

Variational Data Assimilation of Satellite Observations in the Model of Sea Hydrothermodynamics

Eugene I. Parmuzin^{1,2}, Valery I. Agoshkov^{1,3}, Natalia B. Zakharova¹, Victor P. Shutyaev^{1,2}

¹Marchuk Institute of Numerical Mathematics of the Russian Academy of Sciences, Moscow, Russia

²Moscow Institute of Physics and Technology, Dolgoprudny, Russia

³Lomonosov Moscow State University, Moscow, Russia

Corresponding author: eparmuzin@gmail.com

Abstract

Recently, significant results have been achieved in studying and modeling the processes of large-scale sea and ocean variability. This is due to changes in the level of development of tools, methods and equipment. These include: new observational systems (satellites, ARGO buoys, etc.), information analysis methods and numerical algorithms, powerful computers that allow processing large information flows of calculations and observations. We consider the problem of four-dimensional variational data assimilation of satellite data on the sea surface temperature using a model of sea hydrothermodynamics developed at INM RAS in this paper. Problem state and methods for its solution are discussed, the results of numerical experiments with real data of satellite observations are presented, including cases when the observation data are known at the part of the water area.

Keywords: variational data assimilation, satellite observations, sea surface temperature, heat flux, sea hydrodynamics model

ВАРИАЦИОННАЯ АССИМИЛЯЦИЯ ДАННЫХ СПУТНИКОВЫХ НАБЛЮДЕНИЙ В МОДЕЛИ ГИДРОТЕРМОДИНАМИКИ МОРЯ

Е.И. Пармузин^{1,2}, В.И. Агошков^{1,3}, Н.Б. Захарова¹, В.П. Шутяев^{1,2}

¹Институт вычислительной математики им. Г.И. Марчука РАН, Москва, Россия

²Московский физико-технический институт, Долгопрудный, Россия

³Московский государственный университет имени М.В. Ломоносова, Москва, Россия

email: eparmuzin@gmail.com

За последнее время в изучении и моделировании процессов крупномасштабной морской и океанской изменчивости достигнуты существенные результаты. Это связано с изменениями в уровне развития инструментария, методов и оборудования. К ним относятся: новые наблюдательные системы (спутники, буи ARGO и др.), методы анализа информации и численные алгоритмы; мощные компьютеры, позволяющие обрабатывать большие информационные потоки расчетов и наблюдений. В настоящей работе рассматривается задача четырехмерной вариационной ассимиляции данных со спутников о температуре поверхности моря на примере модели гидротермодинамики моря, разработанной в ИВМ РАН. Дано описание постановки задачи и методов ее решения, приводятся результаты численных экспериментов с реальными данными спутниковых наблюдений, в том числе для случаев, когда данные наблюдений известны только на части рассматриваемой акватории.

Ключевые слова: вариационная ассимиляция данных, спутниковые наблюдения, температура поверхности моря, поток тепла, модель гидротермодинамики моря.

Общая постановка задачи ассимиляции данных

Рассмотрим математическую модель, описывающую эволюцию гидродинамической системы в виде

$$\begin{cases} \frac{dx}{dt} = M(x, t), & t > 0 \\ x|_{t=0} = x_0, \end{cases}$$

где x – вектор состояния модели, M – динамический оператор модели, x_0 – вектор начального состояния.

Наблюдения задаются некоторой вектор-функцией $y^0(t)$, которая удовлетворяет уравнению

$$y^0(t) = H(x^t, t) + \varepsilon,$$

где H – оператор наблюдений, x^t – истинное поле, ε – функция ошибки (шум). Функция $y^0(t)$ считается заданной, в то время как информация об ε , как правило, отсутствует.

Для предсказания эволюции в задачах геофизической гидродинамики требуется дополнительная информация о модели (например, начальные условия, неизвестные параметры модели). Эту информацию можно получить с помощью данных наблюдений. Для этого формулируется **задача ассимиляции данных**: при заданной функции наблюдений $y^0(t)$ требуется найти неизвестные параметры модели (например, начальное условие), так, чтобы вектор состояния x удовлетворял исходной задаче, а вектор $H(x)$ был близок в каком-либо

смысле к $y^0(t)$. Найденное в результате решение x называется оценкой состояния и обозначается x^a .

