

Институт Вычислительной Математики Российской Академии Наук

На правах рукописи

Кауркин Максим Николаевич

**Параллельный алгоритм ансамблевой оптимальной интерполяции
усвоения данных наблюдений в модели динамики океана высокого
пространственного разрешения**

05.13.18 «Математическое моделирование, численные методы и комплексы программ»

Диссертация на соискание ученой степени
кандидата физико-математических наук

Научный руководитель:
д.ф.-м.н., чл.-корр. РАН Р.А. Ибраев

Москва — 2017

Оглавление

Оглавление.....	2
Структура диссертации.....	4
Введение.....	6
Задача моделирования динамики океана.....	6
Развитие методов усвоения данных наблюдений.....	9
Основная цель диссертационной работы.....	18
Научная новизна.....	18
Научная и практическая значимость.....	19
Личный вклад автора.....	19
Достоверность результатов.....	21
Публикации по теме диссертации.....	21
Апробация работы.....	22
Благодарности.....	23
Глава 1. Трехмерная математическая модель Мирового океана и компактная вычислительная платформа для моделирования.....	24
1.1 Введение.....	24
1.2 Модель динамики океана.....	25
1.3 Термодинамическая модель льда.....	33
1.4 Компактная вычислительная платформа для моделирования.....	34
1.5 Выводы.....	39
Глава 2. Характеристики данных наблюдений об океане и обзор существующих систем прогноза.....	41
2.1 Введение.....	41
2.2 Данные для усвоения.....	50
2.3 Выводы.....	55
Глава 3. Параллельный алгоритм метода оптимальной интерполяции.....	57

3.1 Введение.....	57
3.2 Методы усвоения данных.....	58
3.3 Метод оптимальной интерполяции (MVOI).....	59
3.4 Параллельный алгоритм для метода MVOI.....	60
3.5 Метод ансамблевой оптимальной интерполяции (EnOI).....	62
3.6 Математическая формулировка метода EnOI.....	63
3.7 Параллельный алгоритм метода EnOI.....	67
3.8 Программная реализация.....	70
3.9 Выводы.....	73
Глава 4. Численные эксперименты по усвоению данных в модели Северной Атлантики.....	74
4.1 Введение.....	74
4.2 Постановка численных экспериментов.....	75
4.3 Проверка метода на синтетических данных.....	78
4.4 Качественный анализ модельного решения.....	82
4.5 Анализ результатов усвоения данных с дрейфтеров ARGO.....	84
4.6 Анализ результатов усвоения данных аномалии уровня AVISO.....	86
4.7 Сравнение с результатами, опубликованными в других работах.....	93
4.8 Выбор параметров метода EnOI.....	94
4.9 Сравнение с независимыми спутниковыми данными ТПО.....	97
4.10 Параллельная эффективность.....	100
4.11 Выводы.....	101
Заключение.....	104
Список сокращений.....	106
Приложение.....	108
Литература.....	111

Структура диссертации

Диссертация состоит из введения, четырех глав, заключения, приложения и списка литературы из 135 наименований. Она содержит 25 рисунков. Каждая глава разбита на разделы, включая введение к главе и выводы из нее.

Во **введении** раскрывается актуальность научной темы, приводится обзор литературы по теме диссертации, формулируются цели и задачи работы, представлена научная новизна и практическое значение работы, сообщается о публикациях и докладах по теме диссертации.

В **первой главе** диссертации приводится описание используемой математической модели динамики океана ИВМИО (основные уравнения термогидродинамических процессов, краевые условия, сетки, алгоритм решения), модели термодинамики льда CICE и вычислительной платформы CMF 3.0, в рамках которой реализован параллельный алгоритм усвоения данных.

Во **второй главе** проводится анализ существующих методов и систем усвоения данных для моделей океана высокого разрешения, а также обзор источников данных наблюдений и их характеристик по следующим параметрам: объем, регулярность, пространственное покрытие; формулируются физические требования к методу усвоения.

В **третьей главе** приводится описание разработанных вычислительных параллельных алгоритмов усвоения данных на основе многомерной оптимальной интерполяции (MVOI) и ансамблевой оптимальной интерполяции (EnOI), и представлены особенности их программной реализации для массивно-параллельных компьютеров с распределенной памятью.

