

Исследование чувствительности алгоритма усвоения малочисленных данных наблюдений в модели динамики океана

М. Н. Кауркин^{1,*}, Р. А. Ибраев^{2,1,3}

¹Институт океанологии им. П. П. Ширшова РАН, Москва, Россия

²Институт вычислительной математики им. Г. И. Марчука РАН, Москва, Россия

³Московский физико-технический институт (национальный исследовательский университет), Долгопрудный, Россия

*E-mail: kaurkinmn@gmail.com

Поступила в редакцию 08.11.2018 г., после доработки – 11.12.2018 г.

Введение. Проведен анализ оригинальной параллельной реализации метода ансамблевой оптимальной интерполяции для усвоения данных наблюдений в модели динамики океана Института вычислительной математики и Института океанологии (модель ИВМИО) с разрешением $0,1^\circ$ для региона Северной Атлантики.

Материалы и методы. На основе известного («истинного») модельного состояния океана были выделены температурные профили (около 70 в сутки, глубина до 1500 м), которые использовались в качестве синтетических данных наблюдений. После возмущения начального условия проведены численные эксперименты для оценки скорости и точности приближения всего модельного решения к «истинному» состоянию океана по мере усвоения температурных профилей.

Анализ результатов. Дан качественный анализ результатов, приведены графики среднеквадратической и средней ошибок модельного решения. Для исследования чувствительности метода к объему используемых данных наблюдений были выполнены эксперименты с уменьшающимся количеством точек наблюдений (1/2, 1/4, 1/8 и 1/16 часть всех температурных профилей), которые показали, что усвоение даже единичных данных способно существенно повысить качество модельного прогноза.

Обсуждение и заключение. Продемонстрировано, что использование метода ансамблевой интерполяции с привлечением относительно малочисленных данных способно существенно улучшить все модельное решение даже для тех регионов, где данных наблюдений было мало или не было вовсе. Только за счет усвоения всего за 3–4 дня среднеквадратическая ошибка для модельного поля ТПО уменьшается на $1,5^\circ\text{C}$, а среднее отклонение становится практически равным нулю по всей расчетной области.

Ключевые слова: моделирование динамики океана, усвоение данных наблюдений, ансамблевая оптимальная интерполяция, вихреразрешающая модель, данные Argo.

Благодарности: работа выполнена в рамках государственного задания ФАНО России (тема № 0149-2018-0020) при частичной поддержке РФФИ (проект № 16-05-01101). При проведении расчетов использовались ресурсы Межведомственного суперкомпьютерного центра РАН.

Для цитирования: Кауркин М. Н., Ибраев Р. А. Исследование чувствительности алгоритма усвоения малочисленных данных наблюдений в модели динамики океана // Морской гидрофизический журнал. Т. 35, № 2. С. 105–113. doi:10.22449/0233-7584-2019-2-105-113

Study of Sensitivity of the Algorithm for Assimilating Small Amount of Data in the Ocean Dynamics Model

M. N. Kaurkin¹, R. A. Ibrayev^{2,1,3}

¹*Shirshov Institute of Oceanology, Russian Academy of Sciences, Moscow, Russia*

²*Marchuk Institute of Numerical Mathematics, Russian Academy of Sciences, Moscow, Russia*

³*Moscow Institute of Physics and Technology (national research university), Dolgoprudny, Russia*

*e-mail: kaurkinmn@gmail.com

Introduction. Analyzed is the original parallel realization of the ensemble optimal interpolation (*EnOI*) method for data assimilation in the ocean dynamics model developed in the Institute of Numerical Mathematics and the Institute of Oceanology (INMIO model) with a resolution 0.1° for the North Atlantic region.

Data and methods. Based on the known (“true”) model state of the ocean, the temperature profiles (about 70 per day, up to 1500 m depth) were chosen and used as synthetic observational data. After the initial condition was perturbed, the numerical experiments were carried out to estimate speed and accuracy of approaching the entire model solution to the “true” state of the ocean as the temperature profiles were assimilated.

