоков указано в Таблице 2 и тексте диссертации

** Меньшее значение показано курсивом

Алгоритм SLS-F (SLS плюс F-робастность) — это еще один полезный инструмент для калибровки оперативных гидрологических моделей. Его не следует рассматривать как альтернативу методу SLS-2L, поскольку эти два подхода предназначены для решения разных проблем. Задача метода SLS-F — поиск наиболее надежных и стабильных во времени параметров. Данное исследование носит пилотный характер, поэтому метод SLS-F использовался при $s=1$ (см. уравнение (12)). Это объясняется тем, что «обычный» метод SLS, в любом случае, предполагает расчет целевой функции в нескольких точках вокруг исследуемого набора параметров. При использовании метода SLS-F (при $s=1$) такой расчет всегда осуществляется ровно в трех точках (за исключением самой первой итерации, которая выполняется в обычном режиме). Общее число запусков модели при этом может возрасти в 5–10 раз, что все равно меньше, чем требуется для «глобального» алгоритма SCE. Увеличение числа шагов осреднения s до 2 ведет к дальнейшему увеличению числа запусков модели (до трех порядков) и, следовательно, в таком случае на оптимизацию расходуется ресурс процессора, соизмеримый с затрачиваемыми при использовании метода «глобальной» оптимизации SCE. Однако практическое использование метода SLS-F все равно имеет смысл, поскольку при такой же загрузке процессора можно получить более стабильные и надежные параметры, чем «глобально» оптимизированные. Примеры результатов экспериментов по автоматическому прогнозированию стока ряда американских и австралийских водосборов представлены в Таблице 2.

Алгоритм SLS-2L (2-Loop Stepwise Line Search — пошаговая линейная оптимизация, 2 цикла) предусматривает замену оригинальной $(N+1)$ -мерной поверхности целевой функции (для модели «Сакраменто» $N=11+16$) более гладкой поверхностью меньшей размерности. Сглаживающий эффект достигается уменьшением числа параметров N .

Таблица 2 – Сравнение результатов оптимизации различными способами на основе SLS

Водоток–створ	Код	Площадь водосбора, км ²	Нач. значение критерия MSOF (метод Корня)	Окончательное значение MSOF				
				SCE	SLS	SLS-F	SLS-2L	SLS-E
Верховые водосборы, находящиеся в штате Техас (США)*								
Onion Creek –Austin	ATIT2	844	23.21	19.36	20.84	20.85	24.01	16.24
Denton Creek – Justin	DCJT2	1039	18.47	16.13	16.57	16.57	18.88	14.99
Greens Bayou – Houston	GBHT2	137	13.82	11.35	11.65	11.66	14.12	9.51
South Fork San Gabriel – Georgetown	GETT2	334	17.39	16.22	16.52	16.54	17.32	16.03
Cowleech Creek – Greenville	GNVT2	212	16.89	14.39	14.60	14.60	17.90	11.72
Brays Bayou – Houston	HBMT2	246	35.69	27.02	28.52	28.53	42.48	24.18
Guadalupe River – Hunt	HNTT2	769	39.50	30.99	31.01	31.02	37.00	28.12
Double Mount Fork – Justiceburg	JTBT2	945	13.73	12.19	12.86	12.89	15.97	10.67
Sandy Creek – Kingsland	KNLT2	904	18.38	11.55	13.67	13.67	13.88	9.66
Davidson Creek – Lyons	LYNT2	508	10.51	10.22	10.37	10.37	10.41	9.10
East Fork Trinity – McKinney	MCKT2	427	16.84	13.87	14.18	14.19	15.31	12.42
Bedias Creek – Madisonville	MDST2	870	33.92	25.79	28.56	28.56	32.50	23.42
Midfield – Tres Palacios	MTPT2	435	35.43	<u>33.92</u>	<u>33.83</u>	33.83	34.00	29.45
Cowhouse Creek – Pidcoke	PICT2	1178	38.99	<u>38.00</u>	<u>37.68</u>	37.70	38.12	22.90
Navidad River – Sublime	SBMT2	896	56.66	53.92	54.57	54.57	55.73	48.83
Верховые водосборы, находящиеся в штате Новый Южный Уэльс**								
Kyeabma Creek – Book Book	KCBV1	145	3.55	0.926	0.914	0.914	1.472	0.798
Kyeamba Creek – Lady Smith	KCLS1	530	3.31	1.140	1.140	1.140	1.272	0.824
Hillis Creek – Mount Adrah	HCMA1	568	2.99	0.991	0.988	0.988	1.045	0.744
Billabong Creek – Aberfeldy	BCAB1	331	1.90	0.786	0.786	0.787	0.800	0.523

* Гидрологические посты, находящиеся в ведении регионального речного прогностического центра Национальной Службы Погоды США (NWS West Gulf RFC)