В **четвертой главе** обсуждаются результаты работы разработанного параллельного алгоритма усвоения данных спутниковой альтиметрии и данных о температуре и солености с дрейфтеров ARGO в модели динамики океана ИВМИО для Северной Атлантики с разрешением 0.1° . Приводится качественная и количественная оценка эффективности работы системы усвоения. Оценивается параллельная эффективность предложенного подхода.

Выполняется сравнение с независимыми данными наблюдений и результатами других научных коллективов. Проводятся численные эксперименты для проверки алгоритма на синтетических данных наблюдений и для выбора оптимальных параметров метода EnOI. Анализируется, как усвоение аномалии уровня океана влияет на те параметры, которые не усваиваются непосредственно - температуру и соленость.

В **заключении** формулируются основные результаты диссертационной работы.

Введение

Задача моделирования динамики океана

Модели динамики океана высокого пространственного разрешения с усвоением непрерывно поступающих данных наблюдений позволяют осуществлять оперативный прогноз трёхмерного состояния морской среды, аналогичный метеорологическим прогнозам погоды, а также являются важной составляющей в исследованиях изменения климата и системах мониторинга состояния окружающей среды.

Высокое разрешение модели океана играет важную роль в достоверном моделировании динамики океана. В частности, исследования динамики Атлантического океана показали, что моделирование с шагом сетки не более 0.1° является критичным для воспроизведения пространственно-временных характеристик течения Гольфстрим [125] и течений в проливах в Северной Атлантике между Европой и Гренландией [60]. Высокое разрешение позволяет явно моделировать вихревую структуру циркуляции, не проявляющуюся на более грубом разрешении. Это явление связано с отношением шага модельной сетки к радиусу деформации Россби. Вопросы необходимости высокого разрешения вычислительных моделей океана посвящены работы [15], [52], [91], [92].

В последние 10 лет с нарастающими темпами увеличивается количество спутниковых данных наблюдений за параметрами состояния океана (температура поверхности океана, уровень океана, концентрация льда и другие геофизические величины), а также появились целые системы буев для контактных измерений (ARGO, PIRATA, TAO, TRITON). Стоит отметить, что количество контактных измерений такого рода (профили параметров океана и т. д.) всего полвека назад составляло $\sim 10^4$ измерений в

год и получалось, в основном, в ходе научных экспедиций, сегодня большинство этих систем работает в оперативном режиме.

Объем цифровых данных об окружающей среде быстро растет. Например, объем архива данных, хранящихся в Национальных Центрах Сбора Информации об Окружающей среде (*eng. National Centers for Environmental Information*, NCEI) организации NOAA (USA), увеличился с 2000 года более чем в 24 раза - почти до 25 петабайт (**рисунок 1a**), в значительной степени из-за увеличения количества спутниковых данных наблюдений и данных модельных расчетов. Ежегодный объем данных, предоставляемых пользователям, также увеличился более чем в 11 раз с 2008 года - до 6 петабайт в год (**рисунок 1b**), и ожидается, что он будет расти еще быстрее [83].

Задача эффективного использования растущих объемов данных наблюдений для понимания океанических процессов является актуальной (см. детальный анализ существующих систем в Главе 2). Например, уже сейчас доступны спутниковые данные наблюдений за поверхностной температурой и уровнем с разрешением менее 1 км (например, *NASA Aqua* [101]). Это соответствует поступлению ~1 ГБ информации за одни сутки. Важно эту информацию правильно и быстро использовать для повышения качества прогноза. Также актуальной и востребованной становится задача прогноза состояния океана.

