

**Секция I. СОВЕРШЕНСТВОВАНИЕ МЕТОДОВ И МОДЕЛЕЙ  
ОЦЕНКИ И ПРОГНОЗИРОВАНИЯ РЕСУРСОВ И РЕЖИМА  
ПОВЕРХНОСТНЫХ И ПОДЗЕМНЫХ ВОД В УСЛОВИЯХ  
НАРАСТАЮЩЕГО АНТРОПОГЕННОГО ВОЗДЕЙСТВИЯ И  
НЕОПРЕДЕЛЕННОСТИ КЛИМАТИЧЕСКИХ И  
ГИДРОЛОГИЧЕСКИХ ХАРАКТЕРИСТИК**

**СОВРЕМЕННЫЕ МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ  
В ГИДРОЛОГИЧЕСКОМ МОДЕЛИРОВАНИИ<sup>5</sup>**

Айзель Г.В.

Институт водных проблем РАН, г. Москва  
hydrogo@yandex.ru

Решение задач, связанных с уменьшением ущербов от затопления освоенной «Водная стратегия Российской Федерации на период до 2020 года», предусматривает ряд мер по реализации опережающего инновационного развития научно-технической и технологической базы водохозяйственного комплекса на основе передовых достижений и технологий. Тем не менее, за период, прошедший с подписания этого документа в 2009 году, примеры практического внедрения лучших доступных технологий и методик в практику водохозяйственного проектирования и управления можно пересчитать по пальцам. Не лучшим образом обстоят дела и с публикационной активностью российских гидрологов: несмотря на рост числа публикаций и показателей их цитируемости, применение в практике расчётов современных методов исследований (математических, географических, гидрологических) остается на низком уровне.

Спектр гидрологических и водохозяйственных задач, требующих внедрения современных методов машинного обучения, охватывает почти всю предметную область, связанную с водными ресурсами суши. Виллиам Сье [1] выделяет три наиболее перспективные области применения указанных методов в приложении к гидрологическим задачам:

1. Обработка космических снимков высокого разрешения с выделением на них основных паттернов распределения характери-

---

<sup>5</sup> Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект № 14-05-00027).

стик почвенного и растительного покровов с последующим построением карт самоорганизующихся признаков;

2. Моделирование процессов тепло - и влагообмена поверхности суши с атмосферой (в том числе с учетом возможного изменения климата);

3. Моделирование водохозяйственных систем оперативного управления располагаемыми водными ресурсами.

В данной работе рассмотрено решение задачи поиска параметров физико-математической модели SWAP [2], необходимых для моделирования процессов формирования стока с неизученных водосборов. Модель SWAP разработана авторским коллективом лаборатории физики почвенных вод Института водных проблем РАН (руководитель д.б.н. Е.М. Гусев) с помощью алгоритма машинного обучения – аппарата искусственных нейронных сетей (ИНС), наиболее распространенного в практике гидрологических исследований [3].

Поскольку информационная база для российских водосборов, необходимая для разработки методологии использования методов машинного обучения, отсутствует, использовались данные наблюдений на 323-х речных водосборах (рис. 1), входящих в перечень объектов, предложенных организаторами проекта MOPEX (MOdel Parameter Estimation Experiment [4]).