** Гидрологические посты, находящиеся в ведении Департамента природных ресурсов штата Новый Южный Уэльс (Австралия)

В нашем случае число идентифицируемых параметров уменьшается с 11 (число параметров модели «Сакраменто») до 7 (число гидравлических параметров почвогрунтов). 8-мерная поверхность целевой функции (образуемая семью гидравлическими параметрами и минимизируемым функционалом J) является более гладкой, чем 12-мерная или (17-мерная) просто по определению, из-за меньшей размерности. Обнаружить оптимумы такой поверхности при помощи SLS весьма легко. Если класс доминирующих почвогрунтов известен, именно их гидравлические параметры используются для расчета стартовой точки. Если этот класс неизвестен, то каждый из 12 известных наборов параметров, (соответствующих перечисленным в Табл. 10 диссертационной работы классам почвогрунтов), может служить в качестве исходной точки для определения региона поиска. Эта процедура является первым циклом оптимизационной процедуры, который позволяет получить 7 оптимальных гидравлических параметров почвогрунтов, по которым рассчитывается стартовая точка для второго цикла (определяемая 11–16 параметрами модели «Сакраменто»). Таким образом, если детальные данные (такие, например, как данные цифровых карт STATSGO и ASRIS) недоступны, оптимизационная процедура начинается с отыскания наиболее подходящего класса доминирующих почвогрунтов и продолжается идентификацией параметров модели «Сакраменто» в обычном порядке (при помощи алгоритма SLS). Это свойство делает алгоритм SLS-2L пригодным для параметризации модели в развивающихся странах и других регионах со скудными или низкокачественными данными. Результаты применения процедуры SLS-2L представлены в Таблице 2.

Алгоритм SLS-E заключается в одновременном генерировании ансамблей «входа» модели «Сакраменто», осуществляемом при помощи датчика случайных чисел, и калибровке модели для каждого элемента ансамбля при помощи метода SLS. Этот подход позволяет не только получать несмещенные значения параметров, но и выполнять постобработку уже выпущенных прогнозов. Метод SLS-E состоит из следующих процедур:

1. Выбрать период для калибровки модели (например, в проведенных численных экспериментах по моделированию стока ручья Куямба в створе Бук-Бук (Kueamba Creek – Book Book) исходные данные были использованы так: (1) «разогрев» модели с 1 января 1995 г. до 31 декабря 1997 г., (2) параметризация модели «Сакраменто» — с 1 января 1998 г. до 31 декабря 2004 г. и, наконец, (3) проверка (верификация) и постобработка прогнозов — с 1 января 2005 г. до 31 декабря 2006 г.);
2. Выполнить параметризацию модели в обычном порядке (например, используя методы, основанные на SLS);
3. Определить стохастические характеристики (например, начальные моменты m_1^t, m_2^t, m_3^t) всех k «входов» I_k и «выходов» O_k модели на каждый момент времени t интервала, используемого для постобработки (post-processing window). Порядок выполнения этой процедуры зависит от специфики решаемой задачи. В частности, непосредственное вычисление моментов приемлемо только в том случае, если изучаемое и моделируемое распределение строго стационарно (такие ситуации встречаются, например, при долгосрочном прогнозировании или моделировании при помощи физически обоснованных динамических моделей). При краткосрочном прогнозировании моменты целесообразно определять в соответствии с пространственно-временным распределением осадков, почвенной влаги, испаряемости, расходов или уровней воды. Другими словами, оценки трех моментов могут быть рассчитаны или измерены как угодно. В нашем случае оценивались только стохастические характеристики осадков (путем анализа их пространственно-временного распределения) и расходов воды (путем анализа их временной изменчивости, рассчитанной по данным автоматизированных наблюдений за уровнем в течение 3-дневного интервала, окружающего полдень суток, к которым относятся найденные моменты, и учитывая погрешность расчета расходов по уровням).
4. Для удобства и упрощения расчета, которое почти не влияет на точность прогноза, выбрать какое-нибудь теоретическое распределение, которое

адекватно описывает остаточные отклонения на «входах» (в большинстве случаев трехпараметрическое гамма-распределение может быть взято для использования по умолчанию);

5. Выбрать генератор случайных чисел, соответствующий выбранному распределению и определенным моментам $(I_k(t); m_{2,k}; m_{3,k})$ и выполнить генерирование некоторого числа ансамблей для каждого из «входов» (как это делается в классическом методе Монте-Карло). Для каждого из ансамблей выполняется квазилокальная калибровка модели методом SLS. (В нашем случае производилось генерирование 125 ансамблей осадков, что вместе с рекалибровкой модели занимает от 3 ч 50 мин до 6 часов 40 мин);
6. Определить, какая из траекторий (или какая комбинация траекторий при $k > 1$) и какой набор параметров позволяют получить «выход», наилучшим образом совпадающий с «целью» (то есть измеренным, прототипным или желательным «выходом»). Степень отклонения выражается при помощи целевой функции J , которая может быть как детерминистической, так и вероятностной. Идеальным попаданием в «цель» может считаться, например, нулевое отклонение прогнозируемой и фактической моды или полное совпадение прогнозируемого распределения и распределения «цели». В нашем случае использовался критерий MSOF;
7. Оценить систематическое смещение выбранной траектории от исходного ряда внутри сравнительно небольшого интервала времени, предшествующего времени прогноза — «окна» постобработки (post-processing window). Для этой цели можно либо просто рассчитать среднее арифметическое отклонение Δ_k , либо определить второй и третий моменты остаточных отклонений (исходного ряда от выбранной траектории), а затем рассчитать их систематическое смещение моды (т.е. наиболее вероятное отклонение r_k элементов исходного ряда от выбранной «правильной» траектории; для трехпараметрического гамма-распределения $r_k = m_{3,k} / 2m_{2,k}$).

8. Применить Δ_k или r_k в качестве поправки для «входа» I_k , определить исправленный «вход» $\hat{F}_k(t) = I_k(t) - r_k$ и перезапустить модель для расчета расхода воды для следующего временного шага.