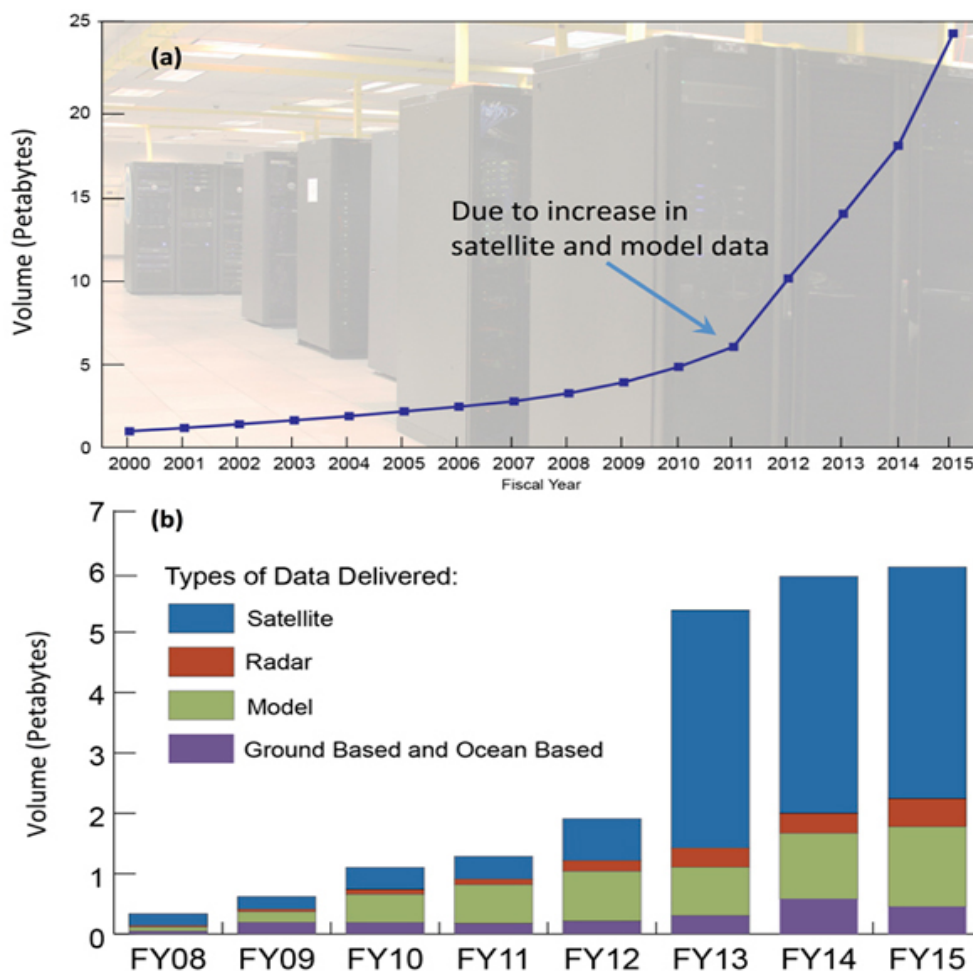


Рисунок 1. (а) Объем данных об окружающей среде, хранящихся на серверах NOAA; **(б)** Запрашиваемые пользователями NOAA объемы данных различных типов (спутниковые наблюдения, данные радиолокации, модельные расчеты, данные измерения наземного и морского базирования) за 2008 - 2015 годы . Рисунок взят из [83]

Для систем прогноза состояния океана, работающих в оперативном режиме, критичным становится вопрос времени решения уравнений модели и выполнения алгоритмов усвоения для построения среднесрочных и краткосрочных прогнозов. Интервал выдачи некоторых современных спутниковых данных наблюдений составляет $O(1)$ часов. Высокое пространственное разрешение моделей океана позволяет моделировать

поведение вихревых структур, а усвоение спутниковых данных наблюдений дает возможность своевременно их обнаружить, что, как следствие, позволяет предсказывать такие природные явления, как штормы, интенсивные вихри и течения.

Решение задачи прогноза состояния океана невозможно без использования параллельных вычислений на компьютерах с распределенной памятью в силу огромного объема получаемой и обрабатываемой информации. Количество ресурсов, необходимых для выполнения расчётов моделями высокого пространственного разрешения, исчисляется сегодня $10^2 - 10^3$ вычислительных ядер для краткосрочных прогнозов и $10^4 - 10^5$ – для средне- и долгосрочных. Поэтому при разработке и внедрении методов усвоения данных наблюдений необходимо решить задачу разработки масштабируемых алгоритмов и программных кодов.