Results. Both qualitative analysis of the results and the graphs of the root-mean-square and mean errors of the model solution are given. To study the method sensitivity to the amount of the observational data, the experiments with the decreasing number of data (1/2, 1/4, 1/8 and 1/16 of all the temperature profiles) were performed. They showed that assimilation even of the isolated data could significantly increase quality of the model forecast.

Discussion and conclusion. The experiments prove that application of the ensemble optimal interpolation method, even in case very few data are assimilated in the model, can significantly improve quality both of the model forecast and the entire model solution for those regions where the observational data are very scarce or absent at all. Thus due to assimilation of the data covering only 3–4 days, the root-mean-square error for the sea surface temperature model field decreases by 1.5°C , and the average deviation becomes equal almost to zero over the entire computational domain.

Keywords: ocean dynamics modeling, observational data assimilation, ensemble optimal interpolation, eddy-resolving model, Argo data.

Acknowledgments: The investigation is carried within the framework of the state tasks of FASO, Russia (theme № 0149-2018-0020) at partial support of RFBR (project № 16-05-01101). Resources of the Interagency supercomputer center of RAS were used in calculations.

For citation: Kaurkin, M. N. and Ibrayev, R. A., 2019. Study of sensitivity of the algorithm for assimilating small amount of data in the ocean dynamics model. *Morskoy Gidrofizicheskiy Zhurnal*, [e-journal] 35(2), pp. 105-113. (in Russian).

Введение

В данной работе мы продолжаем исследования по прогнозу состояния океана с применением ансамблевых методов усвоения. Мы не будем здесь приводить обзор работ по ассимиляции, достаточно подробно об это сказано в наших предыдущих публикациях. В [1] была представлена параллельная реализация метода ансамблевой оптимальной интерполяции (*Ensemble Optimal Interpolation, EnOI*) для усвоения данных наблюдений в вихререзающей модели Мирового океана. Эффективность разработанного параллельного метода была исследована на данных температуры и солёности с дрейфтеров *Argo* [2] и данных спутниковой альтиметрии *AVISO* [3]. Результатом этих работ стало подтверждение, что усвоение данных методом *EnOI* в модели даёт значительное уменьшение ошибок в прогнозируемых трехмерных полях температуры и солёности даже в тех регионах, где данные отсутствуют.

ствовали. Из этого следует, что, возможно, из всего многообразия данных наблюдений о состоянии океана можно выбрать их минимально необходимое количество так, чтобы точность прогноза была максимальной.

Как это ни удивительно, но для восстановления и прогноза мезомасштабной динамики океана в настоящее время существует даже избыточное количество данных. Прежде всего это касается спутниковых данных (температура поверхности океана, альтиметрия). На их основе разрабатываются методы измерения солёности и течений поверхности океана, приповерхностного ветра и т. д. Состояние верхнего двухкилометрового слоя океана монитрится буями *Argo*. Большой объем информации поступает с фиксированных в пространстве измерителей состояния океана. Так, например, в работе [4] приводится методология совместной обработки малочисленных гидрологических и спутниковых альтиметрических наблюдений, которая позволяет весьма успешно восстановить трехмерные поля температуры и солёности. Вся эта информация используется в системах прогноза состояния океана разного временного масштаба (см. обзоры в [5–7]). Также хотелось бы отметить работу [8], в которой проводятся численные эксперименты на базе совместной модели *ECMWF-HOPE* и ансамблевого метода усвоения, а затем анализируются результаты возмущения модельных полей и процессы распространения этих возмущений. В [9] обсуждаются методы расчета динамики океана и Черного моря с использованием усвоения гидрологических данных наблюдений на основе применения фильтра Калмана и принципов адаптации гидродинамических модельных полей.

