

УДК: 004.942

Название: Идентификация и калибровка моделей на гидрометеорологических данных

Авторы:

Никитин Н.О, Университет ИТМО, г. Санкт-Петербург;

Калюжная А.В., Университет ИТМО, г. Санкт-Петербург

Контакты: nikolay.o.nikitin@gmail.com, +79062434402

Научный руководитель: Калюжная А.В, Университет ИТМО, г. Санкт-Петербург

Тезис доклада:

Для решения задач работы с гидрометеорологическими данными большую актуальность имеет идентификация моделей на этих данных. Высокая вычислительная сложность выполнения расчетов с помощью современных численных гидрометеорологических моделей делает необходимым использование таких вспомогательных моделей, построенных на данных, в ходе калибровки.

Калибровка моделей в ситуациях, в которых лишь небольшое количество данных наблюдений может быть использовано для настройки моделей, приводит к идентификации наборов параметров, оптимальных только для обучающей выборки данных – т. е. переобучению. Для решения этой проблемы могут быть предложены робастные стратегии калибровки, т. е. позволяющие находить удовлетворительное решение даже в ситуации с высокой неопределенностью.

В представленной работе реализация робастной калибровки реализована на основе генетических алгоритмов с модифицированной фитнес-функцией. Вместо использования только одной реализации модели, метод Монте-Карло применяется для сценарного зашумления входных данных и создания ансамбля моделей. Соответственно, значение целевой функции определяется как компромисс между ошибкой по реализациям в ансамбле относительно измерений и дисперсией всего ансамбля.

Специализированные модели на данных могут быть использованы как непосредственно для идентификации параметров численных моделей, так и для задач аппроксимации фитнес-функции для уменьшения числа запусков модели, необходимым для решения задач калибровки.

В качестве тестового случая была использована конфигурация модели SWAN для Карского моря, использующая ветровой форсинг, полученный из атмосферного реанализа ERA Interim. Для воспроизведения различных сценариев калибровки были использованы 9 синтетических временных рядов высоты волны в точках, полученных из результатов расчетов с помощью конфигурации модели WaveWatch III для всего Северного Ледовитого Океана. Запуск моделей проводился на высокопроизводительном блейд-сервере в многопоточном режиме.

Результатом работы является метод робастной ансамблевой калибровки модели, который может быть применен для различных предметных областей и использован совместно с вспомогательными моделями на данных.