Выполнив все эти шаги, можно получить обновленные параметры и исправленные «входные» данные (все или только часть) для интервала времени в пределах окна постобработки. Это позволяет интегрально учесть все виды неопределенности: временную и пространственную изменчивость исходных данных, погрешность измерений, несовершенство модели. В итоге дисперсия погрешности прогнозов существенно уменьшается, а моделируемая траектория (гидрограф) остается в пределах доверительного интервала. Необходимо заметить, что при автоматической калибровке гидрологических моделей, используемых для прогнозирования стока большого числа водотоков, описанная процедура может быть значительно упрощена: метод SLS-E используется для идентификации несмещенных параметров, которым соответствует наименьшее значение целевой функции J (например, MSOF), а наилучший «вход» модели и его стохастические характеристики не идентифицируются. Как показали численные эксперименты (Таблица 2), метод SLS-E позволяет получить наилучшие (наименьшие) значения целевой функции MSOF, однако для использования этого метода необходимо значительно большее время, чем для метода SLS (и соизмеримое с временем, затрачиваемым на оптимизацию параметров при помощи метода SCE). Несмотря на это, метод SLS-E является наиболее эффективным способом идентификации несмещенных параметров модели, которые обеспечивают самую высокую точность прогнозирования катастрофических паводков. Пример расчета стока р. Аlikоновки в г. Кисловодске, выполненный для летне-осеннего сезона 2008 г. при предположении максимального влагонасыщения почвы 01.06.2008 и рекалибровкой модели для 125 ансамблей осадков показан на рисунке 6.

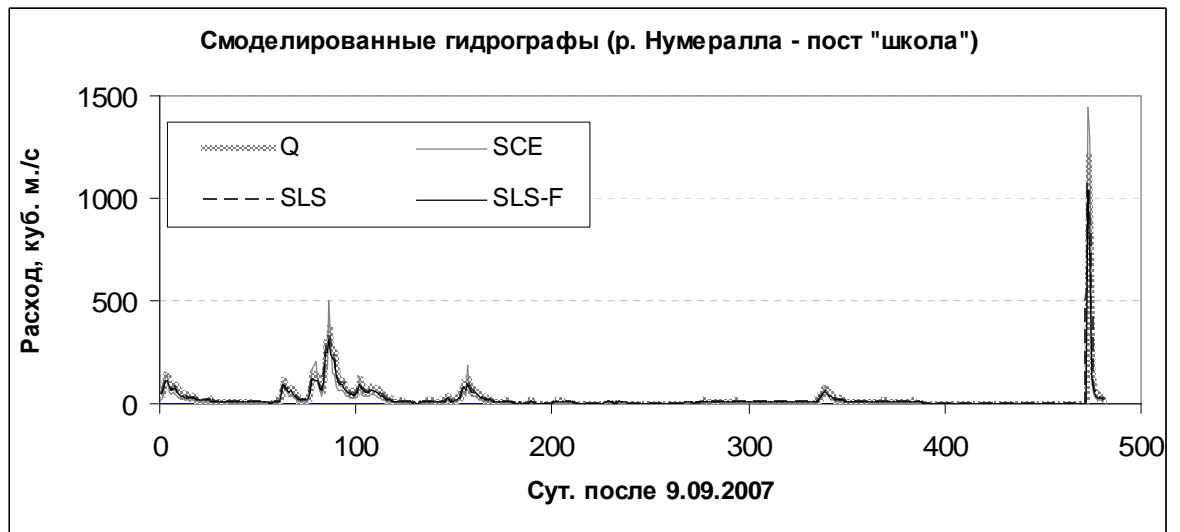
Завершает Главу 4 описание оптимизационных процедур группы SLS третьего поколения SLS-2LF, SLS-EF, SLS-2LE, которые образуются путем совмещения алгоритмических процедур SLS-E, SLS-F и SLS-2L, а также про-

цедуры четвертого поколения SLS-2LEF. В данной работе эти процедуры (за исключением алгоритма SLS-2LF, показавшего высокую эффективность при определении трансферабельных параметров, что подтверждается данными Таблицы 3) практически не использовались, поскольку они требуют больших процессорных ресурсов. Однако есть все основания полагать, что по мере развития вычислительной техники они станут эффективными инструментами калибровки гидрологических моделей (рисунки 7 и 8).

В Главе 5 рассмотрены различные аспекты практического применения разработанных методов и подходов, которые были использованы, в частности, при создании учебно-оперативной автоматической системы прогнозирования паводков «SLS+».

Таблица 3 — Трансферабельность параметров, полученных в результате оптимизации методом SLS-2LF

Водосбор, использованный для калибровки	Водосборы, использованные для проверки			
	1	2	3	4
	Исходные значения MSOF			
	3.55	3.31	2.99	1.90
Kyeabma Creek – Book Book	1.472	1.320	1.659	1.118
Kyeamba Creek – Lady Smith	1.475	1.272	1.734	1.238
Hillis Creek – Mount Adrah	1.801	1.423	1.045	0.999
Billabong Creek – Aberfeldy	1.656	1.512	1.307	0.800



б)