Цель данной работы заключается в валидации параллельного алгоритма усвоения для модели динамики океана ИВМИО. Попытаемся оценить чувствительность точности прогноза к начальному состоянию и некоторым вариациям в количестве усваиваемых данных. Очевидно, что исследования по зависимости точности прогноза от количества, типа, географического распределения данных об океане имеют значение для оптимизации наблюдательных систем. Также обозначим особенности параллельного метода ансамблевой оптимальной интерполяции.

Модель океана и параллельный метод ансамблевой оптимальной интерполяции

Для проведения численных экспериментов будем использовать модель динамики океана ИВМИО, которая была разработана для исследования циркуляции вод Мирового океана и его отдельных акваторий в широком диапазоне пространственных и временных масштабов. Более детальное описание модели и численных экспериментов по ее верификации с разрешением от $0,5^\circ$ до $0,1^\circ$ приведено в работах [10, 11]. Наряду с моделью океана будет использоваться компактная вычислительная платформа *CMF3.0* (*Compact Modeling Framework*) для решения задачи совместного моделирования Земной системы и ее компонентов высокого разрешения на компьютерах с распределенной памятью. Платформа *CMF3.0* используется для межпроцессорных обменов внутри модельных компонентов, интерполяции данных между различными модельными сетками компонентов, работы с файловой системой (ввод/вывод данных) [12].

Метод ансамблевой оптимальной интерполяции *EnOI* подробно описан в работах [13, 14]. В настоящее время он используется для оперативного прогноза состояния океана в ряде систем (см., напр., [15–17]). Подробное описание разработанного и используемого в настоящем исследовании параллельного метода *EnOI* приведено в [1].

