

УДК 551.465

## АЛГОРИТМ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ ТЕМПЕРАТУРНЫХ ОКЕАНИЧЕСКИХ ФРОНТОВ

© 2014 г. В. А. Давидович

ООО “Гроссмейстер”, Москва

e-mail: krolik39@gmail.com

Поступила в редакцию 08.11.2012 г., дата доработки 01.02.2013 г.

Роль и значение фронтов в описании океанской циркуляции давно признаны. Появление спутниковой океанографии привело к разработке методов для автоматического обнаружения фронтов с использованием спутниковых композитных данных. В этой работе мы представляем алгоритм для детектирования фронтов на основе совместного исследования градиентов и изотерм. Алгоритм распознает фронт как последовательность изотерм, обладающих определенными свойствами. Этот подход позволяет обнаружить индивидуальный объект, например, границу склоновой водной массы или северную границу Гольфстрима непосредственно, без предварительного детектирования всех температурных границ в системе. Алгоритм сочетает простоту реализации и высокую вычислительную эффективность с хорошей точностью и высокой надежностью детектирования.

DOI: 10.7868/S0030157414010018

### ВВЕДЕНИЕ

Океанический фронт это узкая зона с относительно высокими значениями горизонтальных градиентов свойств (температуры, солености, питательных веществ, и т.д.), которая разделяет более широкие области, представленные различными водными массами или отличающиеся вертикальной структурой (стратификацией) [3].

Роль и значение фронтов в описании океанской циркуляции давно признаны. Во многих случаях фронты являются главными циркуляционными элементами, определяющими динамику океана [2]. Описание пространственно-временной эволюции фронтов является ключевым компонентом прогноза циркуляции и ассимиляции наблюдений [7]. Примером крупномасштабного океанического фронта является фронт Гольфстрима.

С развитием спутниковой океанографии возросло понимание важности детектирования фронтов и прослеживания их пространственно-временной изменчивости “в реальном времени”. Уникальные возможности для этого предоставляют спутниковые изображения радиометров AVHRR, SeaWIFS, MODIS. В связи с интенсивным развитием различных моделей океанской циркуляции важной задачей становится также детектирование фронтов в модельных данных.

Растущий интерес к автоматическим методам детектирования фронтов с использованием спутниковой радиометрии и модельных полей температуры поверхности океана (ТПО) привел к разработке алгоритмов, основанных на различных подходах, начиная с простых методов, основан-

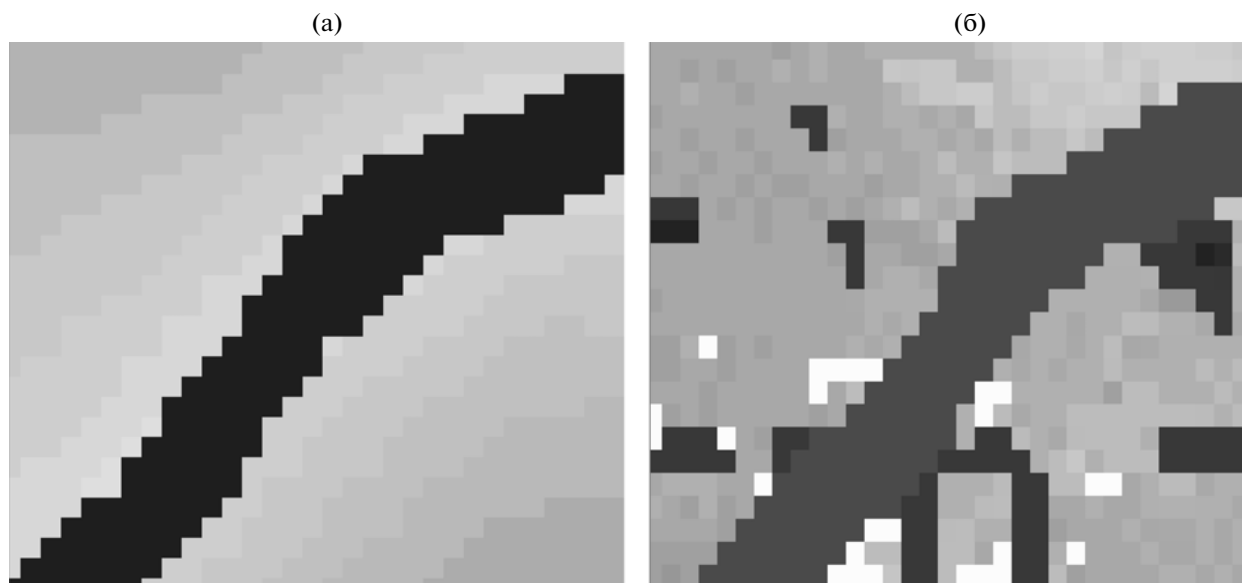
ных на расчетах локальной статистики изменчивости градиентов или аномалий ТПО [6] до более сложных методов, таких как кластерные [9] или основанные на гистограммах [4, 5, 10].