Методы ассимиляции данных

Принято выделять два класса методов ассимиляции данных - статистические и вариационные методы. Опишем несколько методов ассимиляции данных, применяемых в задачах моделирования и прогнозирования состояния морской среды. Одним из первых таких методов являлся метод *объективного анализа*. Первая попытка объективного анализа данных была выполнена Пановски (Panosky, 1949). Суть его метода – двумерная (2-D) полиномиальная интерполяция данных наблюдений (Кибель, 1949). В дальнейшем этот подход был развит Гилкристом и Крессманом (Gilchrist, Cressman, 1954), которые ввели область влияния для каждого наблюдения и предложили использовать так называемое поле “бэкграунда” (поле из предыдущего прогноза). В подходе Бергторссона и Дуза (Bergthorsson, Doos, 1955) поле “бэкграунда” играет более важную роль – их методика ассимиляции основана на анализе разности данных наблюдений и “бэкграунда”, а не самих значений функции наблюдений. Впоследствии модификация этого подхода была дана Крессманом (Cressman, 1959) и состояла в нескольких итерациях анализа – так называемый метод последовательных поправок, или SCM-метод (Successive Correction Method).

Следующий метод в ряду методов ассимиляции данных - это метод *оптимальной интерполяции*. В данном методе наблюдениям присваивают веса, которые связаны с ошибками наблюдений. Поле “бэкграунда” является дополнительным полезным источником информации вместе со своей характеристикой ошибки. Этот подход восходит к Колмогорову (Колмогоров, 1941), работам Винера (Wiener, 1949), а в науках о Земле он стал известен благодаря монографии Гандина (Гандин, 1963). Метод оптимальной интерполяции и его модификации до настоящего времени наиболее широко используются для оперативного анализа данных при предсказании погоды (Lorenz, 1986; Thiebaux, Pedder 1987; Douville et al., 2000), а также при ассимиляции океанографических данных (Carton, Hackert, 1989; Derber, Rosati, 1989; Smith, Cummings, 2012). Большую популярность приобрел метод ансамблевой оптимальной интерполяции (EnOI) (Evensen, 2003; Sakov et al., 2015), который позволяет построить параллельные алгоритмы ассимиляции данных (Кауркин, Ибраев, Беляев, 2016).

В настоящее время самым востребованным статистическим методом ассимиляции данных является метод, называемый *фильтром Калмана*. Фильтр Калмана оперирует понятием вектора состояния системы (набором параметров, описывающих состояние системы на некоторый момент времени) и его статистическим описанием. Существует значительное число модификаций данного метода, так можно выделить расширенный фильтр Калмана - EKF (Ghil et al, 1982; Budgell, 1986), ансамблевый фильтр Калмана - EnKF (Evensen, 2003, 2007; Kalnay et al., 2007; Fertig et al., 2007; Zhang et al., 2009) и некоторые другие. Данный метод нашел широкое применение в моделировании состояния окружающей среды. Так, многоэлементный четырехмерный анализ гидрофизических полей на основе динамико-стохастических моделей разрабатывался в МГИ (Саркисян, Кныш, Демьшев, Коротаев, 1986, 1987). Модификации алгоритма Калмана на основе аппроксимаций ковариационных матриц использовались при моделировании циркуляции Черного моря (Кныш, Коротаев, Мизюк, Саркисян, 2012).

Значительным прорывом в решении задач ассимиляции данных было применение вариационных методов и, в частности, методов оптимального управления. Очень плодотворной оказалась идея минимизировать некоторый функционал, связанный с данными наблюдений, на траекториях (решениях) рассматриваемой модели. Тем самым, задача ассимиляции данных формулируется как задача оптимального управления. Теоретические основы исследования и решения таких задач заложены в классических работах (Беллман, 1957; Понтрягин, 1962; Красовский, 1969; Лионс, 1968; Марчук, 1975). Впервые вариационный формализм был использован в метеорологии Сасаки (Sasaki, 1970), а в задачах динамической океанографии – Прово и Сальмоном (Provost, Salmon, 1986). Как известно, при решении задач минимизации