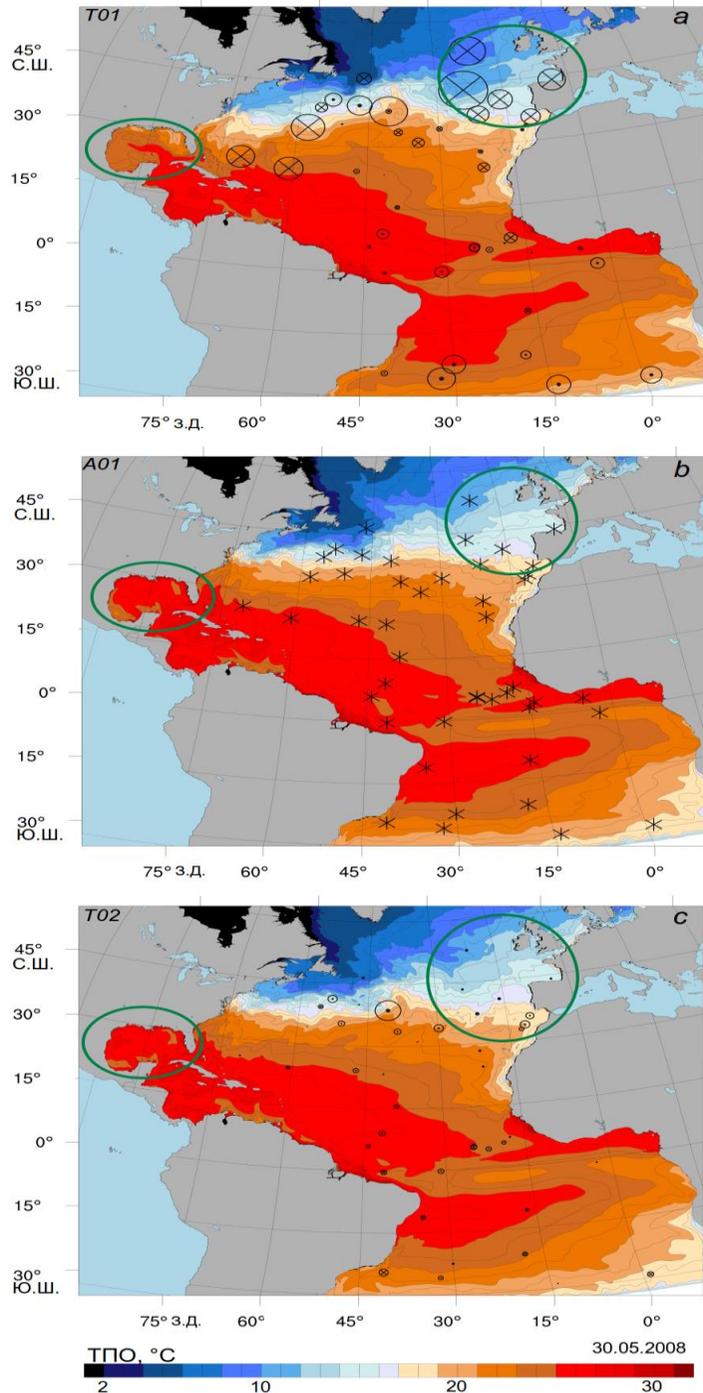
Численные эксперименты

Валидация алгоритма. Для проверки работы системы, состоящей из модели океана и алгоритмов усвоения, проведены численные эксперименты для региона Северной Атлантики. Модель интегрировалась с 01.01.1948 г. по 01.05.2008 г., каждые 10 дней сохранялось полное состояние модели (контрольные точки). Таким образом, в качестве элементов ансамбля для метода *EnOI* при усвоении данных за май 2008 г. использовались по три контрольные точки за каждый май в течение 1985–1998 гг. Суммарно 40 элементов в ансамбле, радиус отсечения не используется, поскольку он не так важен в данном методе [18]. Условия проведения экспериментов такие же, как в работах [2, 3]. Единственная разница заключается в использовании синтетических данных наблюдений, полученных на основе известного модельного состояния океана, которое будем считать «истинным». На основе контрольного эксперимента *A01* были получены малочисленные синтетические данные наблюдений Y_{SOD} (*synthetic observation data*), распределенные по области как данные *Argo* (около 70 профилей в сутки, глубина до 1500 м). Синтетические данные располагались с той же периодичностью и хаотичностью, как и данные измерений *Argo* (не обязательно в узлах модельной сетки). Для проекции в эти точки данных с модельной сетки (т. е. для построения матрицы проекции H) использовалась билинейная интерполяция. Затем были проведены два эксперимента *T01* и *T02* с возмущенным начальным условием (соответствующим модельному решению за другой год), но атмосферное воздействие использовалось за май 2008 г. Однако в эксперименте *T02* будут ежедневно усваиваться данные Y_{SOD} (таблица).

Описание экспериментов Description of the experiments

Эксперимент / Experiment	Дата начального условия / Date of the initial condition	Использование данных наблюдений Y_{SOD} / Use of observational data Y_{SOD}
<i>A01</i>	01.05.2008 г.	Ежесуточное сохранение Y_{SOD} / Daily record of Y_{SOD}
<i>T01</i>	01.05.2000 г.	Не используются / No application
<i>T02</i>	01.05.2000 г.	Ежесуточное усвоение Y_{SOD} методом <i>EnOI</i> / Daily assimilation of Y_{SOD} by the <i>EnOI</i> method