В этой работе мы представляем алгоритм детектирования фронтов на основе совместного анализа градиентов и изотерм, который характеризуется простотой реализации и высокой вычислительной эффективностью, обеспечивая при этом приемлемую точность и высокую надежность (Алгоритм Градиент/Изотерма, АГИ).

Алгоритм позволяет обнаружить индивидуальный объект, например, границу склоновой водной массы или северную границу Гольфстрима, непосредственно, без предварительного детектирования всех температурных границ в системе. Подобная особенность существенно увеличивает скорость работы алгоритма, особенно если нужны не все элементы фронтальной системы, что, например, является обычной ситуацией при ассимиляции данных в моделях циркуляции.

### ДАННЫЕ

Использовано поле ТПО, полученное в результате работы модели MARCOOS/HOPS Real-Time Forecast System. Эта система, функционирующая с 2009 г., основана на “элемент-ориентированном” (feature-oriented) подходе, разработанном Гангопадхья и др. [8] для региона меандров и рингов Гольфстрима. В работе также использованы композитные данные ТПО, основанные на спутниковых изображениях NOAA CoastWatch, West Coast Node, dataset: SST, Blended, Global,



**Рис. 1.** Визуальное отличие модельных данных (а) от композитных (б). Регион захватывает линию фронта, показанную темной линией. Черные и белые пиксели на рисунке (б) показывают сбойные и отсутствующие данные соответственно. Сходные градации полутонов шкалы температур являются результатом преобразования исходных данных в черно-белый формат для целей данной статьи.

EXPERIMENTAL/5-day (<http://thredds1.pfeg.noaa.gov:8080/thredds/Satellite/aggregatBA/ssta/catalog.html?dataset=satellite/BA/ssta/5day>). Композитные изображения ТПО сформированы из микроволновых и инфракрасных изображений. Микроволновые инструменты позволяют измерять температуру поверхности даже в присутствии облаков, но имеют ограничения по разрешению. Они дополнены измерениями инфракрасных датчиков и приведены в 5-дневных интервалах. Данные имеют разрешение примерно 14.98 км/пиксель и покрывают регион 30°–50° с.ш. и 55°–79° з.д.

Подробный анализ различия между модельными и спутниковыми композитными данными не входит в задачи данной статьи, однако важно отметить, что основное отличие с точки зрения алгоритмов распознавания заключается в наличии погрешностей измерений и сбойных пикселей в спутниковых композитных данных. В то же время, в модельных данных, являющихся результатом расчета, локальные тренды характеристик в любом направлении от заданного пикселя могут быть подсчитаны точно и не требуют применения статистических процедур.

На рис. 1 один и тот же регион представлен как модельными, так и композитными данными; видно, что локальное распределение температур в модельных данных существенно отличается от композитных. Это упрощает анализ модельных данных автоматическими методами на уровне отдельных пикселей, однако построение поля ТПО расчетным путем часто приводит к его сглажива-

нию, уменьшает абсолютные значения градиентов и тем самым затрудняет детектирование структурных элементов, в частности, фронтов.

Представленный алгоритм работает как с модельными, так и композитными данными.