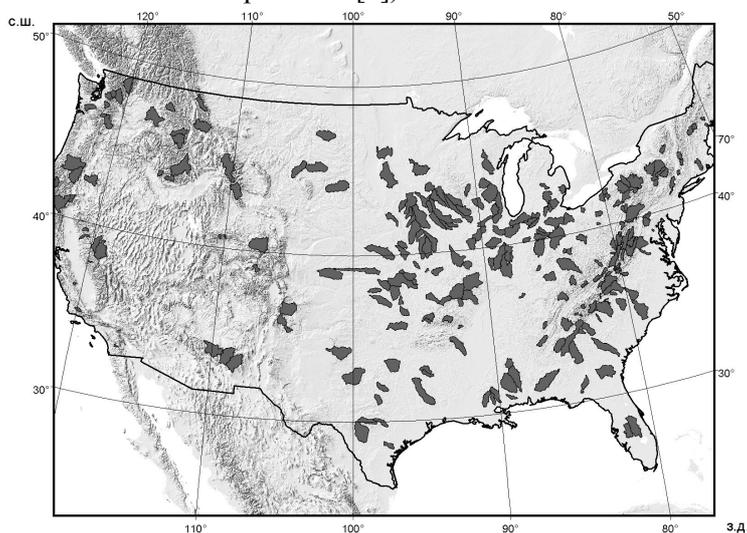


Рис. 1 – Расположение 323-МОРЕХ бассейнов на территории США

ИНС представляют собой систему соединённых и взаимодействующих между собой простых процессоров (искусственных нейронов). Каждый процессор подобной сети имеет дело только с сигналами, которые он периодически получает, и сигналами, которые он периодически посылает другим процессорам. Будучи соединёнными в достаточно большую сеть с управляемым взаимодействием (рис. 2), такие локально простые процессоры вместе способны выполнять довольно сложные задачи.

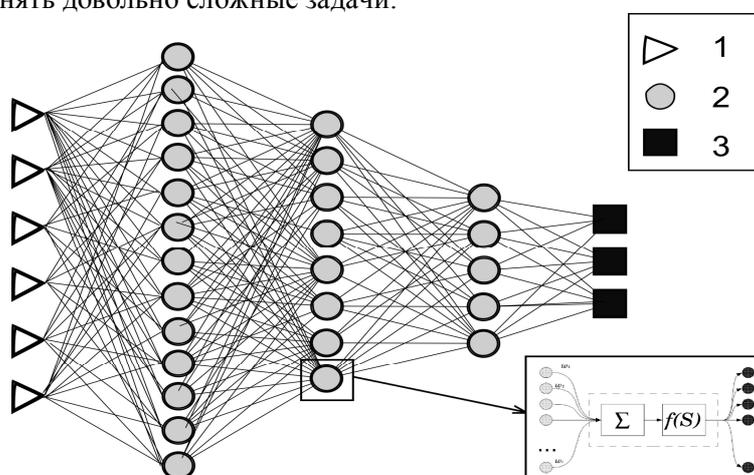


Рис. 2 – Пример структуры полной нейронной сети:  
1 – нейроны входного слоя, 2 – нейроны скрытого (промежуточного слоя),  
3 – нейроны выходного слоя

Нейронные сети не программируются в привычном смысле этого слова, они обучаются. Возможность обучения — одно из главных преимуществ нейронных сетей перед традиционными алгоритмами. Технически обучение заключается в нахождении коэффициентов связей между нейронами. В процессе обучения нейронная сеть способна выявлять сложные зависимости между входными и выходными данными, а также выполнять обобщение. Это значит, что в случае успешного обучения, сеть сможет воспроизвести верный результат на основании данных, которые отсутствовали в обучающей выборке, а также неполных или частично искаженных данных, что характеризует свойство адаптивности нейронных сетей к изменениям окружающей среды.

Задачей, на которую направлено исследование, является построение ИНС, связывающих модельные параметры водосбора с его

природными характеристиками. Соответственно этому, необходимое информационное обеспечение метода ИНС также делится на две группы: предиктанты (калибруемые параметры модели SWAP); предикторы (физико-географические дескрипторы водосбора из глобальных баз данных). В качестве последних были выбраны характеристики почвы (пористость, наименьшая влагоемкость, В-параметр Клэппа и Хорнбергера, коэффициент фильтрации, матричный потенциал почвенной влаги, толщина почвенной колонки); растительности (глубина корнеобитаемой зоны, средний размер листа, высота растительности, относительная площадь листьев, альbedo поверхности суши); климата (амплитуда среднемесячной температуры воздуха, среднегодовая приземная температура воздуха, среднегодовые осадки); геоморфологии (площадь бассейна, уклон, высота над уровнем моря.), взятые из глобальных баз данных метеорологических элементов и параметров подстилающей поверхности (всего 17 предикторов).

Обучение нейронных сетей для каждого из выбранных модельных параметров происходило на группе из 210 водосборов (65% от всей выборки), включающих подгруппу непосредственного обучения и подгруппу кросс-проверки. При построении искусственных нейронных сетей в работе использовались две базовые архитектуры ИНС – многослойный персептрон (МП) и радиальная базисная функция (РБФ). Также было принято решение не ограничиваться одной архитектурной реализацией, а использовать четыре реализации на базе МП и три на базе РБФ, что позволило шире освоить описанный инструментарий решения поставленной задачи и наметить пути установления и обоснования схемы выбора наиболее устойчивого решения.

