

СРАВНИТЕЛЬНЫЕ ИСПЫТАНИЯ АЛГОРИТМОВ РЕДАКТИРОВАНИЯ ДАННЫХ ЭКСПЕРИМЕНТОВ

Хламов М.Г.

Донецкий государственный технический университет, кафедра ПЭ

Abstract

Chlamov M.G. Results measurements characters of algorithm-editor for distorted dates. Method use imitation model for received results.

Рассматриваемый алгоритм редактирования данных эксперимента предназначен для работы в составе гидрофизической информационно-измерительной системы (ГИИС) исследований мелкомасштабной вертикальной структуры гидрофизических полей верхнего деятельного слоя (ВДС) океана. Алгоритм редактирования содержит процедуру коррекции искаженных данных, в которой коррекция выполняется методами интерполяции и связана с классом исследуемых процессов. Формирование состоятельных оценок показателей точности алгоритма на основе экспериментальных данных требует наличия ансамблей реализаций широкого класса гидрофизических процессов. Имеющиеся сведения не отличаются достаточной полнотой и требуют дополнения. Это сопряжено с проведением экспедиций целевого назначения и, следовательно, требует значительных материальных затрат. Кроме этих затруднений, уникальность океанических процессов мелкомасштабной турбулентности, наблюдаемых в ходе экспедиционных исследований, не гарантирует достаточного представительства явлений пространственно-временной изменчивости.

Учитывая отмеченное, в качестве метода исследований использован метод имитационного моделирования, позволяющий в моделируемых условиях воспроизвести широкий спектр вариантов данных и получить результаты более простыми средствами.

Алгоритм редактирования данных эксперимента, описанный в

предыдущей статье, содержит три процедуры: процедуру обнаружения искаженных данных, процедуру их локализации и процедуру коррекции этих данных. Обнаружение и локализация искаженных данных построены на основе линейных адаптивных фильтров первого и второго порядков, используемых в режиме формирователей линейного прогноза. Оценка эффективности алгоритмов обнаружения и локализации выполнена в сравнении с идентичным алгоритмом, построенным на основе метода согласованной фильтрации [1]. Согласованный фильтр настраивается на мелкомасштабную составляющую исследуемых процессов. Разностное уравнение модели согласованного фильтра имело вид:

$$\hat{Y}[i+1] = a_1 \cdot x[i] + a_2 \cdot x[i-1], \quad (1)$$

где a_1 , a_2 - корреляционные коэффициенты, учитывающие характер мелкомасштабного процесса. Половина ширины доверительного интервала прогноза устанавливается, как и в основном алгоритме редактирования с учетом дисперсии прогноза, которая для согласованного фильтра устанавливается по формуле:

$$\bar{D}_\varepsilon \leq \sigma_0^2 \left\{ 1 - [(1 - a\Delta t)^2 + a^2 \cdot \Delta t^2] \cdot \alpha^2 \right\}, \quad (2)$$

где Δt - шаг дискретизации процесса, a - параметр модели мелкомасштабного процесса.

Процедура коррекции искажений данных выполнена в двух модификациях: итерационного алгоритма (СИ-ВП), включающего кубическую сплайн - интерполяцию и итерационный учет взвешенного прогноза; алгоритма однократного применения взвешенного прогноза (ВП).

Оценка метрологических характеристик алгоритма редактирования выполнена методами имитационного моделирования [2]. Имитационная модель функционирования ГИИС включала четыре раздела: модель динамического

гидрофизического параметра (модель данных), в качестве такого параметра выбрана одна из горизонтальных составляющих проекций вектора скорости турбулентного движения водных масс; модель преобразования данных в канале записи-воспроизведения цифрового магнитного регистратора (КЦМЗ), обеспечивающего накопление экспериментальных данных; алгоритма редактирования данных; алгоритма обработки результатов моделирования и расчета метрологических характеристик.