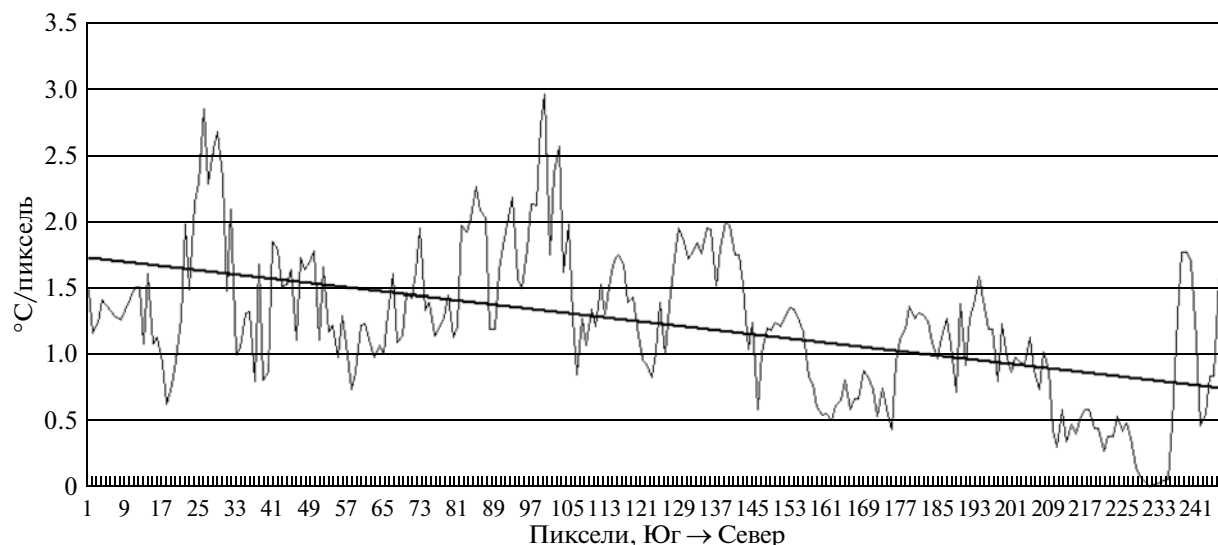
## МЕТОДОЛОГИЯ

Алгоритм основан на том, что, поскольку фронт разделяет водные массы с различными значениями температуры, то местоположение линии фронта совпадает с местоположением определенной изотермы (или участков нескольких последовательных изотерм).

Алгоритм распознает определенные изотермы как объекты, имеющие отношение к искомому фронту, в противоположность к построению искомой линии на основе локальных максимумов градиентов от точки к точке, что характерно для известных алгоритмов этого типа [4, 5, 10]. Наряду с плюсами нашего подхода (быстрота, простота, надежность), у него есть и минус — алгоритм требует предварительного, как правило, многократного, определения некоторых характеристик для каждого детектируемого фронта.

Алгоритм использует специфические особенности дискретных данных. В частности, тот факт, что разные изотермы могут проходить через один и тот же пиксель, в зависимости от разрешения данных и значения интервала температуры между ними.

Алгоритм состоит из трех основных шагов: нахождение точек начала и конца искомой фронталь-



**Рис. 2.** Типичный тренд абсолютных значений поперечных градиентов температуры вдоль северной границы Гольфстрима на примере композитного поля ТПО за 2.05.2011 г.

ной линии; выбор подходящих изотерм; построение фронтальной линии из найденных изотерм.

На первом шаге алгоритм находит точки локальных максимумов градиентов в области начала и конца фронтальной линии в соответствии с индивидуальными особенностями искомого фронта. В большинстве случаев это относительно простая задача, требующая, однако, индивидуального подхода к каждому конкретному фронту. Например, для рассматриваемого региона, типа данных и задачи детектирования северной границы Гольфстрима достаточно найти максимальные локальные градиенты температуры по широте вдоль границ региона (на  $55^\circ$  з.д. и  $79^\circ$  з.д.). В данном случае указанные максимумы находятся непосредственным расчетом по исходным данным без применения каких-либо специальных процедур. Для расчета достаточно взять соседние по широте пиксели для указанной долготы. В силу того, что этот шаг алгоритма носит индивидуальный характер для каждого конкретного фронта, в статье он подробно не рассматривается. Отметим лишь, что в данном случае выбор такого решения основан на анализе данных за продолжительный период. Были использованы еженедельные композитные данные за период с 6 сентября 2010 г. по 11 апреля 2011 г. для региона  $30^\circ$ – $50^\circ$  с.ш.,  $45^\circ$ – $80^\circ$  з.д. Подробное описание использованных для этих целей данных и экспертно-аналитических материалов представлено в [1]. Данные были приведены к разрешению, использованному в этой работе. В процессе анализа определено, что для указанных выше долгот ( $55^\circ$  з.д. и  $79^\circ$  з.д.) такой непосредственный расчет локальных максимумов градиентов дает надежные результаты. Для других участков региона подобный подход может

приводить к существенным ошибкам позиционирования фронта. В общем случае выбор метода определения максимумов градиентов в области начала и конца фронтальной линии требует предварительного анализа исходных данных с учетом особенностей конкретного фронта.

На втором шаге определяется множество изотерм, которые потенциально могут характеризовать искомую фронтальную линию. Это именно те изотермы, которые проходят через точки, найденные на первом шаге.