Для нахождения модельных параметров для проверочной группы из 113 водосборов, отнесенных к группе «не обеспеченных информацией», необходимо в качестве входного слоя найденных ИНС задать предикторы этих водосборов из глобальных баз данных, которые, трансформируясь соответственно оптимизированной архитектурной схеме ИНС, дадут на выходе искомые значения параметров модели SWAP.

Расчеты речного стока с помощью модели SWAP для бассейнов, относящихся к указанной выше группе, с использованием параметров, полученных путем применения методов ИНС, производились для периода с 1986 по 1995 гг. Для всей группы (113 водо-

сборов) были последовательно рассчитаны суточные гидрографы стока, затем для данной выборки вычислены статистики качества расчетов  $Eff$  и  $|Bias|$  (см. рис. 3 и 4).

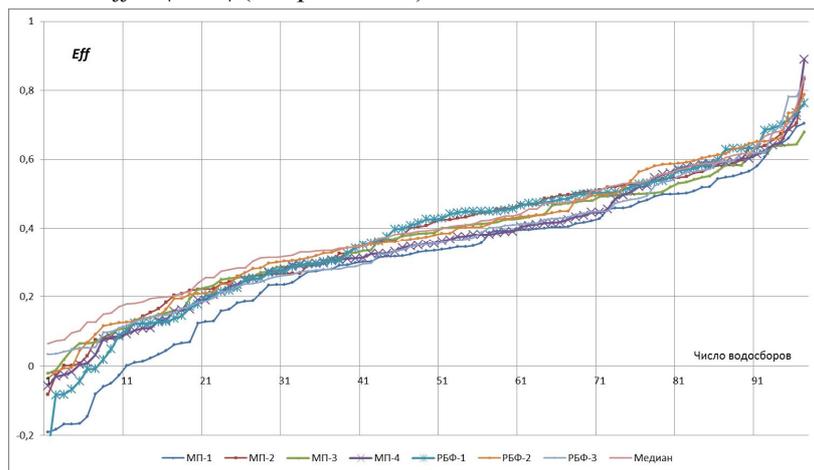


Рис. 3 – Эффективности  $Eff$  расчета суточных значений речного стока моделью SWAP для каждого из 113 водосборов по всем вариантам используемых архитектур ИНС (значения упорядочены по возрастанию)

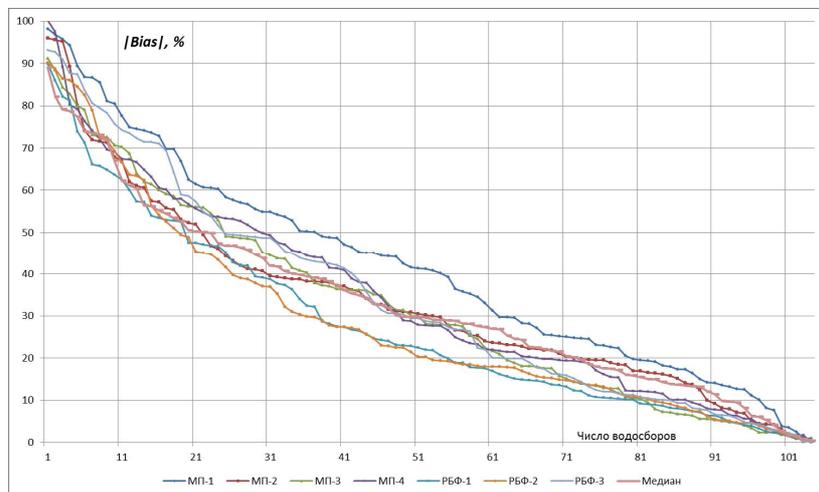


Рис. 4 – Систематические ошибки  $|Bias|$  расчета суточных значений речного стока моделью SWAP для каждого из 113 водосборов по всем вариантам используемых архитектур ИНС (значения упорядочены по убыванию)

Из полученных результатов следует, что лучшее качество расчетов стока для всей выборки «не обеспеченных информацией» бассейнов доступно при применении архитектуры РБФ-2 ко всем модельным параметрам – эффективность расчета 0,348 стремится к нижней границе допустимого «удовлетворительного» интервала, в то время как среднеквадратическая ошибка в 22,5% стремится к своей верхней планке.