возникает необходимость вычислять градиент исходного функционала. Важным шагом в этом направлении было использование теории сопряженных уравнений (Марчук, 1964; Лионс, 1968). Начиная с известных работ (Марчук, Пененко, 1978; Le Dimet, Talagrand, 1986; Lewis, Derber, 1985), применение сопряженных уравнений для исследования и численного решения задач ассимиляции данных (в том числе для вычисления градиента функционала) широко практикуется многими исследователями (Courier, Talagrand, 1987; Lorenc, 1988; Navon, 1986; Агошков, Марчук, 1993; Марчук, Залесный, 1993; Венцель, Залесный, 1996; Шутяев, 2001; Агошков, Пармузин, Шутяев, 2008, 2013, и др.). Первые применения трехмерной вариационной ассимиляции данных (3D-VAR) для операционного анализа были сделаны в Национальном Центре предсказаний NCEP (Parrish, Derber, 1992), а позднее в Европейском Центре прогноза погоды ECMWF и Центр ассимиляции данных NASA (Cohn et al, 1998). В настоящее время все больший интерес вызывает четырехмерная ассимиляция данных (4D-VAR), при которой линеаризованные модели и сопряженные к ним используются для ассимиляции данных наблюдений не в конкретный момент времени, а на заданном временном интервале. Впервые система 4D-VAR была применена в Европейском Центре прогноза погоды (Courtier et al, 1994).

В настоящей работе рассматривается задача 4-мерной вариационной ассимиляции данных на примере модели гидротермодинамики моря. Ниже сформулируем задачу вариационной ассимиляции данных.

Вариационная ассимиляция данных наблюдений

Рассмотрим задачу на интервале $(0, \bar{t})$ и введем функционал от ее решения

$$J(x_0) = \frac{1}{2} (C_1(x_0 - x_0^b), x_0 - x_0^b) + \frac{1}{2} \int_0^{\bar{t}} (C_2(Hx - y^0), Hx - y^0) dt,$$

где H – (линейный) оператор наблюдений, y^0 – функция наблюдений, x_0^b – заданный вектор, C_1, C_2 – весовые операторы, (\cdot, \cdot) – скалярное произведение. Как правило, C_1, C_2 выбираются в виде: $C_1 = B^{-1}$, $C_2 = R^{-1}$, где B, R – ковариационные матрицы векторов $\xi = x_0^b - x^t|_{t=0}$ и ε , соответственно: $B = E(\xi\xi^T)$, $R = E(\varepsilon\varepsilon^T)$.

Предположим, что начальное условие x_0 нам неизвестно. Тогда задача об ассимиляции данных формулируется следующим образом: найти x_0, x такие, что они удовлетворяют системе уравнений и на множестве решений функционал J достигает своего наименьшего значения:

$$\begin{cases} \frac{dx}{dt} = M(x, t), t \in (0, \bar{t}), & x|_{t=0} = x_0 \\ J(x_0) = \inf_v J(v). \end{cases}$$

По определению градиента

$$J' \delta x_0 = (C_1(x_0 - x_0^b), \delta x_0) + \int_0^{\bar{t}} (C_2(Hx - y^0), H \delta x) dt,$$

где δx удовлетворяет системе TLM (tangent linear model) для $t \in (0, \bar{t})$:

$$\frac{d\delta x}{dt} = M'(x, t) \delta x, \quad \delta x|_{t=0} = \delta x_0.$$

Введем сопряженную задачу

$$-\frac{dx^*}{dt} = (M'(x, t))^* x^* - p, \quad x^*|_{t=\bar{t}} = 0,$$

где $p = H^* C_2(Hx - y^0)$. Тогда из соотношения сопряженности $\int_0^{\bar{t}} (p, \delta x) dt = -(x^*|_{t=0}, \delta x_0)$

получаем градиент

$$J' \delta x_0 = (C_1(x_0 - x_0^b), \delta x_0) + \int_0^{\bar{t}} (p, \delta x) dt = (C_1(x_0 - x_0^b) - x^*|_{t=0}, \delta x_0).$$