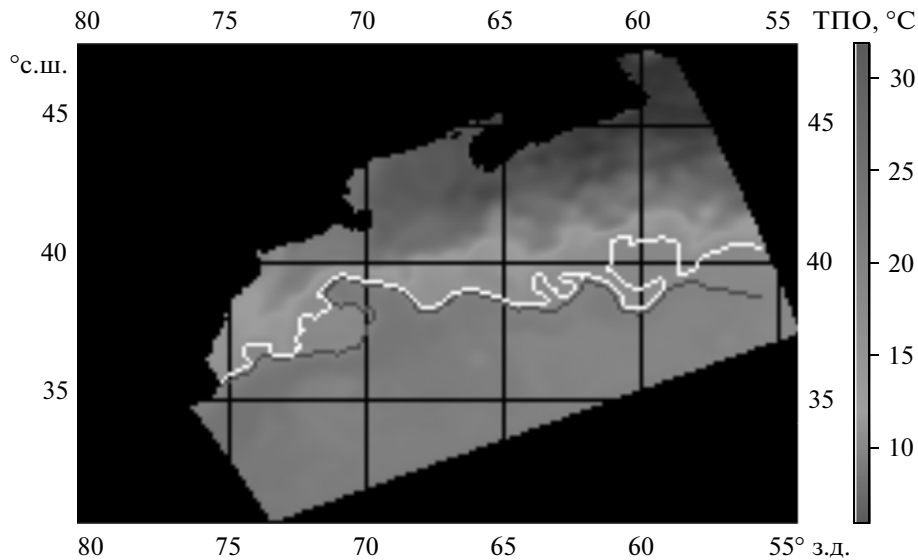
На третьем шаге из множества изотерм выбираются изотермы, характеризующие искомый фронт, и определяются точки, в которых линия фронта переходит с одной изотермы на другую. Фронтальная линия строится на основе выбранных изотерм.

## ПРЕДВАРИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

**Изотермы в дискретных данных.** В дискретном поле ТПО в каждой точке изотермы принципиально присутствует ошибка позиционирования (ОП), поскольку точка не может быть помещена между соседними пикселями.

1. Максимальная ОП (МОП) есть  $\frac{1}{2}$  расстояния между пикселями и, в терминах температуры, может быть оценена как  $|T_1 - T_2|/2$ , где  $T_1$  и  $T_2$  — значения температуры в соседних пикселях (поперек направления линии изотермы в данной точке).

Это означает, что МОП зависит от градиента температуры. Рис. 2 показывает типичный тренд абсолютных значений градиента вдоль северной границы Гольфстрима с юга на север. График по-



**Рис. 3.** Непересекающиеся изотермы с минимальным возможным интервалом температуры для данного разрешения, температуры, региона и типа данных. Модельное поле ТПО за 2.05.2011 г. Изотермы 17.2° (белая) и 19° С (серая). Разрешение – 14.98 км/пиксель. Интервал – 1.8°С. Сходные градации полутонов шкалы температур являются результатом преобразования исходных данных в черно-белый формат для целей данной статьи.

строен по композитным данным (см. выше). Из утверждения 1 можно заключить, что значения МОП покажут такой же уменьшающийся тренд. Алгоритм использует факт наличия этого тренда (см. описание алгоритма ниже).

2. Наличие такого тренда определяет применимость алгоритма к конкретному фронту.

Факт существования МОП говорит о том, что несколько соседних изотерм могут проходить через один и тот же пиксель. Если фронтальная линия характеризуется несколькими изотермами и представлена непрерывной линией, то в каждой ситуации перехода с изотермы на изотерму всегда найдется (с некоторыми ограничениями, см. ниже) как минимум один пиксель, общий для обеих соседних изотерм.

**МОП и разрешение.** МОП находится в обратной пропорциональной зависимости от DPI разрешения ( $R$ ). Если  $R \rightarrow 0$ , то  $\text{МОП} \rightarrow \infty$ . Действительно, представим себе, что разрешение имеет такое значение, что изображение умещается в одном пикселе; тогда все изолинии будут проходить через единственный пиксель. И, наоборот, для любого самого маленького интервала между значениями соседних изотерм всегда существует такое значение разрешения, что эти изотермы не будут иметь общих пикселей. На практике значения интервалов между соседними изотермами ограничены точностью данных. Например, рис. 3 и 4 показывают минимально возможный температурный интервал (МВТИ), при котором изотермы не имеют общих пикселей для данного разрешения, температуры, региона и типа данных.

На рис. 3 показано модельное поле ТПО с разрешением примерно 14.98 км/пиксель. Интервал между изотермами 1.8°С.