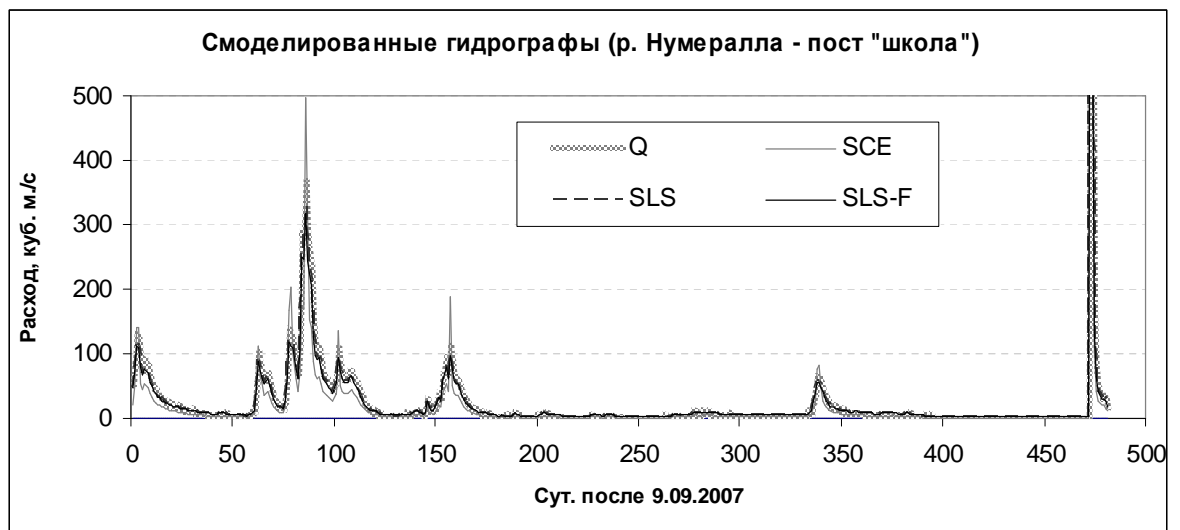


Рисунок 7 — Гидрографы, построенные при помощи модели «Сакраменто», которая была откалибрована различными способами: (а) оригинальный масштаб; (б) увеличенный масштаб.

Основными функциональными элементами разработанной автоматизированной системы прогнозирования паводков «SLS+» являются: 1) концептуальная гидрологическая модель «Сакраменто», которая может быть заменена на любую другую гидрологическую модель (например, модель Гидрометцентра, разработанную Бельчиковым и Корнем, или модель Корня и Кучмента); 2) автоматические оптимизационные процедуры группы SLS, позволяющие выполнять калибровку модели при различной обеспеченности исходными

гидрометеорологическими данными; 3) статическая база данных, содержащая неизменяемую информацию о водосборах; 4) динамическая база данных «INWADA»; 5) блок усвоения данных и 6) блок постобработки выпускаемых прогнозов.

В настоящее время АСП «SLS+» реализуется в двух вариантах: 1) с WINDOWS-совместимым пользовательским интерфейсом, предназначенным для моделирования гидрографа отдельно взятого водотока, и 2) DOS/UNIX/LINUX-совместимая интерактивная программа, которая позволяет выполнять моделирование стока большого числа малых и средних водотоков (например, относящихся к водосбору одной большой реки или находящихся в некотором регионе). Таким образом, в своем окончательном виде первое поколение АСП «SLS+» представляет собой самодостаточный и полностью готовый для практического использования инструмент для автоматического моделирования дождевых паводков на малых и средних водотоках, который может быть использован как в учебных целях студентами гидрометеорологических вузов и университетов, так и для оперативного прогнозирования паводков подразделениями гидрометслужбы (в случае подключения АСП к базе данных Росгидромета), министерства по чрезвычайным ситуациям и другими заинтересованными организациями как внутри Российской Федерации, так и за ее пределами.

Помимо АСП «SLS+», при выполнении описанных в данной диссертационной работе исследований был создан аппаратно-программный комплекс «INWADA», предназначенный для автоматического поиска, копирования и архивирования данных оперативных гидрометеорологических наблюдений (в настоящее время АПК «INWADA» работает в пилотном режиме, производя автоматическое архивирование данных гидрометеорологических наблюдений, производимых в США и Австралии; для этой цели было разработано специальное программное обеспечение, необходимое для управления и эксплуатации динамической базы данных АПК «INWADA»).

В Заключении рассмотрены перспективы развития автоматизированного прогнозирования ОГЯ, а также даны рекомендации по повышению точности прогнозов паводков в Российской Федерации, наиболее важными из которых являются следующие:

1. В районах РФ, подверженных наводнениям в результате прохождения паводков, необходима организация автоматических наблюдений за атмосферными осадками с периодичностью не менее 1 часа;
2. В районах, где паводки вызывают значительный материальный ущерб, целесообразно автоматическое наблюдение за расходом воды с периодичностью не менее 1 часа;
3. В районах РФ с высокой интенсивностью формирования паводков (например, вызванных сильными ливнями или обусловленных рельефом местности) целесообразно использование метеорологических радаров с радиусом действия не менее 250 км; для усвоения радарных данных должны быть организованы синхронные наземные наблюдения;
4. В районах с недостаточной плотностью метеорологических станций и гидрологических постов рекомендуется разработка системы вероятностного прогнозирования паводков по спутниковым данным о содержании влаги в верхнем слое почвы и о перемещении облачных полей, способных вызывать конвективные осадки;
5. В индустриальных и сельскохозяйственных районах оценивание качества прогнозов должно производиться только по матрицам потерь. В связи с этим необходимо обязать органы местной исполнительной власти и заинтересованное руководство промышленных предприятий предоставить всю необходимую информацию для составления таких матриц. В то же время, использование матриц потерь для калибровки гидрологических моделей недопустимо (для этой цели следует использовать мультимасштабные целевые функции);
6. Необходима организация непрерывного фонового прогнозирования речного стока по всей территории РФ в автоматическом режиме. Для этого

необходимо интегрировать в единую прогностическую систему данные наземных, радарных и спутниковых наблюдений, а также выходные данные климатических моделей. В случае выявления зон с повышенной вероятностью возникновения паводков, целесообразно выпускать уточненный прогноз на основе более подробной информации.