Необходимое условие оптимальности (Lions, 1968) приводит задачу к системе для трех неизвестных x_0, x, x^* :

$$\begin{cases} \frac{dx}{dt} = M(x, t), & t \in (0, \bar{t}) \\ x|_{t=0} = x_0, \end{cases} \quad \begin{cases} -\frac{dx^*}{dt} = (M'(x, t))^* x^* - H^* C_2(Hx - y^0), \\ x^*|_{t=\bar{t}} = 0, \\ C_1(x_0 - x_0^b) - x^*|_{t=0} = 0, \end{cases}$$

где $(M'(x, t))^*$ – оператор, сопряженный к производной оператора модели M . Эта система может быть получена также из принципа максимума Понтрягина, сформулированного для исходной задачи минимизации (Marchuk, Zalesny, 1993), или методом множителей Лагранжа (Евтушенко и др., 1997).

Модель гидротермодинамики моря

В качестве примера использования метода вариационной ассимиляции в задачах моделирования акваторий океанов и морей рассмотрим следующую систему. Запишем систему уравнений гидротермодинамики в “приближении Буссинеска и гидростатики”, но сохраняя коэффициенты Ламе соответствующими сферической системе координат [1, 2]. Получаем следующую систему уравнений для функций:

$$\begin{cases} \frac{d\bar{u}}{dt} + \begin{bmatrix} 0 & -f \\ f & 0 \end{bmatrix} \bar{u} - g \cdot \text{grad} \xi + A_u \bar{u} + (A_k)^2 \bar{u} = \bar{f} - \frac{1}{\rho_0} \text{grad} P_a - \frac{g}{\rho_0} \text{grad} \int_0^z \rho_1(T, S) dz', \\ \frac{\partial \xi}{\partial t} - m \frac{\partial}{\partial x} \left(\int_0^H \Theta(z) u dz \right) - m \frac{\partial}{\partial y} \left(\int_0^H \Theta(z) \frac{n}{m} v dz \right) = f_3, \quad \frac{dT}{dt} + A_T T = f_T, \quad \frac{dS}{dt} + A_S S = f_S, \end{cases}$$

где $\bar{f} = g \cdot \text{grad} G$, $r = R - z$, $0 < z < H$, f_T, f_S, f_P – заданные функции “внутренних” источников, $\rho_1(T, S) = \rho_0 \beta_T (T - T^{(0)}) + \rho_0 \beta_S (S - S^{(0)}) + \gamma \rho_0 \beta_{TS} (T, S) + f_P$, $g = \text{const} > 0$, $\rho_0, T^{(0)}, S^{(0)}$ – “невозмущенные” значения плотности воды, температуры, солёности, β_T, β_S – коэффициенты $\beta_{TS}(T, S)$, $P_a, f_3 \equiv f_3(x, y, \xi, t) \equiv f_3(x, y, t)$ – заданные функции, а γ – числовой параметр, $\Theta(z) \equiv r(z)/R$. Граничные условия имеют вид:

$$\begin{aligned} & \left(\int_0^H \Theta \bar{u} dz \right) \bar{n} + \beta_0 m_{op} \sqrt{gH} \xi = m_{op} \sqrt{gH} d_s \text{ на } \partial\Omega, \\ & U_n^{(-)} u - \nu \frac{\partial u}{\partial z} - k_{33} \frac{\partial}{\partial z} A_k u = \tau_x^{(a)} \rho_0, \quad U_n^{(-)} v - \nu \frac{\partial v}{\partial z} - k_{33} \frac{\partial}{\partial z} A_k v = \tau_y^{(a)} \rho_0, \quad A_k u = 0, \quad A_k v = 0, \\ & U_n^{(-)} T - \nu_T \frac{\partial T}{\partial z} + \gamma_T (T - T_a) = Q_T + U_n^{(-)} d_T, \quad U_n^{(-)} S - \nu_S \frac{\partial S}{\partial z} + \gamma_S (S - S_a) = Q_S + U_n^{(-)} d_S, \end{aligned}$$

