

# УСОВЕРШЕНСТВОВАННАЯ ПРОЦЕДУРА ГЕОГРАФИЧЕСКОЙ ПРИВЯЗКИ ДАННЫХ ИЗМЕРЕНИЙ ДРИФТЕРОВ, ПОЛУЧЕННЫХ В ПЕРИОД С 2001 ПО 2007 гг.

А.Л. Холод

Морской гидрофизический институт  
НАН Украины  
г. Севастополь, ул. Капитанская, 2  
E-mail: antonholod@mail.ru

Дрифтеры в настоящее время являются основным инструментом получения измерений *in situ* в Черном море. До 2006 г. обсервационные данные координат дрифтеров определялись по системе ARGOS и шли реже, чем данные измерений. Поэтому необходимо было восстанавливать его координаты на моменты времен измерений. Ранее применялась линейная интерполяция для восстановления координат, однако, при этом допускаются большие погрешности в определении местоположения, а также не воспроизводятся инерционные колебания. В связи с этим была предпринята попытка улучшить восстановление данных. За счет этого планируется улучшить точность при определении скоростей течений и при сопоставлении различных данных с данными дрифтеров.

Один из усовершенствованных методов восстановления географических координат дрифтеров был предложен в статье [1]. Алгоритм, предложенный автором статьи, основывается на представлении функции, описывающей траекторию движения дрифтера, в виде суммы полинома и тригонометрических функций, которые учитывают инерционные колебания движения дрифтера

$$\varphi = a_0 + a_1 t + \dots + a_n t^n + a_{n+1} \cos(\omega t) + a_{n+2} \sin(\omega t) \quad (1)$$

Аналогичная форма представления используется для представления долготы  $\lambda$ .

$$\lambda = b_0 + b_1 t + \dots + b_n t^n + b_{n+1} \cos(\omega t) + b_{n+2} \sin(\omega t) \quad (2)$$

Проблема заключается в выборе оптимальной степени полинома. На основе эвристического подхода формулируются правила для разбиения временного ряда данных на сегменты. Длина сегмента зависит от частоты поступления данных, а также от пространственно-временной разницы между соседними измерениями. Для каждого сег-

мента определяются оптимальные коэффициенты. Сам алгоритм является достаточно сложным, но, тем не менее, он был нами реализован и проверен на имеющихся у нас данных.

Существуют и другие методы решения поставленной задачи, например, метод локальной аппроксимации (МЛА). На наш взгляд этот метод является более предпочтительным. МЛА можно рассматривать обобщением таких методов, как метод обратных расстояний, метод Кrigинга, применяемый в Surfer и др. Основное отличие от них заключается в том, что в окрестности точки приближения аппроксимируемая функция представляется не константой, а имеет более сложной вид за счет добавления степенных, тригонометрических или других функций [2]. Благодаря этому достигается более высокая точность приближения данных. Метод позволяет решить эту задачу на основе варьирования как самого отрезка на котором производится аппроксимация, так и степени полинома, с использованием для этих целей метода перекрестного скользящего экзамена.

В соответствии с МЛА с каждым  $x$  связывается свой набор наблюдений  $z^{(s)}$ , соответствующих узлам  $x^{(s)}$ , из определенной окрестности  $x$ . Таким образом, набор наблюдений оказывается меняющимся, определяемым каждый раз как некоторая окрестность  $x$ . В каждой такой окрестности строится аппроксимация и определяются оптимальные коэффициенты. Оценка, получаемая по МЛА, определяется на основе минимизации следующего функционала

$$\begin{aligned} \hat{y}_N(x, \delta) &= \hat{c}_N^T \varphi(0), \hat{c}_N = \arg \min_c J_N(x, c, \delta) \\ J_N(x, c, \delta) &= \frac{1}{N} \sum_{s=1}^N \rho\left(\frac{x - x^{(s)}}{\delta}\right) F(z^{(s)} - c^T \varphi(x - x^{(s)})), \end{aligned} \quad (3)$$

где  $\varphi(x) \in R_m$  – вектор заданных координатных функций;  $c \in R_m$  – вектор неопределенных коэффициентов;  $\hat{c}_N$  – оценка этого вектора, определяемая минимизацией функционала  $J_N$  по  $c$ ;  $\delta$  – параметр локальности;  $F(\cdot)$  – неотрицательная функция потерь;  $\rho(u)$  – функция локальности, имеющая максимум в нуле и стремящаяся к нулю при  $\|u\| \rightarrow \infty$ ;  $\hat{y}_N(x, \delta)$  – оценка значения  $y(x)$  в точке  $x$ ;  $t$  – символ транспонирования.