В разделе «Выводы» обобщены основные результаты выполненных исследований.

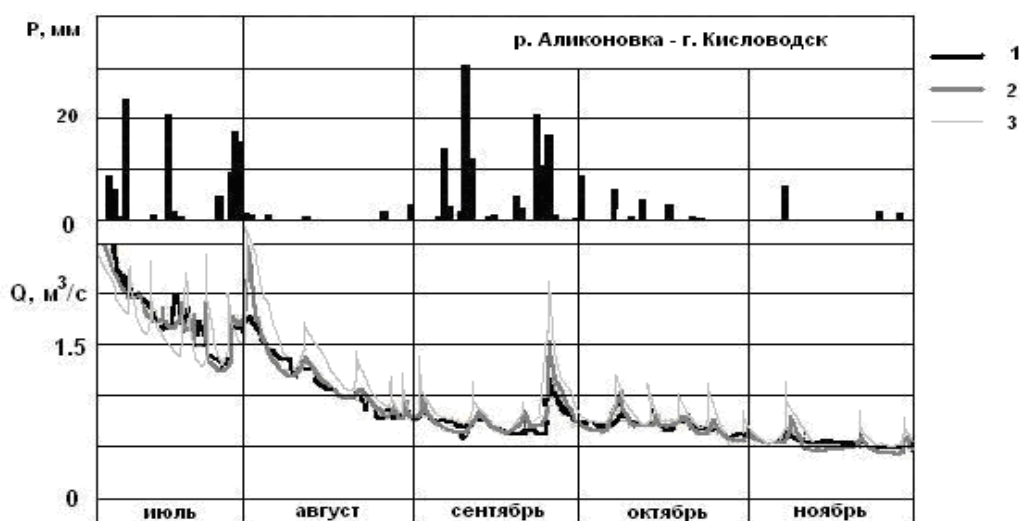


Рисунок 8 — Пример гидрографа р. Аlikоновки (в створе г. Кисловодска, 2008 г.), рассчитанного только для теплого периода года (в начале расчетного периода почва считается максимально насыщенной влагой). Модель откалибрована за теплый период 2007 г. методом SLS-E (1 — фактический гидрограф; 2 — смоделированный гидрораф (125 ансамблей осадков), 3 — гидрограф без уточнения осадков).

ОСНОВНЫЕ РЕЗУЛЬТАТЫ И ВЫВОДЫ

В результате проведенных исследований решена крупная научно-техническая проблема, заключающаяся в получении возможности моделирования и прогнозирования паводочного стока с малоизученных и неизученных малых и средних водосборов. Успешное решение этой проблемы имеет важ-

ное хозяйственное и экономическое значение для повышения точности прогнозов катастрофических паводков и половодий. Поставленная цель достигнута за счет создания, развития и совершенствования комплекса научных методов автоматической калибровки гидрологических моделей, основанных на выявленных особенностях поведения многомерных поверхностей целевых функций, свойственных для гидрологических моделей с взаимозависимыми параметрами. В результате проделанной работы было установлено, что:

1. Метод SLS является наиболее подходящим в случаях, если априорная (начальная) точка при квазилокальной оптимизации задана корректно, а водосбор достаточно хорошо освещен гидрометеорологическими наблюдениями; во всех остальных случаях рекомендуется использование модификаций метода SLS – алгоритмов второго поколения SLS-E, SLS-2L, SLS-F и алгоритма третьего поколения SLS-2LF (остальные алгоритмы могут быть использованы по мере совершенствования вычислительной техники);
2. Метод SLS-2L рекомендуется для водосборов, на которых имеется недостаточность данных о почвогрунтах и их гидравлических свойствах;
3. Метод SLS-F рекомендуется в случаях, когда обучающие выборки сравнительно коротки, в результате чего положение поверхности целевой функции (и, следовательно, локальных оптимумов) нестабильна во времени;
4. Метод SLS-E может быть использован при низкой плотности метеорологических станций (осадкомеров). Метод SLS-E позволяет ликвидировать смещенность параметров, вызванную неравномерностью выпадения осадков, поэтому он может быть использован в качестве инструмента для постобработки смоделированного гидрографа; предметом стохастического генерирования (в виде ансамблей) может выступить любой или все элементы «входа» модели;
5. Метод SLS-2LF позволяет получить трансферабольные параметры, которые могут быть использованы на неизученных водосборах;
6. Комбинации методов (SLS-2LF, SLS-2LE, SLS-EF и SLS-2LEF) являются чрезвычайно ресурсоемкими и при современном уровне развития вычислительной техники в оперативном режиме использоваться не могут; по мере

увеличения ресурсов процессора целесообразно в первую очередь использовать оптимизационные алгоритмы группы SLS-E (т.е. SLS-E, SLS-2LE, SLS-2EF и SLS-2LEF), как обеспечивающие несмещенные параметры, что исключительно важно для расчета начальных условий перед наступлением очередного паводка;