где $\bar{U} = (u, v, w) \equiv (\bar{u}, w)$, $U_n^{(-)} = (|U_n| - U_n)/2$. Граничные функции d_T, d_S, Q_T, Q_S тоже могут быть неизвестными. Если u, v, ξ, T, S найдены, то функции w и P определяются по формулам:

$$w(x, y, z, t) = \frac{1}{r} \left(m \frac{\partial}{\partial x} \left(\int_z^H r u dz' \right) + m \frac{\partial}{\partial y} \left(\frac{n}{m} \int_z^H r v dz' \right) \right), \quad (x, y, t) \in \Omega \times (0, \bar{t}),$$

$$P(x, y, z, t) = P_a(x, y, t) + \rho_0 g (z - \xi) + \int_0^z g \rho_1(T, S) dz'.$$

Вариационная ассимиляция температуры поверхности моря

Пусть дополнительной неизвестной ("управлением") является функция полного потока Q . Введем функционал стоимости вида:

$$J_\alpha \equiv J_\alpha(Q, \phi) = \frac{1}{2} \int_0^{\bar{t}} \int_\Omega \alpha |Q - Q^{(0)}|^2 d\Omega dt + J_0(\phi), \quad J_0(\phi) = \frac{1}{2} \int_0^{\bar{t}} \int_\Omega m_0 |T - T_{obs}|^2 d\Omega dt.$$

Здесь: $\alpha \equiv \alpha(\lambda, \theta, t)$ – функция, играющая роль "регуляризатора" или "штрафной функции" (возможен случай, когда $\alpha(\lambda, \theta, t) = \text{const} \geq 0$) и которая может быть размерной величиной, T_{obs} – данные наблюдений, а $Q^{(0)} \equiv Q^{(0)}(\lambda, \theta, t)$ – заданная функция (которая может быть также и тривиальной). Задача вариационной ассимиляции формулируется следующим образом: *требуется найти решение $\phi = (u, v, \xi, T, S)$ и функцию Q , такие, чтобы на них функционал принимал наименьшее значение.*

В работах [3, 4] показано, что решение системы оптимальности, которая определяет решение сформулированной задачи вариационной ассимиляции данных, сводится к последовательному решению некоторых подзадач на $t \in (t_{j-1}, t_j)$, $j = 1, 2, \dots, J$.

Рассмотренный алгоритм вариационной ассимиляции данных спутниковых наблюдений применялся для решения задачи вариационной ассимиляции данных о температуре поверхности моря с целью восстановления потоков тепла на поверхности.

Обработка данных наблюдений со спутников

В задачах вариационной ассимиляции данных важную роль играют используемые данные и их качество. От обработки данных наблюдений о состоянии исследуемых сред зависит численное решение всей задачи математического моделирования. Под обработкой понимаются получение данных, интерполяция на регулярные сетки и расчетные сетки численной модели, а также верификация получаемых данных. Для указанных задач применяются современные технологии по работе с большими данными, позволяющие эффективно обрабатывать и анализировать гидрофизические данные наблюдений [5].

Алгоритмы для обработки данных позволяют также вычислять статистические характеристики исследуемых полей. На основе многолетних данных со спутников о среднесуточной температуре поверхности моря вычисляются статистическое усреднение и среднеквадратическое отклонение. Указанные статистические характеристики используются как коэффициенты в ковариационных матрицах ошибок наблюдений, обратные к которым включаются в качестве весовых операторов в исходный функционал стоимости при решении задачи вариационной ассимиляции данных [2].

Численные эксперименты по вариационной ассимиляции данных

Для экспериментов использовалась разработанная в ИВМ РАН глобальная трехмерная модель гидротермодинамики океана [3]. Модель была дополнена блоком ассимиляции данных наблюдений о температуре поверхности моря T_{obs} . В качестве объекта моделирования рассматривается акватория Черного моря. Параметры рассматриваемой области и ее географические координаты могут быть описаны следующим образом: сетка $286 \times 159 \times 27$ точки (широта \times долгота \times глубина). Шаги сетки по координатам x и по y постоянны и равны соответственно 0.05° и 0.04° . Шаг по времени $\Delta t = 5$ минут. В качестве T_{obs} использовались данные со спутников о температуре поверхности Черного моря за 2008 г. В качестве $Q^{(0)}$ использовался среднеклиматический поток за январь, полученный по данным реанализа NCEP (National Centers for Environmental Prediction), а также поток, вычисляемый в прямой модели динамики океана. Ниже проиллюстрирован результат ассимиляции в модели гидротермодинамики спутниковых данных наблюдений.