Развитие методов усвоения данных наблюдений

Модель динамики океана описывает временную эволюцию полей физических величин (температуры, солёности, двух компонентов скорости и т.д.). В океанологии под усвоением данных в модели океана понимается процесс корректировки модельного решения, описывающего состояние реального океана, данными наблюдений, чтобы уменьшить расхождения между модельным полем и этими данными, т.е. приблизить решение к реальному состоянию океана. По определению, модель океана является несовершенным представлением реальной системы, использующим упрощенные математические параметризации для сложных процессов и содержащим ошибки, связанные с разрешением модельной сетки и используемых конечно-разностных методов. Даже в верно

сформулированной и реализованной модели океана возникает расхождение с данными наблюдений из-за неопределенностей в начальных условиях и атмосферном форсинге (воздействии). Расхождение модели с реальным состоянием океана являются фундаментальным и неизбежным фактом. Разумеется, наблюдения также не обеспечивают полную картину состояния океана из-за ошибок измерений и малочисленности этих данных.

Усвоение данных можно рассматривать как динамический подход к экстраполяции редких океанических данных для использования их в модели. Это можно сделать различными методами, например линейной интерполяцией, который является простым, но не эффективным способом использования данных наблюдений, поскольку она не учитывает пространственно-временные корреляции в полях модельных величин, которые возникают в результате проявления геофизически значимых явлений, таких как мезомасштабные вихри. Использование модели динамики океана, которая воспроизводит такие корреляции, должно повысить эффективность такой экстраполяции и позволит получить более точную оценку состояния океана в местах, где нет данных.

Основная идея усвоения данных похожа на задачу построения теоретической кривой (модельного решения), которая должна соответствовать предполагаемой физической закономерности (модели), на основе малочисленных экспериментальных точек (данных наблюдений), которая регулярно возникает в экспериментальной физике. При решении этой задачи нужно минимизировать несоответствие между модельными значениями и данными. Эта задача может быть формализована по-разному: нахождение минимума заданного функционала-ошибки (вариационный подход) или поиск оптимальной оценки, называемой также оптимальным фильтром в смысле минимума дисперсии неизвестного состояния океана (которое может рассматриваться как случайный вектор)

по данным наблюдений и траектории модели (динамико-стохастический подход).

При решении данной задачи различными методами возникают схожие проблемы из-за значительного размера вектора состояния модели и малочисленности и разрозненности данных наблюдений, на основании которых этот вектор нужно скорректировать. Например, если взять модель океана, которая имеет $10^3 \times 10^3$ горизонтальных точек сетки с 50 вертикальными уровнями, то ее вектор состояния, состоящий только из пяти модельных полей (две горизонтальные скорости, температура, соленость и функция уровня океана) будет иметь размерность более 2×10^8 .

Изначально проблема корректировки модельного прогноза данными измерений возникла в задаче численного прогноза погоды [7]. Основное различие в решении задачи усвоения данных в моделях атмосферы и океана заключается в значительной разнице в количестве доступных данных наблюдений, точнее во временной и пространственной регулярности измерений этих данных. Сеть метеорологических станций наблюдений существует по всему миру более сотни лет, в то время как появление данных наблюдений за океаном до недавнего времени не носило регулярный характер. А чем меньше информации о состоянии Земной системы доступно, тем сложнее спрогнозировать реальное состояние этой системы, и тем более сложные методы нужно использовать для решения данной задачи.

Существенный прогресс в области прогноза состояния океана, и, прежде всего, в методах усвоения данных наблюдений, достигнутый в последние десятилетия, в значительной степени связан со стартовавшим в 1997 проектом GODAE (*eng. Global Ocean Data Assimilation Experiment*) [62] и последовавшим за ним в 2008 году проектом GODAE OceanView

[72]. Эти проекты стали площадкой для обсуждения проблем мониторинга и прогноза состояния океана.

В настоящее время существует несколько алгоритмов усвоения данных, которые применяются в задачах прогноза погоды и в оперативной океанологии, см. обзор в [121]. Используемые подходы можно разделить на вариационные (3DVar, 4DVar) [4], [5], [13], [40] и динамико-стохастические [47] – преимущественно ансамблевые фильтры Калмана (*eng. Ensemble Kalman Filter, EnKF*) и ансамблевая оптимальная интерполяция (*eng. Ensemble Optimal Interpolation, EnOI*) [66]. Однако можно отметить и ряд работ, в которых используются методы усвоения, несколько отличные от тех подходов, которые приведены выше. Так, например, в [131] используется метод частиц, который является частью реализации так называемого Лагранжевого подхода. В работе [130] используется Байесовский метод максимизации условной вероятности. Есть также ряд гибридных методов, в которых вариационный подход совмещен тем или иным способом с теорией фильтрации, например [54]. Тем не менее, большинство работ по усвоению, публикуемые в настоящий момент, так или иначе относятся к вышеупомянутым основным классам.