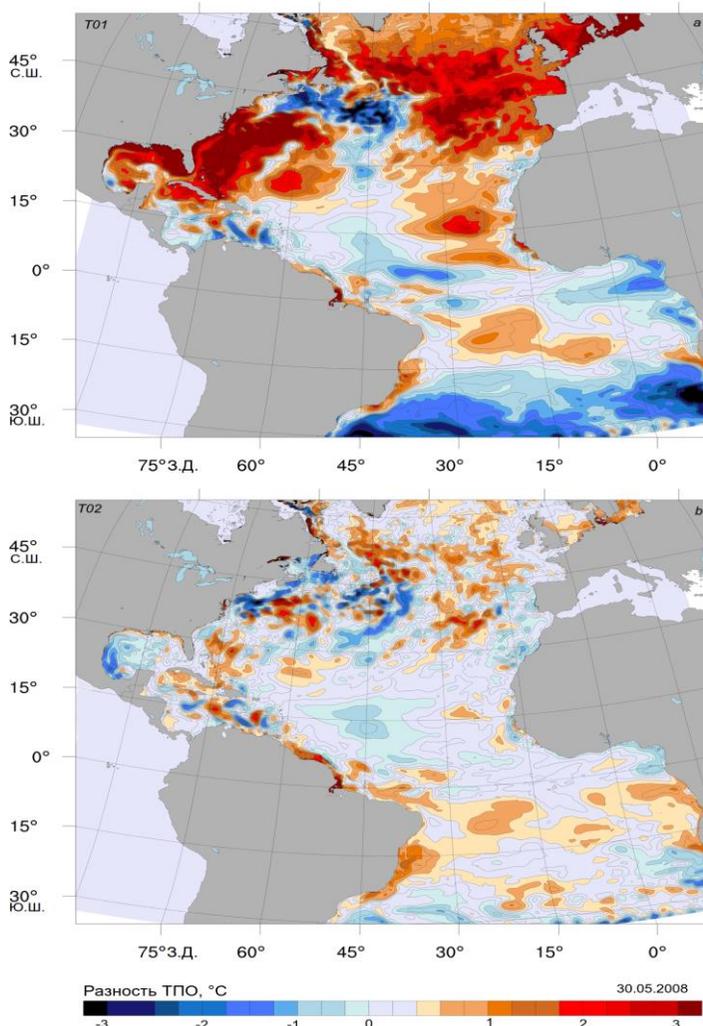
Примечание. Используется атмосферный форсинг *COREII* за период 01.05. – 30.05.2008 г.
Note. Atmospheric forcing *COREII* for 01.05.–30.05.2008 is used



Р и с. 1. Температура поверхности океана в модели Северной Атлантики за 30.05.2008 г. в «истинном» эксперименте A01 (b), а также в экспериментах T01 (a) и T02 (c) с возмущенным начальным условием

Fig. 1. Sea surface temperature (SST) in the North Atlantic Ocean model on 30.05.2008 in the “true” experiment A01 (b) and also in the experiments T01 (a) and T02 (c) with the perturbed initial condition

На рис. 1, где показана модельная температура поверхности океана (ТПО) для различных экспериментов, приведено расположение данных Y_{SOD} , чтобы создать представление об их количестве. Размер кругов на рис. 1, *a*, *c* пропорционален разности между температурой в точке наблюдения и модельной. Крестик в круге означает, что модельная температура ниже температуры профиля, точка в круге – выше. Для удобства сравнения на рис. 1, *b* звездочками отмечены данные Y_{SOD} , которые генерируются в эксперименте *A01*. На фрагментах обведены две области, в которых эффект от усвоения данных Y_{SOD} наиболее заметен. Стоит обратить внимание, что данных наблюдений в Мексиканском заливе не было на протяжении всего эксперимента, однако ТПО в нем была существенно скорректирована за счет усвоения.

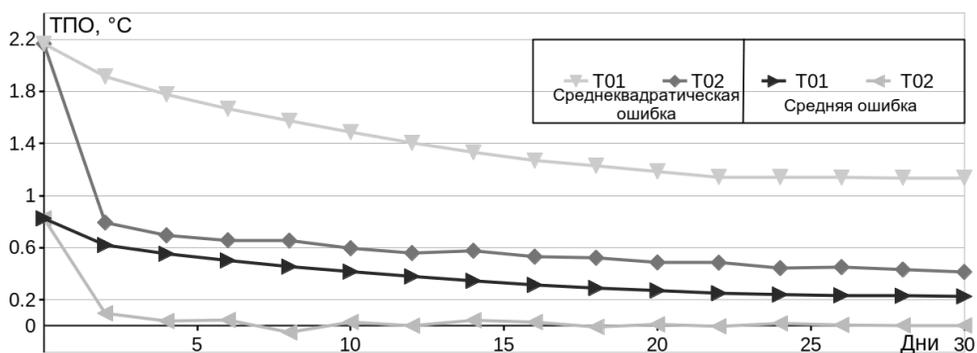


Р и с. 2. Разность между модельными полями ТПО за 30.05.2008 г.: ТПО_{A01} – ТПО_{T01} (*a*); ТПО_{A01} – ТПО_{T02} (*b*)