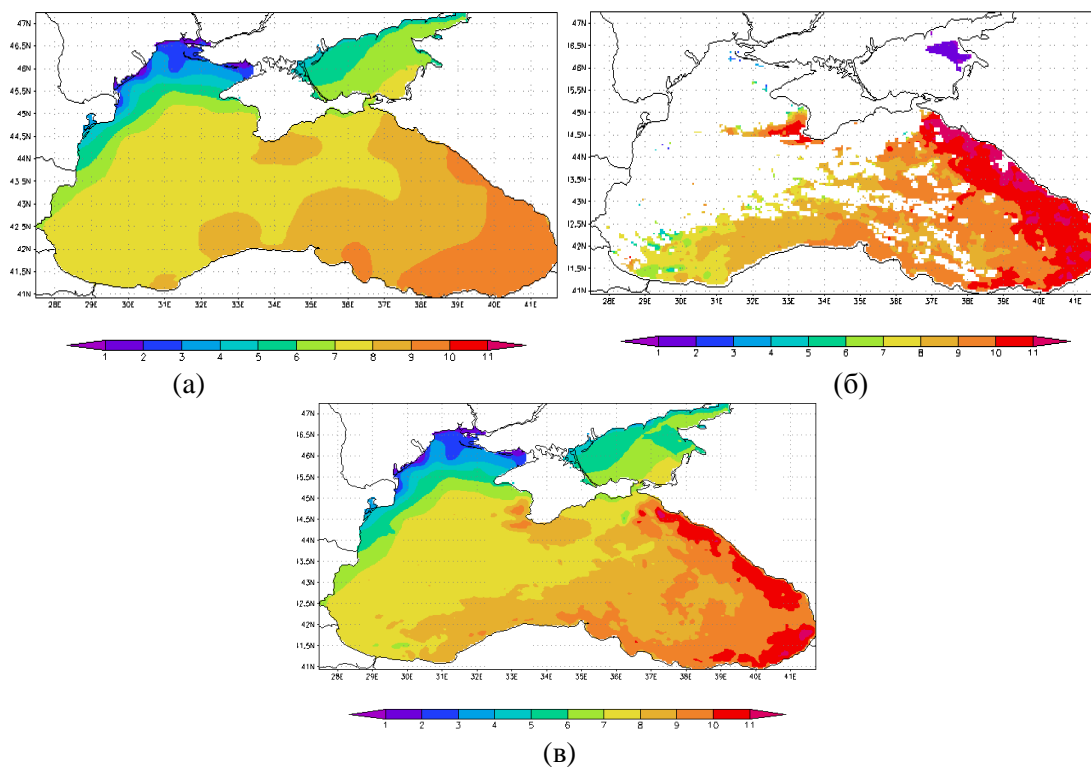


Рис. 1. Температура поверхности моря (среднесуточная): (а) расчет модели до ассимиляции данных; (б) данные наблюдений; (в) расчет модели после ассимиляции данных наблюдений

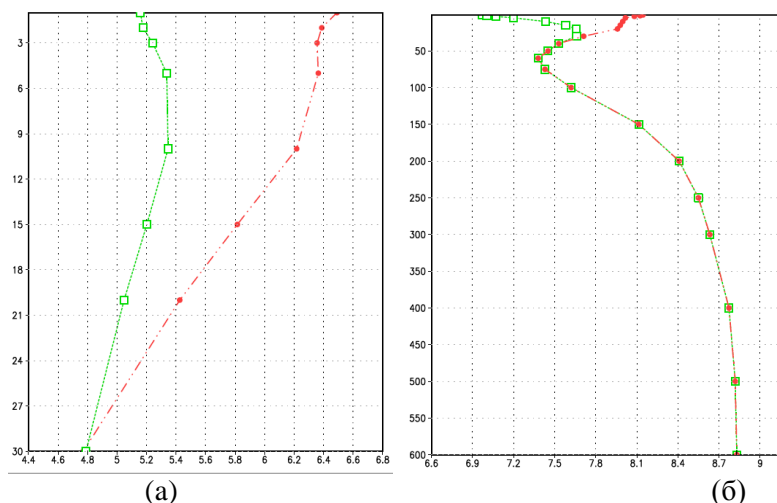


Рис. 2. Профиль температуры. Пунктирная линия с квадратами – расчет по модели до ассимиляции данных, пунктирная линия с точками – расчет по модели после ассимиляции данных наблюдений: (а) точка с координатами 30°в.д., 45°с.ш., (б) точка с координатами 32°в.д., 44°с.ш.