В вариационном подходе решается задача поиска минимума функционала ошибки, состоящего из суммы двух функционалов: ошибки прогноза – J_b (1) и ошибки данных наблюдений – J_o . Формула (2) задает этот функционал в методе 3DVar, а (3) – для 4DVar.

$$J_b(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) \quad (1)$$

$$J(\mathbf{x}) = J_b(\mathbf{x}) + \frac{1}{2}(\mathbf{y}_o - \mathbf{H}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}_o - \mathbf{H}(\mathbf{x})) \quad (2)$$

$$J(\mathbf{x}) = J_b(\mathbf{x}) + \frac{1}{2} \sum_{i=0}^N (\mathbf{y}_o^i - \mathbf{H}^i(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}_o^i - \mathbf{H}^i(\mathbf{x})) \quad (3)$$

В формулах (1) – (3):

x – искомый вектор состояния модели;

y_0 – вектор данных наблюдений. Индекс i здесь и ниже показывает то, что данные доступны для разных моментов времени;

x_b – вектор-состояние модельного прогноза;

H – оператор проекции из пространства модельного решения в пространство данных наблюдений;

B и R – ковариационные матрицы ошибки модельного прогноза и данных наблюдений.

Таким образом, в случае 4DVar задача минимизации решается внутри некоторого отрезка времени – окна усвоения. Схематически работа вариационных методов показана на **рисунке 2**. Подобный класс задач в метеорологии был впервые описан в работе [120]. Задачи вариационного усвоения данных с использованием сопряженных уравнений в моделях динамики атмосферы рассматривались в работах Г.И. Марчука и его учеников, одними из первых работ по этой теме были [41], [93]. В последние годы учеными ИВМ РАН достигнут существенный прогресс в вопросах математического обоснования, развития и экспериментальной проверки новых подходов к решению данной задачи [4], [5], [13]. Разработанный метод сопряженных уравнений показал свою универсальность и успешно применялся для решения задач оптимального управления и усвоения в ядерной энергетике, физике атмосферы и океана, метеорологии, прибрежной океанографии, экологии и др. [37], [38].

