

СОПОСТАВЛЕНИЕ ТЕМПЕРАТУРЫ ПОВЕРХНОСТИ ЧЕРНОГО МОРЯ, ПОЛУЧЕННОЙ ПО ДАННЫМ ИЗС NOAA, С ИЗМЕРЕНИЯМИ ДРИФТЕРОВ В 2005 – 2006 ГОДАХ

E.B. Плотников, Ю.Б. Ратнер

Морской гидрофизический институт
НАН Украины
г. Севастополь, ул. Капитанская, 2

*E-mail: plotnikov_e_v@mail.ru;
yubrat@mail.ru*

Излагаются результаты сопоставления спутниковых и контактных данных о температуре поверхности Черного моря, полученных в 2005 – 2006 годах при помощи SVP дрифтеров и радиометра AVHRR, установленного на бортах ИЗС NOAA. Рассматриваются особенности и результаты применения усовершенствованной методики фильтрации недостоверных данных из тестовых массивов совместных измерений, используемых для сопоставления.

Введение. В работе изложены результаты сопоставления данных о температуре поверхности Черного моря, полученных в 2005 – 2006 годах при помощи SVP – дрифтеров и измерений радиометра AVHRR, установленного на бортах ИЗС серии NOAA. Целью такого сопоставления является оценка эффективности методики восстановления ТПМ по данным спутниковых измерений. Эта методика постоянно модернизируется, и для оценки адекватности вносимых усовершенствований необходимо сравнение получаемых результатов с данными, полученными при помощи контактных измерений. Как результат усовершенствования алгоритмов восстановления ТПМ следует ожидать уменьшение систематического и среднеквадратичного отклонений между спутниковой температурой и температурой, измеренной с помощью дрифтеров по сравнению с аналогичными результатами, полученными ранее. Одна из задач этой работы состоит в подтверждении этих предположений. При сопоставлении спутниковых данных с измерениями дрифтеров возникает ряд специфических проблем, приводящих к несовпадению сопоставляемых величин, не связанному с ошибками применяемых

методик. Поля ТПМ, получаемые по спутниковым данным и измеряемые с помощью дрифтеров, осреднены по разным пространственным масштабам. Спутниковый радиометр воспринимает температуру поверхностного слоя моря толщиной порядка 10–100 мкм, осредненную по площади порядка 1 км². Дрифтер же измеряет температуру на глубине 0.25 м в точке с фиксированными горизонтальными координатами. Кроме этого, имеют место несовпадения моментов времени измерений, ошибки, как географической привязки спутниковых данных, так и определения местоположения дрифтеров, различия в технических характеристиках радиометров, установленных на разных спутниках, сбои в работе аппаратуры и прочее. Подробно влияние этих факторов изучено в [1].

В большинстве случаев, ошибки, связанные с перечисленными факторами, лежат в допустимых пределах, соответствующих условиям эксплуатации аппаратуры и ожидаемым точностям применяемых методик обработки измерительной информации. К другой категории относятся ошибки, имеющие большие величины и для которых требуется изучение причин их возникновения. Поэтому одним из этапов работы по сопоставлению данных является изучение подобных аномалий.

При выполнении данной работы была применена методика выявления аномалий в массивах данных, основанная на принципах робастной статистики. Эта методика представляет собой попытку повысить уровень автоматизации выделения некорректных данных в тестовых массивах, что будет способствовать повышению надежности и сокращению времени выполнения этих операций.

Процедура построения тестовых массивов. Для формирования массива спутниковых данных использовалась информация, полученная в 2005 – 2006 годах на приемной станции МГИ НАНУ с бортов ИЗС NOAA-16, NOAA-17, NOAA-18. Был применен усовершенствованный алгоритм фильтрации облачности, который будет изложен в отдельной статье, готовящейся к публикации. Контактные данные о величинах ТПМ в 2005 году были получены с помощью двух дрифтера типа SVP-BTC60 и двух дрифтеров типа SVP-BT. Измерения

проводились с 24 января по 29 августа. В 2006 году были использованы два дрифтера типа SVP-BTC80. Измерения проводились с 3 июля по 26 сентября. Шаг дискретизации по времени составлял приблизительно один час. Подробные сведения о технических характеристиках и особенностях применения указанных типов дрифтеров изложены, например, в [2]. При формировании массивов данных выбирались только те измерения, для которых:

1. Разность времен между спутниковыми и дрифтерными данными не превышала 1/2 часа.