Отметим также, что все вариации архитектуры РБФ показывают в среднем лучшие результаты, чем сети, построенные на архитектуре МП. Это может быть объяснено тем, что структура РБФ не производит экстраполяцию данных, в противовес МП. Таким образом, разброс модельных параметров, полученных с помощью РБФ, будет всегда в рамках соответствующего разброса на обучающей выборке. Данные результаты можно считать удовлетворительными. Географический охват исследуемых водосборов настолько широк, разнообразие стокоформирующих комплексов настолько велико, что применение единой архитектуры ИНС для расчета всех бассейнов дает большие ошибки в итоговых расчетах.

Следуя концепции «исключительности» водосбора, заключающейся в соответствии каждому из них своей единственной базовой архитектуры ИНС, применение которой позволяет получить лучший и приемлемый результат, для каждого водосбора была выделена архитектура ИНС, дающая лучший результат в терминах  $Eff$  и  $|Bias|$  (рис. 5).

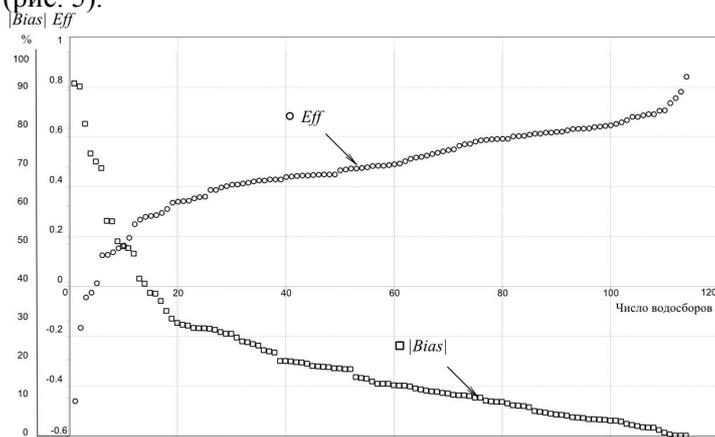


Рис. 5 – Эффективности  $Eff$  (о) и модули систематической ошибки  $Bias$  (□) расчета суточных значений стока по модели SWAP для каждого из 113 водосборов

Рассчитанные медианы  $Eff$  и  $|Bias|$  равны 0,484 и 13,7%, что подтверждает удовлетворительное качество вычислений.

Показанная в работе перспектива получения модельных параметров с высокой степенью достоверности их определения, необходимых для оценки характеристик речного стока для неизученных водосборов, открывает широкие горизонты по улучшению методик оценки водно-ресурсного потенциала территории, оценки ресурсов поверхностных вод, в том числе и в условиях изменяющегося климата.

### Литература

1. *Hsieh W.W.* Machine learning methods in the environmental sciences: neural networks and kernels. – Cambridge university press, 2009. p. 364.

2. *Гусев Е.М., Насонова О.Н.* Моделирование тепло- и влагообмена поверхности суши с атмосферой. М.: Наука, 2010. 327 с.

3. *Rasouli K., Hsieh W.W., Cannon A.J.* Daily streamflow forecasting by machine learning methods with weather and climate inputs // Journal of Hydrology. 2012. Т. 414. С. 284-293.

4. *Schaake J., Cong S., Duan Q.* The US MOPEX Data Set // IAHS Publications. 2006. № 307. С. 9-28.

## СОВРЕМЕННЫЕ ИЗМЕНЕНИЯ КЛИМАТИЧЕСКИХ ХАРАКТЕРИСТИК В БАССЕЙНЕ АМУРА

Болгов М.В., Трубецкова М.Д., Филиппова И.А.

Институт водных проблем РАН, г. Москва  
trubets@mail.ru

**Введение.** Катастрофическое наводнение, произошедшее на Амуре летом 2013 г., поставило перед исследователями вопрос о необходимости оценки возможности повторения явлений такого рода в будущем. В настоящее время отличительной особенностью климата на планете является его нестационарный характер, причем в различных районах климатические изменения различны, они неравномерны во времени, и интенсивность их неодинакова. Для оценки возможности повторения экстремального паводка 2013 г. на р. Амур, прежде всего, необходимо получить представление о современных тенденциях изменения климата на территории бассейна