F i g. 2. Difference between the sea surface temperature (SST) model fields on 30.05.2008: SST_{A01} – SST_{T01} (*a*); SST_{A01} – SST_{T02} (*b*)

На рис. 2 приведена разность между модельными полями ТПО в экспериментах *T01*, *T02* и в «истинном» эксперименте *A01* за 30.05.2008 г. (модельная дата). При сравнении рис. 2, *a*, *b* наглядно видно, как усвоение исправляет решение во всей расчетной области, а не только в точках, где находятся данные Y_{SOD} . Особенно сильно заметен эффект от усвоения данных в районе течения Гольфстрим.

Графики среднеквадратической и средней ошибок ТПО, вычисленных по всем расчетным узлам модельной сетки, приведены на рис. 3. Они показывают, что усвоение Y_{SOD} значительно приближает модельное поле ТПО в эксперименте *T02* к модельному полю в эксперименте *A01* по всей расчетной области. Это подтверждает корректную работу реализованного метода усвоения данных, использование которого с привлечением малочисленных данных наблюдений Y_{SOD} (они расположены с той же плотностью, что и данные *Argo*) способно приблизить все модельное решение к «истинному» состоянию океана (эксперимент *A01*).

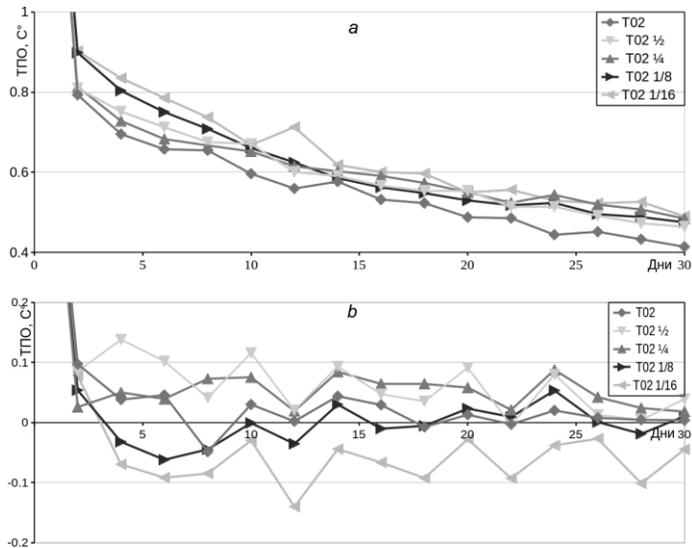


Р и с. 3. Среднеквадратическая и средняя ошибки ТПО в экспериментах *T01* и *T02* по сравнению с «истинным» полем ТПО из эксперимента *A01*

F i g. 3. Root-mean-square (RMS) and average (AVG) errors in the experiments *T01* and *T02* in comparison to the SST “true” field from experiment *A01*

Исследование чувствительности метода к объему используемых данных.

Для этой цели кроме эксперимента *T02* проведены аналогичные эксперименты по усвоению синтетических данных Y_{SOD} , но с уменьшающимся количеством точек наблюдений. Суммарно кроме эксперимента *T02* было проведено еще 4 эксперимента с 1/2, 1/4, 1/8 и 1/16 частью всех температурных профилей. В эксперименте с 1/16 частью данных происходило усвоение лишь 4–5 профилей в сутки. Графики среднеквадратической и средней ошибок модельного поля ТПО для этих экспериментов приведены на рис. 4. Из анализа этого рисунка можно сделать вывод о высокой эффективности используемого метода при малочисленных данных наблюдений, так что даже усвоение единичных данных способно значительно повысить качество моделирования.