2. Перепад величин на карте спутниковой термодинамической температуры по пространственному окну 3x3 пикселя с центром в рассматриваемом пикселе не превышал 0.5 градуса.

3. В пространственном окне, эквивалентном по масштабу фрагменту карты спутниковой термодинамической температуры размером 5x5 пикселей с центром в рассматриваемом пикселе, отсутствовала облачность.

В итоге были сформированы два тестовых массива, содержащих измерения за 2005 и 2006 годы. По информации, содержащейся в тестовых массивах, были рассчитаны отклонения между дистанционными и контактными измерениями ТПМ. Образец временного ряда отклонений за 2005 год показан на рисунке 1.

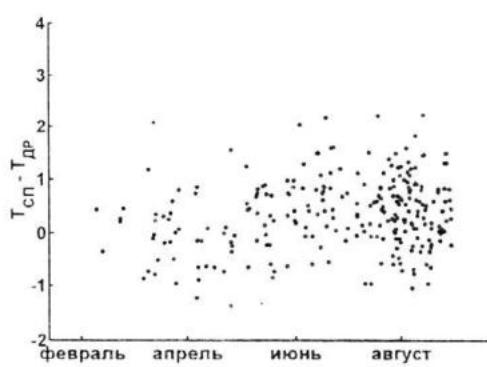


Рисунок 1 – График временного ряда отклонений, рассчитанных по данным 2005 года

Методика выделения некорректных данных. Анализируя полученные временные ряды, можно заметить небольшое число точек, подозрительных на аномальные отклонения в данных. Для выделения и изуче-

ния этих точек применялась методика, состоящая из следующих шагов:

1. Выделение аномальных измерений при помощи итеративно-взвешенного метода наименьших квадратов [3].

2. Изучение причин возникновения резко выделяющихся отклонений с привлечением дополнительной спутниковой, метеорологической и гидрологической информации.

3. Исключение из тестового массива аномальных измерений, причины возникновения которых удалось установить.

В ходе реализации этой процедуры в результате работы итеративно-взвешенного МНК рассчитываются веса отклонений, изменяющиеся в диапазоне от 0 до 1. Зависимость этих весов от абсолютных величин отклонений является монотонно убывающей. Чем меньше величина веса, тем большее степень аномальности соответствующего отклонения. Следовательно, ограничив веса снизу некоторым пороговым значением, можно выделить аномальные отклонения. Для определения этого порогового значения рассматривается гистограмма весов. Гистограмма, как правило, имеет более или менее выраженную двухмодовую структуру. Первая мода сосредоточена в окрестности больших весов, вторая – малых. В большинстве случаев форма гистограммы позволяет достаточно уверенно сделать вывод о величине искомого порога. На рисунках 2 и 3 проиллюстрирован этот этап методики.

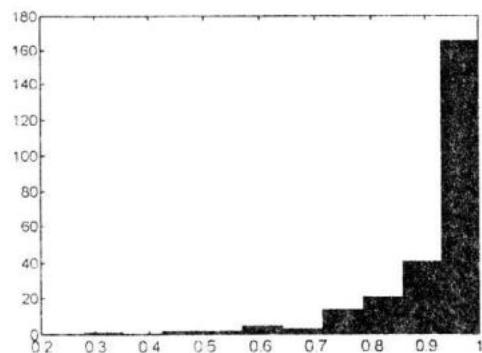


Рисунок 2 – Гистограмма весов отклонений, построенная по данным 2005 года

На рисунке 2 показана гистограмма весов отклонений для данных 2005 года. Здесь двухмодовая структура выражена нечетко.