Пусть фронтальная линия характеризуется несколькими последовательными изотермами. Для любой фронтальной линии всегда можно определить два критических уровня разрешения: низкое DPI разрешение (НР), при котором все соответствующие изотермы проходят через одни и те же пиксели и высокое DPI разрешение (ВР), при котором соответствующие изотермы не имеют общих пикселей.

3. Из определения НР и ВР следует, что, если  $\text{НР} < R < \text{ВР}$ , существует хотя бы один пиксель, принадлежащий обеим изотермам.

Существует одно важное следствие из утверждения 3: в случае  $R = \text{НР}$  достаточно иметь хотя бы один “надежный” пиксель фронтальной линии, чтобы определить всю линию, выделив соответствующую изотерму; если  $\text{НР} < R < \text{ВР}$ , достаточно иметь хотя бы один “надежный” пиксель для каждой изотермы фронта, чтобы найти всю фронтальную линию путем выделения соответствующих изотерм и нахождения общих пикселей.

С другой стороны, если, для изотерм  $I$  и  $J$  и составной фронтальной линии  $K_{IJ}$ ,

$$\text{МВТИ}_{IJ} > \text{МАХ}(\text{МОП}_K), \quad (1)$$

существует по крайней мере один пиксель  $P_{IJ}$  такой, что  $P_{IJ} \in I$  и  $P_{IJ} \in J$ . В нашем примере (рис. 2) МВТИ в интересующей области составляет приблизительно 1.8°С, в то время как  $\text{МАХ}(\text{МОП})$



Рис. 4. Значения температуры вдоль изотерм 17.2 и 19°C. Линейный тренд показывает точность детектирования. Модельное поле ТПО за 2.05.2011 г. Разрешение – 14.98 км/пиксель.

северной границы Гольфстрима (см. определение МОП и рис. 2) – приблизительно 1.5°C. Это означает, что для данного разрешения, интервала температур, региона и типа данных изотермы, характеризующие данный фронт, имеют общие пиксели. Неравенство (1) может использоваться, чтобы определить применимость этого алгоритма к определенному полю ТПО. С другой стороны неравенство (1), вычисленное для всего региона, является более сильным, чем это может быть необходимо практически. Вычисления можно ограничить определенной областью, где изотермы могут иметь общие пиксели. В этом случае важно, чтобы ограничение, содержащееся в утверждении 2 было выполнено. Это означает, что, если неравенство (1) не выполнено для целой области, оно может быть выполнено для его определенной части.

#### ОПИСАНИЕ АЛГОРИТМА АГИ

Пусть для требуемой фронтальной линии определен начальный пиксель  $S$  и конечный пиксель  $E$ . В соответствии с концепцией МОП имеются два множества изотерм:  $I_S$  и  $I_E$ , таких, что  $\forall i (S \in I_{Si})$  и  $\forall j (E \in I_{Ej})$ , где  $i$  и  $j$  – номера изотерм в соответствующих множествах.

Если  $\forall i$  и  $\forall j (I_{Si} \cap I_{Ej}) = \{\emptyset\}$ , то данная фронтальная линия не может быть построена из 2-х изотерм и потребуется дополнительный (промежуточный) “надежный” пиксель. Для простоты примем, что искомая фронтальная линия характеризуется позиционированием двух изотерм; данное описание может быть легко обобщено для большего количества изотерм путем их последовательного попарного рассмотрения.

Пусть  $I_G = I_S \cap I_E \neq \{\emptyset\}$ ;

пусть  $P_J$  есть множество пикселей изотермы  $I_{GJ} \in I_G$ ;

пусть  $P = \cup_K P_K$ ,  $P = \{p_1, \dots, p_n\}$ ;

пусть  $X_m, Y_m$  есть координаты пикселя  $p_m$ ;

пусть  $D_m$  есть расстояние между пикселями  $p_m$  и  $E$ .

4. Пусть  $M_L = \text{MIN}(D_n)$ , где  $L$  – индекс пикселя в множестве  $P$ ;

тогда  $p_L$  будет пикселем, в котором происходит смена изотерм, определяющих позиционирование данной фронтальной линии.