7. При моделировании гидрографа с 1-часовой дискретностью все рассмотренные методы расчета единичного гидрографа показали соизмеримое улучшение начального значения целевой функции (12,97% – по площадям стекания, 11,91% – по синтетическому единичному гидрографу, 9,42% – по биномиальной кривой); при моделировании гидрографа с 24-часовой дискретностью все рассмотренные методы расчета единичного гидрографа также показали соизмеримое улучшение начального значения целевой функции (43,43% – по площадям стекания, 40,29% – по синтетическому единичному гидрографу, 39,37% – по биномиальной кривой);

8. Во всех случаях и на всех водотоках наилучший результат соответствует способу расчета по площадям стекания, вторым по результативности является способ уточнения синтетического гидрографа, а худшим – аппроксимация гидрографа кривой биномиального распределения, однако различия между полученными результатами незначительна; при моделировании суточного стока в регионах с недостаточным количеством исходных данных экономически целесообразно использовать аппроксимацию единичного гидрографа кривой биномиального распределения, поскольку этот способ не требует никаких дополнительных затрат; при наличии карт, позволяющих построение схем стекания, целесообразно использовать более точный метод построения гидрографа по площадям стекания;

9. Наибольшая степень улучшения значения целевой функции при моделировании суточного гидрографа объясняется эффектом осреднения, поскольку внутрисуточное распределение стока в этом случае играет меньшую роль;

10. В связи с тем, что на практике возможно возникновение необходимости моделирования стока с различными временной дискретностью и обеспеченно-

стью исходными данными, в АСП «SLS+» должно быть предусмотрено возможность использования всех рассмотренных методов расчета единичного гидрографа; по умолчанию целесообразно использовать биномиальную кривую;

11. Оптимизационная процедура SLS вполне робастна по отношению к шагу оптимизации; во всех рассмотренных случаях дискретность от 5 до 100 интервалов позволила получить соизмеримые результаты; в случае, если различные размеры шага оптимизации приводят к сходным решениям, целесообразно назначать наиболее экономичную дискретность (то есть такую, при которой нахождение оптимума затрачиваются наименьшие ресурсы процессора, выражаемые в числе пусков модели). Наиболее экономически выгодной является дискретность оптимизации от 5 до 10 интервалов, однако в ряде случаев для достижения приемлемого результата необходима более высокая дискретность – до 50 интервалов (т.е. шаг оптимизации равен 2% амплитуды каждого из параметров). Дискретность выше 100 интервалов (при которой шаг оптимизации равен 1% амплитуды) нецелесообразна, поскольку в этом случае поиск может остановиться слишком быстро (в незначительной «депрессии» вблизи исходной точки); в качестве дискретности оптимизации, принимаемой по умолчанию, можно рекомендовать дискретность 20, 33 или 50 интервалов, однако в некоторых случаях (при наличии экспериментального обоснования) она может быть уменьшена до 5 или 10 интервалов;

12. При калибровке модели «Сакраменто» различными способами выявлена повышенная чувствительность к длине обучающей выборки параметров, оптимизированных методом SCE, что подтверждает предварительные выводы о специфике формы дна поверхности целевой функции. Это означает, что метод SCE в данном случае неэффективен. «Случайный» характер изменения значений параметров, найденных методом SCE свидетельствует о значительной невыпуклости поверхности целевой функции и неустойчивости оптимумов;

13. Метод SLS и его F-робастная модификация SLS-F показали близкие значения оптимальных параметров; гидрографы, соответствующие этим парамет-

рам практически не отличаются. Значения критерия MSOF для методов SLS и SLS-F в целом сопоставимы, однако метод SLS-F менее экономичен (в смысле затрат ресурсов процессора), чем «обычный» SLS, но он позволяет получить более стабильные параметры. Он рекомендуется для использования в тех регионах, где имеются достаточно надежные данные о почвогрунтах;

14. Для работы АСП «SLS+» в режиме реального времени необходима разработка динамической базы данных; удобной альтернативой которой может быть прямое подключение пользователя к национальным базам данных;

15. Наиболее перспективными направлениями для дальнейших исследований являются: 1) прогнозирование наводнений, вызванных наложением паводков и половодий, 2) адаптация разработанных методов, алгоритмов и процедур к другим моделям (например, модели талого и дождевого стока) и 3) создание методологии выбора наиболее информативных частот целевой функции MISOF.

Представленные в работе методики, процедуры, алгоритмы и рекомендации использованы в учебном процессе ГОУ ВПО «Российский государственный гидрометеорологический университет». Кроме того, они использованы при разработке программного обеспечения автоматизированной системы прогнозирования паводков «SLS+», предназначенной для экспорта в страны Юго-Восточной Азии, Африки, Центральной и Южной Америки. Метод автоматической калибровки SLS используется Национальной Службой Погоды США (U.S. National Weather Service, NOAA) для параметризации модели «Сакраменто», применяемой для оперативного прогнозирования дождевых паводков. Сделанные в результате проделанной работы выводы и рекомендации могут быть использованы для усовершенствования системы прогнозирования катастрофических паводков и половодий в рамках программы «Модернизация и техническое перевооружение учреждений и организаций Росгидромета».

Заключительная часть работы выполнена в рамках аналитической ведомственной целевой программы Министерства образования и науки Российской Федерации «Развитие научного потенциала высшей школы (2009-2010)»

(проект 2.1.1/3962 «Автоматизированная система краткосрочного прогнозирования паводков и половодий»); отдельные элементы диссертационной работы выполнены при поддержке Технического Университета г. Дельфт (Нидерланды), Национального управления США по исследованию океана и атмосферы (NOAA), Национального аэрокосмического агентства США (NASA), Метеобюро Правительства Австралии и проекта eWater (Австралия). Разработанная АСП «SLS+» признана лучшим инновационным продуктом 2009 г. по направлению «рациональное природопользование», а возглавляемый соискателем авторский коллектив награжден за эту работу Дипломом Комитета по науке и высшей школе Правительства Санкт-Петербурга.