В модели данных имитационными методами отображался класс нестационарных процессов с изменяющимся средним и дисперсией и содержащих стационарную составляющую, описанную нормированной спектральной плотностью. Подобный подход и методы моделирования использованы в работе [3]. Изменяющееся среднее моделировало вертикальный профиль скорости ветрового течения и описывалось уравнениями Экмана. Изменениями дисперсии моделировались процессы:

- генерации турбулентности механизмами опрокидывания поверхностных волн и гидродинамической неустойчивости волновых движений, а также ее диссипацией в верхнем перемешанном слое (ВПС) океана; изменение с.к.о. скорости турбулентного движения здесь описывалось экспоненциально убывающей функцией;
- возрастания скорости турбулентного движения в области термоклина (ТК), убывания к нижней его границе и носящей характер перемежающихся всплеск описывалось огибающей, покрывающей всю область ТК и описываемой функцией вида $\sin^2(x)/x^2$, и вписанных под огибающей всплеск, описываемых аналогичными функциями, но покрывающими часть ТК в соответствии со статистическими данными;
- в тонкой вертикальной структуре гидродинамических полей нижней области ВДС в виде слоевых всплесках турбулентности описывались подобно всплескам турбулентности в ТК со статистическими характеристиками присущими для этой области.

Нормированная спектральная плотность стационарной составляющей

процесса с учетом пространственно-временных масштабов и аппаратурных ограничений ГИИС описывалась в виде:

$$S_x(\omega) = \frac{2}{\pi} \cdot \frac{a^3}{(a^2 + \omega^2)^2}, \quad (3)$$

где параметра a и вид зависимости отражают пространственно-временной масштаб процесса.

Стационарный процесс формируется цифровым фильтром из входного белого шума. Разностное уравнение, описывающее нормированный стационарный процесс, представляется в виде:

$$x[k] = U[k] \cdot \Theta_0 + U[k-1] \Theta_1 + x[k-1] \cdot \eta_1 + x[k-2] \cdot \eta_2, \quad (4)$$

где $x[k]$, $x[k-1]$, $x[k-2]$ - текущие и предшествующие значения временного ряда, описывающего нормированный стационарный процесс; $\Theta_0, \Theta_1, \eta_1, \eta_2$ - весовые коэффициенты; $U[k], U[k-1]$ - текущее и предшествующее значения нормального белого шума. С учетом параметров процесса и ГИИС, получены следующие значения параметров модели:

$$\Theta_0=0.104; \Theta_1=0; \eta_1=1.846, \eta_2=-0.852; D_U=0.039, m_U=0.$$

Нестационарный процесс формировался из стационарного путем манипуляции его статистическими параметрами, описанными выше функциями. Наихудший вариант данных представлен реализациями с максимальным значением дисперсии и неизменным средним, что соответствует верхней точке ВПС.

Модель преобразования данных в КЦМЗ представляет собой имитационную модель преобразования нестационарного сообщения, представленного в двоичном формате блоками по l слов с признаками начала

и конца блока, в нестационарном канале. Искажения двоичной информации описывались воздействием комплекса возмущающих факторов, моделирующих специфические морские условия, на функционирование ГИИС. В отличие от работы [4], при моделировании рассматривалось прохождение стохастических сообщений по цифровому каналу.

В процессе моделирования устанавливались следующие величины:

- показатели эффективности алгоритмов обнаружения, к которым отнесены коэффициент обнаружения искажений - $K_{ОИ}$ и коэффициент ложных тревог - $K_{ЛТ}$, поскольку оба коэффициента устанавливались, как частотные и могут интерпретироваться как вероятностные характеристики, то для них устанавливались доверительные границы при доверительной вероятности $P=0.95$;
- параметры погрешности данных, которые оценивались их статистическими характеристиками (выборочными средним и стандартным отклонением (с.к.о.), а также границами доверительных интервалов).

Для решения задачи обнаружения искажений наилучшими параметрами обладает алгоритм на основе согласованного фильтра ($K_{ОИ} = 0.61 \pm 0.02$, $K_{ЛТ} = 0.02 \pm 0.01$). Незначительно ему уступает алгоритм на основе адаптивного фильтра второго порядка ($K_{ОИ} = 0.58 \pm 0.02$, $K_{ЛТ} = 0.03 \pm 0.01$). С ростом доверительной вероятности и ширины доверительного интервала уменьшается число ошибок всех алгоритмов обнаружения.