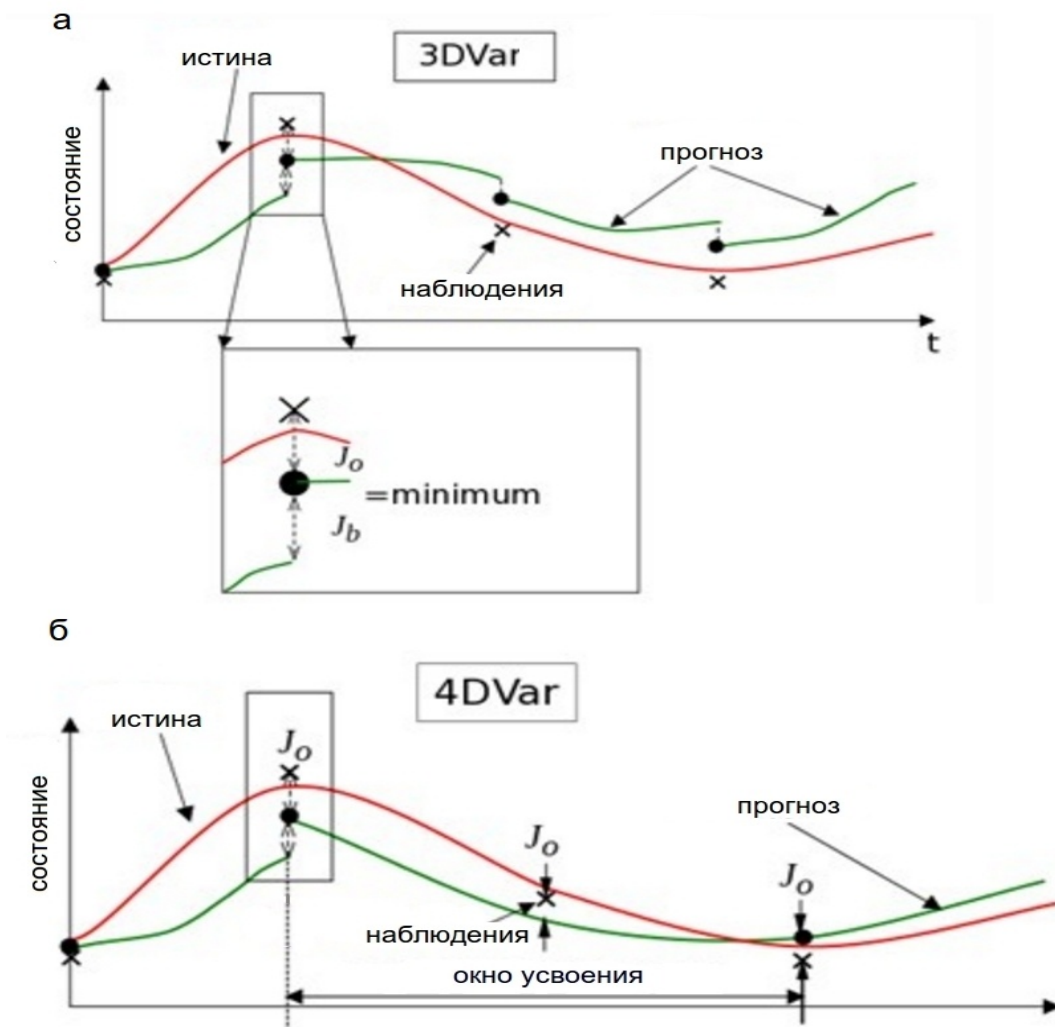


Рисунок 2. Схема работы вариационных алгоритмов усвоения данных наблюдений (а) 3DVar (б) 4DVar.

Системы усвоения, основанные на динамико-стохастическом подходе (ансамблевый фильтр Калмана, ансамблевая оптимальная интерполяция), используются в ведущих прогностических центрах состояния океана (ТОРАЗ, Bluelink, REMO, GODAS) [121].

Альтернативные методы, основанные на применении обобщенной Калмановской теории фильтрации [116], развиваются в работах ученых МГИ РАН по моделированию циркуляции Черного моря [33]. В работе [34]

выполняется ассимиляция¹ температуры и солёности на основе алгоритма адаптивной статистики прогноза ошибок для восстановления климатической сезонной изменчивости течений Каспийского моря. В работе [32] ассимилируются поля температуры и солёности, подготовленные на каждые модельные сутки с учетом известного из наблюдаемой альтиметрии со спутников уровня моря и измерений дрейфтеров ARGO. Ассимиляция восстановленных трехмерных полей в уравнениях переноса диффузии тепла и соли осуществлялась методом релаксации, посредством включения в правую часть этих уравнений ассимиляционных источников.

В последнее десятилетие появилось большое число работ, посвященных усвоению данных альтиметрии [87] в различных моделях динамики океана. Например, стоит отметить работу [127], в которой данные альтиметрии, полученные со спутников AVISO, усваивались в модели NUCOM [55] различными методами, в том числе и методом объективной ансамблевой интерполяции (EnOI). Отличительной чертой методов, основанных на EnKF и EnOI, является то, что они не требуют построения сопряженного оператора модели и решения обратной задачи, что для модели с большим числом параметров весьма затруднительно, и, вообще, используют модель как «черный ящик» [81].

До сих пор нет ясности, какой из подходов, 4DVar, EnKF или какие-то гибридные схемы предпочтительнее в плане точности [56]. В каждом конкретном случае ищется оптимальное решение задачи усвоения (минимизация дисперсии, максимум правдоподобия при Байесовском подходе, минимум вариации и т.д.). Результаты применения того или иного

¹ Понятия «усвоение данных» и «ассимиляция данных» является аналогичными.

метода зависят как от модели, так и от региона наблюдений и конкретных данных.