5. Пусть  $i$  есть номер изотермы в множестве; пусть  $I_{GS}$  есть множество изотерм, таких, что  $(\forall I_{GSi} \in I_{GS}) (p_L \in I_{GSi})$ ;

пусть  $I_{GE}$  есть множество изотерм, таких, что  $(\forall I_{GEi} \in I_{GE}) (p_L \in I_{GEi})$ ;

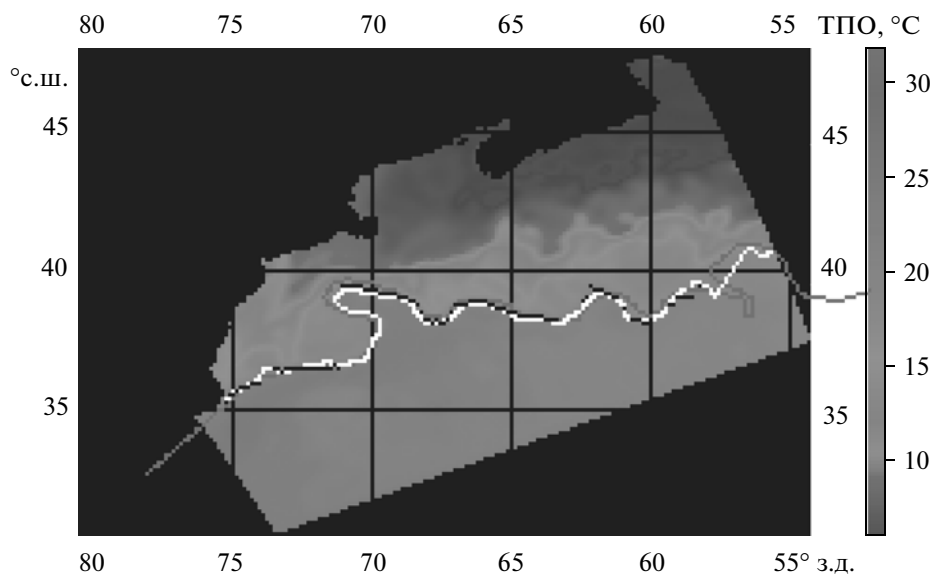
пусть  $I_{GS}^* \in I_{GS}$  есть изотерма с максимальным (минимальным для холодных течений) значением температуры;

пусть  $I_{GE}^* \in I_{GE}$  есть изотерма с максимальным (минимальным для холодных течений) значением температуры.

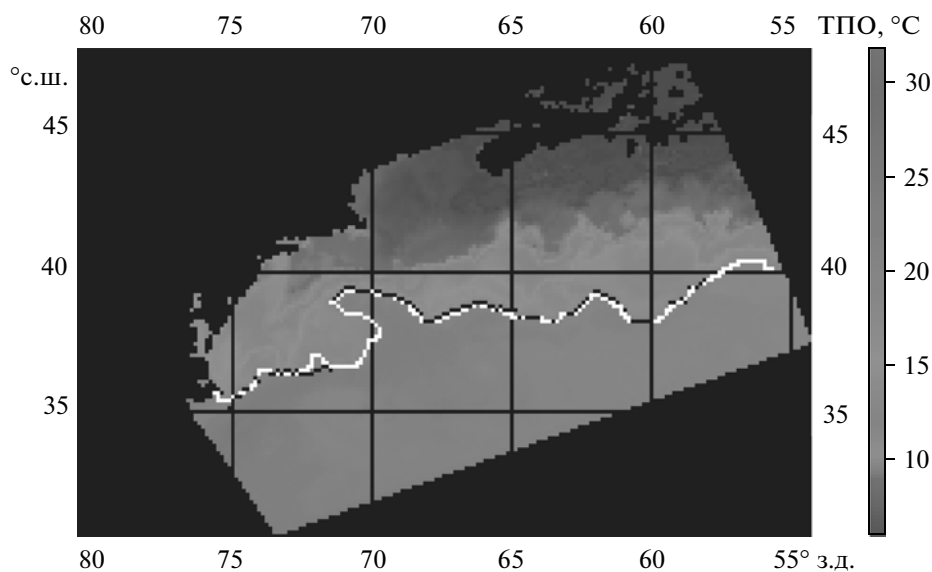
Построим изотерму  $I_{GS}^*$  от пикселя  $S$  до пикселя  $p_L$ ; построим изотерму  $I_{GE}^*$  от пикселя  $p_L$  до пикселя  $E$ .

Требуемая фронтальная линия построена. Она состоит из сегментов двух изотерм, имеющих один общий пиксель. В случае  $I_{GS}^* = I_{GE}^*$  примем  $p_L = E$ .

“Надежные” начальный и конечный пиксели фронтальной линии обычно определяются исхо-



**Рис. 5.** Северная граница Гольфстрима по результатам детектирования. Модельное поле ТПО за 2.05.2011 г. Разрешение – 14.98 км/пиксель. Белая линия показывает искомую фронтальную линию, серая – представленную экспертом-аналитиком; черные пиксели показывают локальные максимумы градиентов. Сходные градации полутонов шкалы температур являются результатом преобразования исходных данных в черно-белый формат для целей данной статьи.