Р и с. 4. Среднеквадратическая (a) и средняя (b) ошибки модельного поля ТПО, вычисленные по всем расчетным узлам модельной сетки, в экспериментах на базе T02 с усвоением всего объема, 1/2, 1/4, 1/8 и 1/16 части всех температурных профилей по отношению к «истинному» полю ТПО в эксперименте A01

F i g. 4. Root-mean square (RMS) (a) and average (AVG) errors of the sea surface temperature (SST) model field calculated in all the mesh nodes in the experiments based on T02 with assimilation of all the temperature profiles, and 1/2, 1/4, 1/8 and 1/16 parts of them in comparison to the sea surface temperature (SST) “true” field in experiment A01

Выводы

В настоящей работе продолжено исследование предложенного ранее параллельного метода усвоения данных наблюдений на основе ансамблевой оптимальной интерполяции *EnOI* на примере модели динамики океана высокого разрешения ИВМИО для Северной Атлантики. Проведены численные эксперименты с привлечением малочисленных синтетических данных наблюдений. Особенностью такого подхода перед использованием натуральных данных измерений (спутниковых или дрейферных) является возможность оценить скорость приближения модельного решения к «истинному» состоянию океана, которое неизвестно в натуральных экспериментах. Показано, что использование метода *EnOI* с привлечением относительно малочисленных данных способно существенно улучшить все модельное решение даже в тех регионах, где данные наблюдений отсутствовали. Только за счет усвоения малочисленных данных наблюдений всего за 3–4 дня среднеквадратическая ошибка для модельного поля температуры уменьшается на $1,5^{\circ}\text{C}$, а среднее отклонение становится практически равным нулю по всей расчетной области.

СПИСОК ЛИТЕРАТУРЫ

1. Kaurkin M., Ibrayev R., Koromyslov A. EnOI-Based Data Assimilation Technology for Satellite Observations and ARGO Float Measurements in a High Resolution Global Ocean Model Using the CMF Platform // Supercomputing. RuSCDays 2016 / Eds. V. Voevodin, S. Sobolev. Cham : Springer, 2016. P. 57–66. https://doi.org/10.1007/978-3-319-55669-7_5
2. Кауркин М. Н., Ибраев Р. А., Беляев К. П. Усвоение данных АРГО в модель динамики океана с высоким разрешением по методу ансамблевой оптимальной интерполяции (EnOI) // Океанология. 2016. Т. 56, № 6. С. 852–860. doi:10.7868/S0030157416060058