Вычислительные проблемы при усвоении данных

По мере увеличения мощностей вычислительных ресурсов, требуется постоянная разработка и использование численных алгоритмов, эффективно использующих эти ресурсы. Например, для решения задач геофизики фактически стало стандартом использование компьютеров параллельной архитектуры с распределенной памятью, поэтому большинство наиболее распространенных моделей динамики океана (например MOM [73] или POM [57]) получили программную реализацию с использованием библиотеки MPI [129]. Как следствие, в силу увеличения разрешения пространственных сеток в моделях, а также из-за роста объема данных наблюдений, аналогичные проблемы необходимо решать и для методов усвоения данных.

Для решения задачи усвоения данных в последнее десятилетие стали появляться программные продукты, которые распространяются как программы с открытыми исходными кодами (eng. *open-source software*) и используются для учебных и исследовательских целей в задачах моделирования океана и атмосферы [64], [102], [103]. Однако большая часть программных реализаций наиболее развитых методов усвоения данных, на создание которых требуются значительные усилия, особенно в применении для компьютеров параллельной архитектуры, остаются закрытыми. Это обусловлено тем, что подобные методы и системы представляют коммерческий интерес, например для нефтедобывающих компаний (система REMO [128]) или для военно-морского флота (система NCODA [126]). Таким образом, разработка разных систем усвоения

данных ведется относительно независимо, в том смысле, что обмен знаниями происходит лишь через научные статьи, а не передачей готовых программных продуктов. Что позволяет более объективно оценивать преимущества тех или иных методов усвоения.

Среди новых отечественных работ по усвоению данных автор хотел бы отметить работу [123], в которой уделено пристальное внимание именно параллельной реализации ансамблевых фильтров Калмана для полуЛагранжевой модели атмосферы ПЛАВ [49].

Важность разработки параллельных алгоритмов усвоения данных наблюдений в рамках вариационного подхода в формулировке со слабыми ограничениями и применением схем расщепления для изучения атмосферы, океана и окружающей среды отмечается в работах [39], [40].

Основная цель диссертационной работы

Основными целями диссертационной работы являются:

1. Разработка параллельного алгоритма усвоения данных наблюдений на основе динамико-стохастического подхода в модели динамики океана высокого пространственного разрешения.
2. Тестирование алгоритма усвоения данных спутниковых и дрейферных наблюдений с применением модели динамики океана ИВМИО для региона Северной Атлантики.

Научная новизна

Предложенная система усвоения данных наблюдений, основанная на ансамблевом подходе, имеет эффективную параллельную реализацию для компьютеров с распределенной памятью и практически не зависит от математической формулировки модели динамики океана. Система усвоения реализована в качестве программного сервиса усвоения данных вычислительной платформы CMF3.0 для использования в моделях высокого пространственного разрешения на массивно-параллельных компьютерах, пришедшего на смену использовавшемуся ранее интегрированному в модель океана методу многомерной оптимальной интерполяции.

Представленная в работе система усвоения данных океанических наблюдений, состоящая из математической модели динамики океана ИВМИО, вычислительной платформы CMF 3.0 и работающего на ее базе программного сервиса усвоения данных наблюдений на основе метода ансамблевой оптимальной интерполяции, является первой в России

системой, способной работать на пространственных Глобальных сетках с пространственным разрешением 0.1° .

Научная и практическая значимость

Разработанная система усвоения данных позволяет эффективно усваивать различные данные спутниковых и дрейферных наблюдений и корректировать данные модельных расчетов, значительно повышая качество моделирования. Эффективность разработанного и реализованного параллельного метода усвоения данных наблюдений подтверждена тестами на современных суперкомпьютерах («Ломоносов», «МВС-10П», «РСК Торнадо» и др.). С применением программного сервиса усвоения данных в модели Северной Атлантики ИВМИО с пространственным разрешением 0.1° были усвоены данные спутниковой альтиметрии AVISO и данные о температуре и солености с дрейферов ARGO. Показано, что ошибки прогноза после усвоения по сравнению с контрольным расчётом уменьшаются почти в два раза и в целом эффективность всей системы находится на мировом уровне, в плане качества моделирования и параллельной эффективности. Предложенная система усвоения данных наблюдений на базе модели океана ИВМИО ориентирована на использование в среднесрочном и долгосрочном прогнозе.