На рис. 1 представлены поля температур поверхности моря (среднее за день), рассчитанные с помощью численной модели на 1 февраля 2008 г., данные наблюдений за 1 февраля 2008 г. и расчет модели после ассимиляции этих данных. Легко заметить, что модель ассимилирует данные и приближает решение модели к данным наблюдений.

На рис. 2 изображены профили температуры в различных точках Черного моря в один из моментов времени 1 февраля 2008 года. На каждом из графиков различными маркерами обозначены профили температуры из численной модели гидротермодинамики до и после ассимиляции данных.

Как показали численные эксперименты, рассмотренный выше алгоритм ассимиляции температуры позволяет получать более точное поле температуры на поверхности моря по сравнению с результатами, когда процесс ассимиляции не применяется.

Заключение

В настоящей работе проведен обзор методов ассимиляции данных, как статистических, так и вариационных. Дана общая постановка метода вариационной ассимиляции данных. Алгоритм вариационной ассимиляции данных рассмотрен на примере решения задачи о восстановлении потока тепла при ассимиляции данных наблюдений за температурой поверхности моря. Сформулирован численный алгоритм решения задачи вариационной ассимиляции данных спутниковых наблюдений, который был применен для численных расчетов в акватории Черного моря. Предложенный алгоритм ассимиляции температуры позволяет получать более точное поле температуры поверхности моря по сравнению с результатами, когда процесс ассимиляции не применяется.

Работа выполнена при финансовой поддержке РФФ (проект № 19-71-20035, постановка задачи ассимиляции и численные эксперименты), РФФИ (проект № 19-01-00595) и гранта Президента РФ для молодых ученых (проект № МК-3228.2018.5, обработка данных).

References

- [1] Agoshkov V.I., Assovskiy M.V., Parmuzin E.I., Zakharova N.B., Zalesny V.B., Shutyaev V.P. Variational assimilation of observation data in the mathematical model of the Black Sea taking into account the tide-generating forces, *Russ. J. Numer. Anal. Math. Modelling*, 2015, V. 30, No. 3, pp. 129–142. DOI: 10.1515/rnam-2015-0013
- [2] Agoshkov, V.I., Parmuzin, E.I., Zakharova, N.B., Shutyaev, V.P. Variational assimilation with covariance matrices of observation data errors for the model of the Baltic Sea dynamics, *Russ. J. Numer. Anal. Math. Modelling*, 2018, v. 33, no. 3, pp. 149-160. DOI: 10.1515/rnam-2018-0013
- [3] Agoshkov V.I., Gusev A.V., Diansky N.A., Oleinikov R.V. An algorithm for the solution of the ocean hydrothermodynamics problem with variational assimilation of the sea level function data, *Russ. J. Numer. Anal. Math. Modelling*, 2007, 22(2), pp. 1-10. DOI: 10.1515/RJNAMM.2007.007
- [4] Agoshkov V.I., Parmuzin E.I., Shutyaev V.P. A numerical algorithm of variational data assimilation for reconstruction of salinity fluxes on the ocean surface, *Russ. J. Numer. Anal. Math. Modelling*, 2008, 23(2), 135-161. DOI: 10.1515/RJNAMM.2008.009
- [5] Захарова Н.Б., Зотов А.Э. Обработка данных дистанционного зондирования о температуре поверхности Балтийского моря // Региональные проблемы дистанционного зондирования Земли: материалы V Междунар. Науч. конф., Красноярск, 11-14 сентября 2018 г. – Красноярск: Сиб. Федер. Ун-т, 2018. С. 318-321.