3. *Кауркин М. Н., Ибраев Р. А., Беляев К. П.* Усвоение данных альтиметрии в модели динамики океана методом ансамблевой интерполяции // *Известия РАН. Физика атмосферы и океана.* 2018. Т. 54, № 1. С. 64–72. doi:10.7868/S0003351518010075
4. *Коротчаев Г. К., Лишаев П. Н., Кныш В. В.* Восстановление трехмерных полей солёности и температуры Черного моря по данным спутниковых альтиметрических измерений // *Исследование Земли из космоса.* 2016. № 1–2. С. 199–212. doi:10.7868/S0205961416010073
5. *Operational oceanography in the 21st century / Eds. A. Schiller, G. B. Brassington.* Dordrecht, The Netherlands : Springer, 2011. 745 p. doi:10.1007/978-94-007-0332-2-18
6. *Assessing the impact of observations on ocean forecasts and reanalyses: Part 1, Global studies / P. R. Oke [et al.] // Journal of Operational Oceanography.* 2015. Vol. 8, suppl. 1. P. 49–62. doi:10.1080/1755876X.2015.1022067
7. *Ocean Data Assimilation Systems for GODAE / J. Cummings [et al.] // Oceanography.* 2009. Vol. 22, no. 3. P. 96–109. doi:10.5670/oceanog.2009.69
8. *Belyaev K. P., Tuchkova N. P., Cubash U.* Response of a coupled ocean-ice-atmosphere model to data assimilation in the tropical zone of the Pacific Ocean // *Oceanology.* 2010. Vol. 50, iss. 3. P. 306–316. doi:10.1134/S0001437010030021
9. *Assimilation of hydrological observation data for calculating currents in seas and oceans / V. V. Knysh [et al.] // Izvestiya, Atmospheric and Oceanic Physics.* 2012. Vol. 48, iss. 1. P. 57–73. doi:10.1134/S0001433812010057
10. *Ибраев Р. А., Хабеев Р. Н., Ушаков К. В.* Вихреразрешающая 1/10° модель Мирового океана // *Известия РАН. Физика атмосферы и океана.* 2012. Т. 48, № 1. С. 45–55. doi:10.1134/S0001433812010045
11. *Ушаков К. В., Ибраев Р. А., Калмыков В. В.* Воспроизведение климата Мирового океана с помощью массивно-параллельной численной модели // *Известия РАН. Физика атмосферы и океана.* 2015. Т. 51, № 4. С. 416–436. doi:10.1134/S0001433815040131
12. *Compact Modeling Framework v3.0 for high-resolution global ocean-ice-atmosphere models / V. V. Kalmykov [et al.] // Geoscientific Model Development.* 2018. Vol. 11, no. 10. P. 3983–3997. <https://doi.org/10.5194/gmd-11-3983-2018>
13. *Evensen G.* The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation // *Ocean Dynamics.* 2003. Vol. 53, no. 4. P. 343–367.
14. *Evensen G.* Data Assimilation: The Ensemble Kalman Filter. Berlin : Springer, 2009. 307 p. <http://dx.doi.org/10.1007/978-3-642-03711-5>
15. *The REMO Ocean Data Assimilation System into HYCOM (RODAS_H): General Description and Preliminary Results / C. A. S. Tanajura [et al.] // Atmospheric and Oceanic Science Letters.* 2014. Vol. 7, no. 5. P. 464–470. doi:10.3878/j.issn.1674-2834.14.0011
16. *Ocean data assimilation: a case for ensemble optimal interpolation / P. R. Oke [et al.] // Australian Meteorological and Oceanographic Journal.* 2010. Vol. 59. P. 67–76. <https://doi.org/10.22499/2.5901.008>
17. *Cummings J. A.* Operational multivariate ocean data assimilation // *Quarterly Journal of the Royal Meteorological Society.* 2005. Vol. 131, iss. 613. P. 3583–3604. doi:10.1256/qj.05.105
18. *Oke P. R., Sakov P., Corney S. P.* Impacts of localisation in the EnKF and EnOI: experiments with a small model // *Ocean Dynamics.* 2007. Vol. 57, iss. 1. P. 32–45. doi:10.1007/s10236-006-0088-8

Об авторах:

Кауркин Максим Николаевич, научный сотрудник, группа моделирования изменчивости климата океанов и морей (физическое направление), Институт океанологии им. П. П. Ширшова РАН (117997, Россия, г. Москва, Нахимовский проспект, д. 36), кандидат физико-математических наук, **ResearcherID: S-1416-2016**, kaurkinmn@gmail.com

Ибраев Рашид Ахметзиевич, главный научный сотрудник, группа исследования крупно- и мезомасштабной динамики вод Мирового океана и окраинных морей России на основе моделирования и анализа данных наблюдений, Институт вычислительной математики им. Г. И. Марчука РАН (119333, г. Москва, ул. Губкина, д. 8), член-корреспондент РАН, доктор физико-математических наук, **ResearcherID: S-6750-2016**, ibraev@math.ru