Основное содержание диссертации опубликовано в работах:

из перечня ведущих научных рецензируемых журналов и изданий, в которых по рекомендации Высшей аттестационной комиссии Министерства образования и науки Российской Федерации должны быть опубликованы основные научные результаты диссертации на соискание ученой степени доктора наук:

1. Кузьмин В. А. Отбор и параметризация прогностических моделей речного стока // Метеорология и гидрология, 2001, № 3, с. 85–90.
2. Кузьмин В. А. Краткосрочное прогнозирование катастрофических половодий и паводков // Метеорология и гидрология, 2001, № 6, с. 89–95.
3. Кузьмин В. А. Основные принципы автоматической калибровки многопараметрических моделей, используемых в оперативных системах прогнозирования дождевых паводков // Метеорология и гидрология, 2009, № 6, с. 92–104.
4. Кузьмин В. А. Алгоритмы автоматической калибровки многопараметрических моделей, используемых в оперативных системах прогнозирования паводков // Метеорология и гидрология, 2009, № 7, с. 74–85.

5. Кузьмин В. А., Заман А. Постобработка и корректировка прогнозов паводков, выпускаемых при помощи автоматизированных систем // Метеорология и гидрология, 2009, № 8, с. 80–90.
6. Кузьмин В. А. Фундаментальные основы автоматизированного прогнозирования катастрофических дождевых паводков // Естественные и технические науки, 2009, №6.
7. Кузьмин В. А. Аппаратно-программный комплекс «INWADA» по поиску, передаче и хранению гидрометеорологической информации в целях фоновое прогнозирование опасных гидрологических явлений // Естественные и технические науки, 2010, №1.
8. Kuzmin, V., Seo, D.-J., Koren, V. Fast and efficient optimization of hydrologic model parameters using a priori estimates and stepwise line search // Journal of Hydrology, vol. 353, iss. 1–2, 2008, pp. 109–128.

В других изданиях:

9. Кузьмин В. А. Краткосрочное прогнозирование весеннего стока неизученных рек. СПб, Изд. РГГМИ, 1997, 32 с.
10. Карасёв И. Ф., Кузьмин В. А. Мода и среднее как основные гидрологические характеристики // Труды Академии Водохозяйственных Наук, СПб, 1995.
11. Кузьмин В. А. Использование регрессионных уравнений для прогнозирования стока р. Волги // Моделирование и прогнозирование речного стока. СПб, Изд. РГГМУ, 1999, с. 54.
12. Кузьмин В. А. Краткосрочное прогнозирование катастрофических наводнений при недостаточных исходных данных // Новые технологии в гидрометеорологии. СПб, Изд. РГГМУ, 2001.
13. Кузьмин В. А. Унифицированная самообучающаяся процедура моделирования речного стока // Труды конференции «Современные проблемы стохастической гидрологии». Москва, ИВП, 2001, с. 207

14. Кузьмин В. А. Прогнозирование стока р. Оки на основе модели кинематической волны. Деп. в ВИНТИ 14.08.95, №2439–В95.
15. Кузьмин В. А. Применение множественной корреляции для прогнозирования уровня Чебоксарского водохранилища. Деп. в ВИНТИ 14.08.95, №2440–В95.
16. Кузьмин В. А. О выборе критерия оценивания методик гидрологического прогнозирования. Деп. в ВИНТИ 14.08.95, №2441–В95.
17. Кузьмин В. А. О принципах параметризации математических моделей в прикладной гидрологии. Деп. в ВИНТИ 02.07.98, №2069–В98.
18. Кузьмин В. А. Выделение идентифицируемых и инерционных компонент речного стока // Материалы Итоговой сессии ученого совета РГГМУ, 2001, с. 82–86.
19. Коваленко В. В., Лубяной А. В., Старостин В. Л., Кузьмин В. А., Вакс Ф. М., Подрядов Д. А. Применение динамических и стохастических моделей водных объектов в прогностических целях // Тезисы Международного Симпозиума «Расчеты речного стока», Санкт-Петербург, 1995.
20. Кузьмин В. А. Критерий отбора моделей при прогнозировании экстремальных половодий // Материалы Итоговой сессии ученого совета РГГМУ, 2000, с. 52.
21. Кузьмин В. А., Аксой Х., Определение робастных статистик низкого стока рек Турции // Материалы Итоговой сессии ученого совета РГГМУ, 2000, с. 73.
22. V. Kuzmin, P. van Gelder, H. Aksoy, I. Kucuk. Application of the stochastic self-training method for the modeling of extreme floods // The Extremes of the Extremes, IAHS Publ. No. 271, 2002, pp. 317–322.
23. Z. Zhang, M. Smith, V. Koren, S. Reed, F. Moreda, V. Kuzmin, R. Anderson. A study of the relationship between rainfall variability and the improvement of using a distributed model // Eds. Y. Chen, K. Takara, I. D. Cluckie, H. F. de Smedt, GIS and Remote Sensing in Hydrology, Water Resources and Environment, IAHS Publ. No. 289, 2004.