Было решено выбрать пороговое значение, равное 0.6. На рисунке 3 показана гистограмма весов отклонений для данных 2006 года. Здесь двухмодовая структура выражена явно. В данном случае пороговое значение равно 0.5.

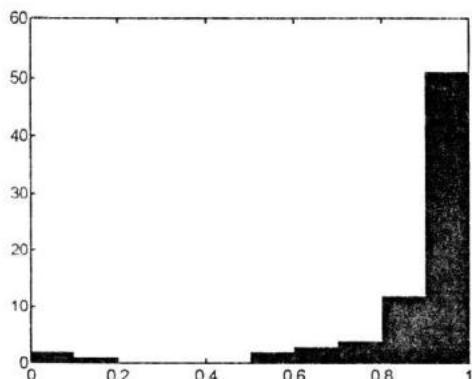


Рисунок 3 – Гистограмма весов отклонений, построенная по данным 2006 года

На втором этапе привлекаются географическом положении изучаемой точки, данные о скорости ветра, мелкомасштабной изменчивости поля ТПМ в рассматриваемом районе. Также рассматриваются характеристики поля ТПМ, полученные с карт, построенных в ходе близких по времени дистанционных съемок данного региона. Далее аномальные данные, причины возникновения которых удалось установить, исключаются из массива данных. При рассмотрении изучаемого тестового массива причинами возникновения аномальных ошибок в основном оказывались эффекты, связанные с дневным прогревом и неоднородностью поля температуры. Следует отметить, что в данном случае в силу слабой статистической обеспеченности были исключены из рассмотрения данные, полученные в сентябре 2006 года.

В результате применения вышеописанной процедуры к тестовым массивам исключены аномальные данные. Они показаны

крестиками на рисунках 4 и 5.

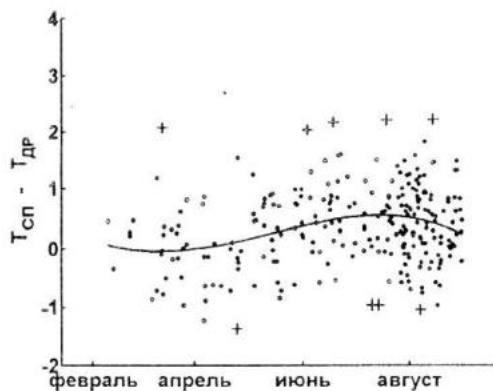


Рисунок 4 – График временного ряда отклонений спутниковых и контактных ТПМ, рассчитанных по данным 2005 года

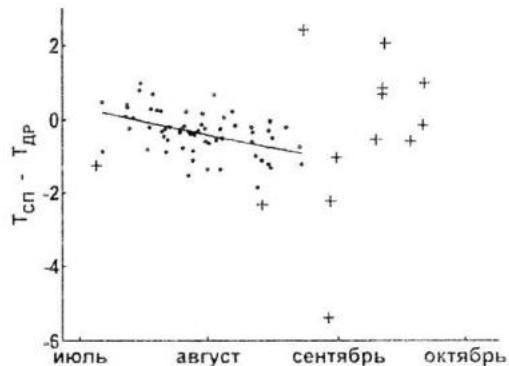


Рисунок 5 – График временного ряда отклонений спутниковых и контактных ТПМ, рассчитанных по данным 2006 года

Результаты. Средние и среднеквадратичные отклонения между спутниковыми и контактными ТПМ, а также коэффициенты линии регрессии, построенной по нормализованной независимой переменной

$\tilde{x} = \frac{x - \mu_x}{\sigma_x}$, где x – время, прошедшее с начала года, и измеряемое в сутках, приведены в таблице 1.