Для реализации процедуры локальной аппроксимации вводится специальная весовая функция  $\rho(u)$  – функция локальности. Она обеспечивает наибольший вес в сумме слагаемым, соответствующим значениям  $x^{(s)}$  наиболее близким к центру аппроксимации. На рисунке 1 приведены примеры функций локальности.

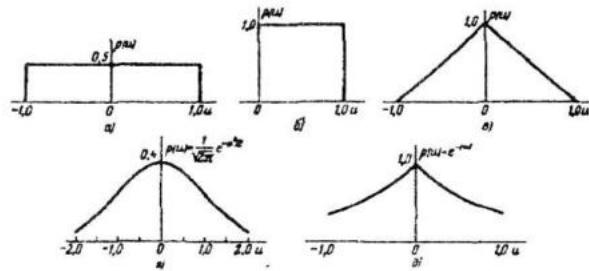


Рисунок 1 – Примеры функций локальности

Рациональный выбор параметра локальности –  $\delta$  является одним из ключевых вопросов непараметрического оценивания. Эффективным средством для нахождения оптимальности параметра локальности является деление наблюдений на два множества: обучающее и контрольное. Оценка функции  $\hat{y}_N(x, \delta)$  строится по совокупности  $\{z^{(s)}, x^{(s)}\}$ , образующей обучающее множество. Контроль точности выполняется по наблюдениям другого, контрольного множества. Формы контроля могут быть различными, в частности, базироваться на методах анализа суммы квадратов невязок оценок  $e^{(s)} = z^{(s)} - \hat{y}_N(x^{(s)}, \delta)$  и вычисляемых в узлах, образующих контрольное множество. При малом объеме выборки деление ее на два подмножества – недопустимое расточительство, ведущее к тому, что и оценка, построенная по уменьшенному обучающему множеству, будет плохой, да и контроль качества мало достоверным. Для выбора оптимального параметра локальности используется метод скользящего экзамена (МСЭ). Он заключается в формировании обучающего и контрольного множеств из одного массива данных путем многократного разбиения этого массива на две части, одна из которых используется для построения оценки, а оставшаяся, вторая – для контроля. МСЭ требует громоздких вычислений. Существенно сокращения объема вычислений можно добиться в случае, когда в условиях многократного разбиения данных на обучающее и контрольное множества, последнее содержит

жит только один элемент. Данный способ деления использован в настоящей работе.

Определим оптимальную степень полинома и параметр локальности на основе сопоставления полученных данных с контрольным массивом измерений. Исходными для вычислений являются данные координат дрифтера, полученные системой ARGOS. Контрольный массив представляет аналогичные данные, полученные системой GPS.

На графике, представленном на рисунке 2, показаны четыре кривые, отражающие зависимость среднеквадратической погрешности вычислений от параметра локальности. Кривые отличаются максимальной степенью полинома, включенного в базисные функции. Оптимальное значение параметра локальности относительно слабо зависит от максимальной степени полинома базисных функций и величины параметра локальности в интервале [0.3–0.5].

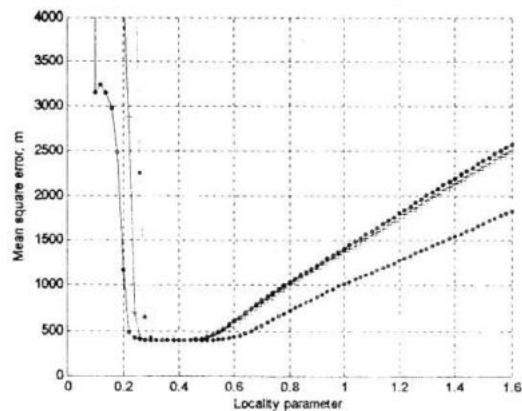


Рисунок 2 – Зависимость среднеквадратической погрешности вычислений от параметра локальности при прямом сопоставлении

Проверим эффективность процедуры подбора оптимального параметра локальности методом скользящего экзамена.

Из рисунка 3 видно, что с повышением максимальной степени полинома зависимость точности от параметра локальности смещается вправо. При этом величина оптимального параметра локальности возрастает. Значение последнего лежит в интервале [0.3–0.5]. Величина среднеквадратической погрешности увеличилась примерно в два раза. Это связано с тем, что когда точность оценивается по тем же данным ARGOS, сами контрольные данные имеют погрешность, существенно превышающую ошибки координат, полученных с помощью GPS.