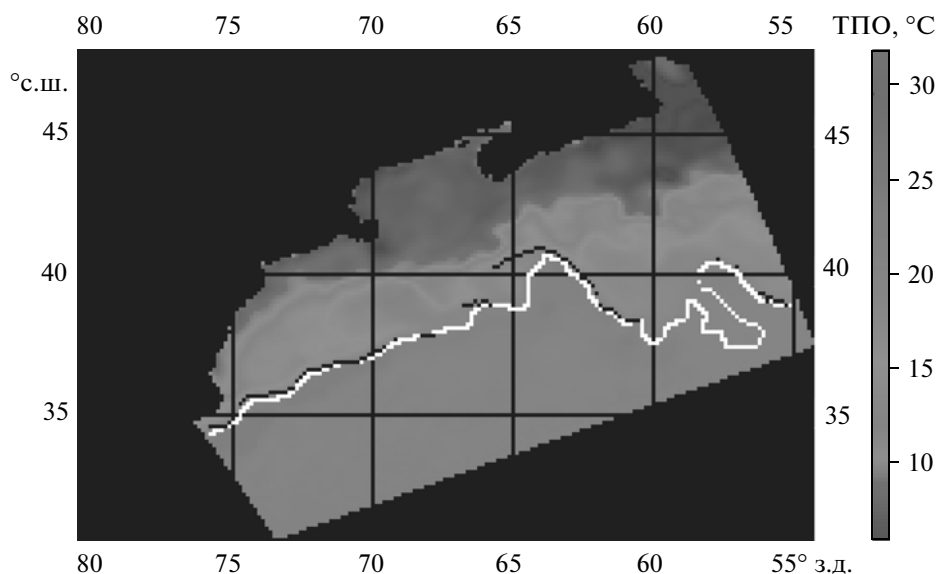
**Рис. 6.** Северная граница Гольфстрима по результатам детектирования. Композитное поле ТПО за 2.05.2011 г. Применен медианный фильтр. Разрешение – 14.98 км/пиксель. Белая линия показывает искомую фронтальную линию; черные пиксели показывают локальные максимумы градиентов. Сходные градации полутонов шкалы температур являются результатом преобразования исходных данных в черно-белый формат для целей данной статьи.

для из конкретного фронта, разрешения, типа исходных данных и региона на основе выявления пикселей с максимальными градиентами, отвечающих определенным условиям.

Выбор изотермы, ближайшей к центру потока течения (утверждение 5), образующего искомый фронт, гарантирует, что выбрана граница именно этого течения – например, в нашем случае, север-

ная граница Гольфстрима, а не граница склоновой водной массы.

Выбор общего для изотерм пикселя, наиболее близкого к “конечной точке”  $E$ , в соответствии с утверждением 4 обеспечивает оптимальное позиционирование фронтальной линии по отношению к “начальному пикселю”  $S$ , который имеет большее значение МОП. С учетом утверждения 5 это обеспечивает более надежный результат.



**Рис. 7.** Северная граница Гольфстрима по результатам детектирования. Модельное поле ТПО за 13.12.2010. Разрешение – 14.98 км/пиксель. Белая линия показывает искомую фронтальную линию; черные пиксели показывают локальные максимумы градиентов. Сходные градации полутонов шкалы температур являются результатом преобразования исходных данных в черно-белый формат для целей данной статьи.

Построение изотерм в дискретных данных является относительно простой процедурой, поскольку подразумевает поиск на основе параметра, являющегося константой, в то время как градиенты демонстрируют существенную изменчивость вдоль фронтальной линии.

#### ОЦЕНКА ОШИБКИ ПОЗИЦИОНИРОВАНИЯ

**Ошибка позиционирования изотермы.** Для изотермы  $I$  со значением температуры  $T$  и пикселя  $P_j \in I$

$\Delta t_j = |t_j - T|$ , где  $t_j$  – значение температуры для пикселя  $P_j$ ;

абсолютная ошибка (км):  $E_j = (\Delta t_j / \text{MPE}) * (\text{Res} / 2)$ , где  $\text{Res}$  – разрешение (км/пиксель);

средняя ошибка для изотермы  $I$  (км):  $AE_I = (\sum_{j=1, n} E_j) / n$ , где  $n$  – количество пикселей в линии.