Таблица 1 – Статистические характеристики отклонений между спутниковыми и контактными ТПМ

Год	μ	σ	α_1	α_2	α_3	α_4
2005	0.35°	0.6°	0.4961 ± 0.1212	0.1853 ± 0.142	-0.2079 ± 0.1418	-0.0832 ± 0.0801
2006	-0.31°	0.61°	-0.2273 ± 0.1386	-0.3552 ± 0.1375	–	–

В таблице 1 μ и σ обозначают соответственно систематическое смещение и среднеквадратичное отклонение. Линия регрессии, построенная по данным 2005 года и показанная на рисунке 4 – кубическая, с коэффициентами $\alpha_1 - \alpha_4$. Кривая регрессии, построенная по данным 2006 года, и показанная на рисунке 5, является линейной с коэффициентами $\alpha_1 - \alpha_2$. Выбор вида регрессии осуществлен на основе анализа 95% доверительного интервала для оценок коэффициентов α_i . Все четыре коэффициента кубической регрессии, построенной по данным 2005 года, значимы по доверительному уровню 95%.

Анализируя представленные временные ряды, можно заметить, что они могут иметь тренды. Оба тренда характеризуются понижением величин отклонений от лета к осени. Для тренда, рассчитанного по данным 2005 года, характерно повышение отклонений от зимы к лету.

Выводы. Анализируя результаты, приведенные в таблице, следует отметить уменьшение среднеквадратичной погрешности по сравнению с аналогичными результатами, полученными в 2003 году [1]. Обращает на себя внимание существенное различие между систематическими смещениями по данным 2005 и 2006 годов. Сравнение полученных результатов с аналогичными за 2003 год позволяет говорить о примерно 10 процентном уменьшении отклонений между данными дистанционных и контактных измерений полей ТПМ и, следовательно, об улучшении методики восстановления температуры поверхности моря по спутниковым данным.

Анализ полученных данных показал различия в статистических характеристиках распределения отклонений между спутниковыми и контактными ТПМ, полученными по данным 2005 и 2006 годов. Кроме того замечена зависимость величин возникающих ошибок от сезона, в котором производятся измерения. Эти данные требуют дальнейшего изучения. В ходе дальнейших ис-

следований необходимо более детально выяснить причины различий в систематических отклонениях и различий в трендах.

Предложенная процедура фильтрации аномальных измерений позволила ускорить и повысить достоверность их выделения из формируемых тестовых массивов.

В перспективе предполагается продолжать работу по сопоставлению спутниковых и контактных ТПМ на систематической основе, по мере поступления новых данных и усовершенствований методики.

ЛИТЕРАТУРА

1. Ю.Б. Ратнер, Д.М. Соловьев, Е.И. Канин. Сопоставление величин поверхностной температуры Черного моря, полученных по данным аппаратуры AVHRR-3 ИЗС NOAA и SVP-дрифтеров в марте–августе 2003 г. // Экологическая безопасность прибрежной и шельфовой зон и комплексное использование ресурсов шельфа. – Севастополь. ЭКОСИ – Гидрофизика. Вып. 11. – 2004. – С. 155 – 173.
2. А.П. Толстошев, Е.Г. Лунев, Г.К. Коротаев, С.В. Мотыжев Термопрофилирующий дрейфующий буй. . // Экологическая безопасность прибрежной и шельфовой зон и комплексное использование ресурсов шельфа. – Севастополь. ЭКОСИ – Гидрофизика. Вып. 11. – 2004. – С. 143 – 154.
3. Holland, P.W., and R.E. Welsch (1977), "Robust regression using iteratively reweighted least-squares," Communications in Statistics: Theory and Methods, A6. – Р. 813 – 827.
4. Ю.Б. Ратнер, Т.М. Баянкина Сопоставление величин поверхностной температуры, полученных по модели динамики Черного моря и данным SVP-дрифтеров в марте–августе 2003 г. // Экологическая безопасность прибрежной и шельфовой зон и комплексное использование ресурсов шельфа. – Севастополь. ЭКОСИ – Гидрофизика. Вып. 11. – 2004. – С. 51 – 62