24. V. Kuzmin. Short-Term forecasting of flood transformation along floodplain // J. of Floodplain Management, vol.2, No.1, 2000, pp. 50–57.
25. V. Kuzmin. Forecasting of large rivers floods at absence of hydrometric data whereby robust self-training model // NATO ARW Stochastic models of hydrological processes and their applications to problems of environmental preservation, Moscow, November 23-27, 1998, p. 50–55.
26. V. Kuzmin, P. van Gelder. The principles of catastrophic flood forecasting // River Flood Defence, Eds. F. Toensmann and M. Koch, Vol. 2, pp. G137-G144, 2000, Kassel, Germany.
27. P. van Gelder, V. Kuzmin, P. Visser. Analysis and statistical forecasting of trends of hydrological processes in climate changes // River Flood Defence, Eds. F. Toensmann and M. Koch, Vol. 1, pp. D13-D22, 2000, Kassel, Germany.
28. M. Smith, V. Koren, Z. Zhang, S. Reed, D.-J. Seo, F. Moreda, V. Kuzmin, Z. Cui, R. Anderson. NOAA/NWS Distributed Hydrologic Modeling Research and Development // NOAA Technical Report NWS 45, April, 2004.
29. V. Kuzmin, A. Seed, J. Walker. Australian Government Bureau of Meteorology forecast and real-time observational hydrometeorological data for hydrologic forecasting // eWater CRC Technical Report, eWater CRC, Canberra, 2007.
30. D. Barrett, V. Kuzmin, J. Walker, T. McVicar, C. Draper. Improved stream flow forecasting by coupling satellite observations, in situ data and catchment models using data assimilation methods // eWater CRC Technical Report, 2008, 77 p.
31. T. McVicar, V. Kuzmin, D. Barrett, and J. Walker. eWater Research highlights. D1: Predicting and Forecasting Streamflow // eWater CRC, 2008.
32. R. M. Anderson, V. Koren, S. Reed, M. Smith, and V. Kuzmin. Regionalization of Rainfall-Runoff Model Parameters // International Union of Geodesy and Geophysics, Sapporo, Japan, July 9, 2003.
33. V. Kuzmin. A stochastic filter for asymmetric hydrological distributions // Proc. Water Down Under 2008, Adelaide, April 14–17, 2008.

34. V. Kuzmin, N. Sanotskaya, I. Vinokurov. Automatic Calibration of the Sacramento Soil Moisture Accounting Model in Data Sparse Regions // Proc. the 2nd International Multidisciplinary Conference on Hydrology and Ecology (HydroECO 2009), Eds. J. Bruthans, K. Kovar, P. Nachtnebel, Vienna, Austria, April 20–23, 2009.
35. H. Aksoy, P. van Gelder, V. Kuzmin. Generation of the daily streamflow hydrograph // Proc. of the International Workshop on River Runoff: minima and maxima, , Eds. V. Kuzmin, P. van Gelder and C. Clark, June 6–8, 2001, Saint Petersburg, Russia.
36. V. Kuzmin and P. van Gelder. Identification of the Fokker-Plank-Kolmogorov equation: an advanced approach // Proc. of the International Workshop on River Runoff: minima and maxima // Eds. V. Kuzmin, P. van Gelder and C. Clark, June 6–8, 2001, Saint Petersburg, Russia.
37. V. Kuzmin. Stochastic analysis of remote sensed data in flood modelling // Proc. 5th International Workshop on Application of remote sensing in hydrology, Montpellier, France, October 2–5, 2001, p. 39.
38. V. Kuzmin. Extreme flood forecasting in developing countries // AGU 2001 SM, Boston, 2001.
39. V. Kuzmin. Stochastic forecasting of extreme flood transformation // Extreme Value Theory and Applications, December 10–15, 2001, Nordic Folk Academy, Goteborg, Sweden, p. 9.
40. V. Koren, S. Reed, M. Smith, Z. Zhang, D.-J. Seo, F. Moreda, and V. Kuzmin. Use of spatially variable data in river flood prediction // Proc. AGU-EGS-EUG Assembly, 2003, Nice, France.
41. V. Kuzmin, V. Koren, D.-J. Seo. Does the Global Minimum From Automatic Calibration Provide an 'Optimal' Solution? // Proc. American Geophysical Union Spring Meeting 2004, H41D–03.
42. V. Koren, D.-J. Seo, V. Kuzmin. A priori parameters, uncertainties, and calibration in watershed modeling // Proc. American Geophysical Union Spring Meeting 2005, H44B–01.

43. V. Kuzmin, J. Walker. A Stochastic Filter for Asymmetrically Distributed Streamflow // CAHMDA-III International Workshop, January 9–11, 2008, Melbourne, Australia.
44. V. Kuzmin. Use of the Sacramento soil moisture accounting model in areas with insufficient forcing data, EGU General Assembly 2009, April 19–23, 2009, Vienna, Austria.
45. S. Reed, V. Koren, M. Smith, Z. Zhang, F. Moreda, D.-J. Seo and DMIP Participants (incl. V. Kuzmin). Overall distributed model intercomparison project results // Journal of Hydrology, Vol. 298, Iss. 1–4 (Special Issue on The Distributed Model Intercomparison Project (DMIP)), 2004, pp. 27–60.
46. V. Kovalenko, A. Lubyanyoy, V. Starostin, V. Kuzmin, F. Vaks, D. Podryadov. Application of dynamic and stochastic models of water bodies for the forecast purposes // Runoff computations for water projects IHP–V. Technical document in Hydrology, No. 9, Paris, 1997.

Соискатель



(Кузьмин В. А.)