Например, для изотермы 19°C (рис. 3) средняя ошибка позиционирования составляет примерно 4.42 км, в то время как разрешение изображения составляет примерно 14.98 км/пиксель, или около 30% расстояния между соседними пикселями.

**Ошибка позиционирования фронтальной линии.** Оценка может быть сделана на основе наличия областей с относительно большими градиентами, расположенными вдоль фронтальной линии. Также можно воспользоваться наличием “надежных” пикселей, найденных с помощью анализа градиентов в отдельных регионах, через которые

проходит фронтальная линия, если таковые могут быть найдены. В других регионах ошибка может быть оценена на основе оценки ошибки позиционирования соответствующих изотерм (см. выше).

#### ВЫВОДЫ

Алгоритм тестировался на модельном поле ТПО для ежедневных изображений с 02 мая 2011 г. по 09 мая 2011 г. и с 13 декабря 2010 г. по 20 декабря 2010 г. Алгоритм также тестировался на 5-дневном композитном поле ТПО за тот же период. Предварительно применялся медианный фильтр. Регион: 30°–50° с.ш. и 55°–79° з.д. Искомая фронтальная линия: северная граница Гольфстрима. Примеры результатов представлены на рис. 5, 6 и 7. Белая линия показывает искомую фронтальную линию; серая линия на рис. 5 – представленную экспертом-аналитиком; черные пиксели показывают локальные максимумы градиентов. Алгоритм демонстрирует хорошее соответствие результатов экспертно-аналитическим материалам и высокую надежность. Небольшой объем вычислений и высокая скорость работы являются следствием относительной простоты алгоритма.

В перспективе детектирование таким способом южной границы Гольфстрима и границы склоновой водной массы может быть использовано в качестве исходных данных для алгоритма автоматического распознавания океанических структур на картах фронтов [1].

Автор благодарит А. Гангопадхья за иницирование данного проекта и конструктивное сотрудничество в процессе создания алгоритма.

### СПИСОК ЛИТЕРАТУРЫ

1. *Давидович В., Гангопадхья А.* Алгоритм автоматического распознавания океанических структур на картах фронтов // *Океанология*. 2012. Т. 52. № 3. С. 462–472.
2. *Федоров К.Н.* Физическая природа и структура океанических фронтов. Л.: Гидрометеиздат, 1983. 296 с.
3. *Belkin I.M.* Front. *Interdisciplinary Encyclopedia of Marine Sciences* / Eds. Nybakken et al. Grolier, Danbury, CT. 2002. P. 433–436.
4. *Cayula J.F., Cornillon P.* Edge detection Algorithm for SST images // *J. Atmos. Oceanic Technol.* 1992. V. 9. № 1. P. 67–80.
5. *Cayula J.F., Cornillon P.* Multi-image edge detection for SST images // *J. Atmos. Oceanic Technol.* 1995. № 12. P. 821–829.
6. *Cornillon P., Watts D.R.* Satellite thermal infrared and inverted echo sounder determinations of the Gulf Stream northern edge // *J. Atmos. Oceanic Technol.* 1987. № 4. P. 712–723.
7. *Gangopadhyay A., Robinson A.R.* Feature oriented regional modeling of Oceanic Fronts // *Dynamics of Atmosphere and Oceans*. 2002. V. 36. № 1–3. P. 201–232.
8. *Gangopadhyay A., Robinson A.R., Arango H.G.* Circulation and Dynamics of the Western North Atlantic, I: Multiscale Feature Models // *J. Atmos. Oceanic Technol.* 1997. V. 14. № 6. P. 1314–1332.
9. *Holyer R.J., Peckinpaugh S.H.* Edge detection applied to satellite imagery of the oceans // *IEEE Trans. Geosci. Remote Sens.* 1989. № 27. P. 46–56.
10. *Ullman, D.S., Cornillon P.C.* Evaluation of front detection methods for satellite-derived SST data using in situ observations // *J. Atmos. Oceanic Technol.* 2000. № 17. P. 1667–1675.

## Algorithm of Automatic Recognition of Temperature Oceanic Fronts

V. A. Davidovich

The recognition of fronts in describing ocean circulation has been long appreciated. Advent of satellite oceanography led to developing methods for automatic front detection using satellite composite SST data. In this work we present an algorithm for frontline detection based on gradients/isotherms analyses. The algorithm detects frontline as a sequence of isotherms having some special characteristics. This approach makes it possible to detect an individual frontlines, for instance, slope water edge, or northern edge of Gulf Stream directly, without preliminary detection of all frontlines in the system. The algorithm combines simplicity of implementation and high computational efficiency with good accuracy and high reliability.