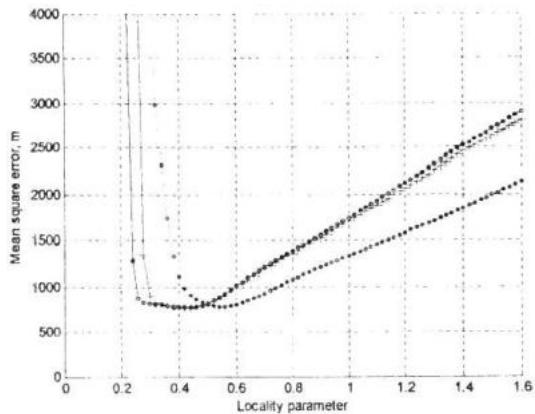


Рисунок 3 – Зависимость среднеквадратичной погрешности, от параметра локальности, вычисленного по МСЭ

Методика предполагает вычисления оптимального параметра локальности МСЭ и использование его в дальнейшем в МЛА. Произведя вычисления по описанной выше методике (максимальная степень полинома равна 3, параметр локальности равен 0.52), и сопоставив их с контрольным массивом, получили значение среднеквадратической погрешности, представленное в таблице I.

Таблица I – Среднеквадратическая ошибка восстановления данных различными методами

| Метод восстановления координат | Среднеквадратическая ошибка, м |
|--------------------------------|--------------------------------|
| Линейная интерполяция          | 549.5                          |
| Метод, предложенный в [1]      | 489.1                          |
| МЛА                            | 407.9                          |

Полученные результаты отражают тот факт, что метод локальной аппроксимации обеспечивает более точное восстановление траектории движения дрифтера в морской среде. Для иллюстрации данного преимущества в качестве примера приведем участок траектории движения дрифтера, где видны инерционные колебания.

На рисунке 4 пунктирной линией изображена траектория движения, восстановленная с помощью линейной интерполяции; штрихпунктирной линией – траектория движения, восстановленная с помощью МЛА; сплошной линией – траектория движения дрифтера, полученная по данным GPS.

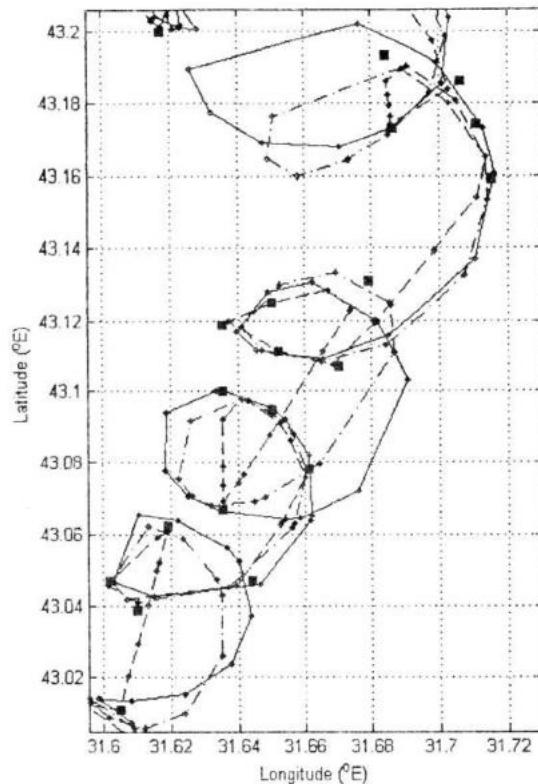


Рисунок 4 – Участок траектории движения дрифтера, восстановленный различными методами

**Выводы.** Создана улучшенная процедура восстановления координат дрифтеров, на основе метода локальной аппроксимации. Произведено сопоставление результатов работы данной процедуры с аналогичной процедурой, предложенной в статье [1] и методом линейной интерполяции. Восстановление траектории движения дрифтера по методу локальной аппроксимации лучше воспроизводит инерционные колебания, но, тем не менее, и при его применении могут наблюдаться заметные искажения.

## ЛИТЕРАТУРА

1. Bograd et.al. On sampling strategies and interpolation schemes for satellite-tracked drifters.– Journal of atmospheric and oceanic technology. – July 1999. – P. 893–904.
2. В.Я. Катковник. Непараметрическая идентификация и сглаживание данных: метод локальной аппроксимации. – М.: Главная редакция физико-математической литературы, 1985. – 336 с.