Погрешность канала ($\sigma_{КДМЗ}$, $\Delta_{КДМЗ}$) представляется двумя составляющими: не обнаруживаемой погрешностью ($\sigma_{НО}$), вызванной искажениями, не превышающими доверительного интервала прогноза, и погрешностью восстановления алгоритмами коррекции искаженных участков данных ($\sigma_{ВО}$). Вторая составляющая включает погрешность обусловленную "ложной тревогой". Результаты вычисления статистических характеристик погрешностей измерены в единицах младшего разряда двоичного кода (ЕМР) и сведены в таблицу.

Таблица — Показатели точности алгоритмов редактирования

Алгоритм обнаруж.	Алгоритм коррекции	$\sigma_{HO}, \text{EMР}$	$\sigma_{BO}, \text{EMР}$	$\sigma_{KЦMЗ}, \text{EMР}$	$\Delta_{KЦMЗ}, \%$
$\sigma_{MMT} = \text{var}$					
АД1		21.2 ± 0.9	-	-	-
АД2	ВП(АД2)	11.4 ± 0.4	2.7 ± 0.1	6.9 ± 0.3	0.3
ОФ	ВП(ОФ)	4.8 ± 0.2	2.3 ± 0.1	3.4 ± 0.2	0.2
АД2	СВ-ВП	11.4 ± 0.4	1.9 ± 0.1	6.8 ± 0.3	0.3
$\sigma_{MMT} = \text{max} (5 \text{ см/с})$					
АД1		23.3 ± 1.0	-	-	-
АД2	ВП(АД2)	11.7 ± 0.4	5.8 ± 0.2	10.0 ± 0.4	0.5
ОФ	ВП(ОФ)	7.3 ± 0.3	4.7 ± 0.2	9.6 ± 0.3	0.5
АД2	СВ-ВП	11.7 ± 0.4	5.1 ± 0.2	6.1 ± 0.3	0.3

Выборочные средние всех составляющих погрешностей не превысили единицы младшего разряда кода. Это дает основание считать их нулевыми. Из параметров, характеризующих нестационарность гидрофизических процессов, влияние на погрешности оказывало изменение среднеквадратичной скорости мелкомасштабной турбулентности. Влияния других параметров не обнаружено. Наибольшие погрешности достигались при максимальной среднеквадратичной скорости турбулентности. Оценки σ_{HO} слабо связаны со всеми параметрами сообщений и КЦМЗ и определяются только шириной доверительных интервалов прогнозов. Эта составляющая погрешности превышает σ_{BO} в 2...5 раз, но в долевого отношении составляет 30...40 %. Погрешность КЦМЗ при изменениях во всем диапазоне мелкомасштабной скорости не превысила 0.3%. В моделируемых условиях эта погрешность возрастает до 0.5% ,при этом не обнаруживаемая составляющая достигает 0.6%.

Анализ показателей точности алгоритмов коррекции позволяет сделать вывод об эквивалентности алгоритмов ВП на основе прогнозирующих фильтров и некотором преимуществе итерационного алгоритма (СИ-ВП). Алгоритм редактирования, построенный на методе адаптивной фильтрации, незначительно уступая по точности аналогичному алгоритму, использующему согласованную фильтрацию, существенно превосходит его по мобильности, так как не требует для своей организации априорных сведений о классе и

параметрах исследуемых процессах. Алгоритм базируются на интенсивном анализе информации, содержащейся в данных, и ее использовании для адаптивной подстройке параметров модели. Вариант процедуры коррекции по методу взвешенного прогноза примечателен тем, что во всех трех процедурах алгоритма редактирования для выполнения различных функций: обнаружения искаженных данных, их локализации и коррекции, - используется один и тот же универсальный алгоритм адаптивной цифровой фильтрации.

Литература

1. Казакевич Д.И. Основы теории случайных функций в задачах гидрометеорологии. - Л.: Гидрометеиздат, 1989. - 232 с.
2. Цифровое моделирование систем стационарных случайных процессов / Е.А. Гриндина, А.Н. Лебедев, Д.Д. Недосекин, Е.А. Чернявский. - Л. Энергоатомиздат. Ленинградское отделение, 1991. - 144 с.
3. Калацкий В.И. Моделирование вертикальной структуры деятельного слоя океана. - Л.: Гидрометеиздат, 1978. - 216 с.
4. Вичес А.И., Горон А.И., Смирнов В.А. Моделирование канала магнитной записи на ЭВМ.-М.: Радио и связь, 1984. - 184с.