Личный вклад автора

Представленная диссертация является самостоятельным законченным трудом автора. Научные результаты диссертации, выносимые на защиту, получены лично автором, являются новыми. Результаты других авторов, упомянутые в тексте диссертации, а также полученные в соавторстве,

отмечены соответствующими ссылками и принадлежат соавторам в равной степени.

Автором были получены следующие результаты:

1. Разработан параллельный алгоритм многомерной оптимальной интерполяции (MVOI) для усвоения данных наблюдений ARGO в модели динамики океана ИВМИО.

2. Создана программная архитектура и разработан параллельный алгоритм ансамблевой оптимальной интерполяции (EnOI), масштабируемый для нерегулярных по пространству спутниковых и дрейферных данных наблюдений.

3. Написана программная реализация параллельного алгоритма усвоения данных EnOI в виде программного сервиса вычислительной платформы совместного моделирования CMF 3.0. Вычислительная эффективность и параллельная масштабируемость разработанной системы подтверждена тестами на различных современных суперкомпьютерах массивно-параллельной архитектуры.

4. На основе серии численных экспериментов для модели динамики Северной Атлантики высокого разрешения ИВМИО с усвоением данных наблюдений о температуре и солености с дрейферов ARGO и данных спутниковой альтиметрии AVISO получены качественные и количественные оценки точности решения на основе независимых данных наблюдений, которые соответствуют результатам ведущих научных коллективов.

Достоверность результатов

Достоверность результатов диссертационной работы обоснована использованием строгих математических выводов со ссылками на статьи других авторов, валидацией результатов прогноза состояния океана по данным наблюдений, вычислительными экспериментами по масштабируемости разработанной программы и сравнением характеристик решения с результатами полученными в других работах.

Публикации по теме диссертации

Было опубликовано **10** работ, из них **6** публикаций, в которых излагаются основные научные результаты диссертации на соискание ученой степени кандидата наук:

1. 3 статьи в журналах из перечня ВАК [10], [24], [25].
2. 1 статья в зарубежном сборнике², выпуски которого индексируются Web of Science [84].
3. 2 свидетельства о государственной регистрации программ для ЭВМ³ [26], [27].

² Высшая аттестационная комиссия Минобрнауки России информирует, что начиная с 21 апреля 2008 года к периодическим изданиям, включенным в перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертации на соискание ученой степени доктора и кандидата наук, относятся зарубежные издания, включенные в одну из трех систем цитирования *Web of Science: Science Citation Index Expanded*...

³ В соответствии с пунктом 13 Положения о присуждении ученых степеней, утвержденного постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842, к **публикациям, в которых излагаются основные научные результаты диссертации** на соискание ученой степени, приравниваются патенты на изобретения, патенты (свидетельства) на полезную модель, патенты на промышленный образец, патенты на селекционные достижения, **свидетельства на программу для электронных вычислительных машин**, ...

4. 4 работы в различных сборниках, индексируемых РИНЦ (Российский индекс научного цитирования) [28], [29], [30], [31].

Апробация работы

Материалы, вошедшие в диссертационную работу, представлялись на научных семинарах, международных и всероссийских конференциях:

- Международная суперкомпьютерная конференция «Научный сервис в сети Интернет: многообразие суперкомпьютерных миров». г. Новороссийск, пос. Абрау-Дюрсо. 2014 г.

- Школа-семинар «Современные методы дистанционных исследований и прогноза параметров среды в Арктике». Таруса Калужской обл., ИКИ РАН. 2015 г.

- Всероссийская конференция «Комплексные исследования морей России: оперативная океанография и экспедиционные исследования». г. Севастополь, ФГБУН МГИ. 2016 г.

- Всероссийская Конференция «Научный сервис в сети интернет». г. Новороссийск, пос. Абрау-Дюрсо. 2016 г.

- Доклад на Ученом совете физического направления ИО РАН (03.06.2016)

- Конференция «Мировой океан: модели, данные и оперативная океанология». г. Севастополь: ФГБУН МГИ, октябрь 2016 г.

- Конференция «Суперкомпьютерные дни в России». г. Москва, МГУ. 2016 г.

- Доклад на семинаре Лаборатории гидрометеорологии Арктики Гидрометцентра России (18.10.2016)

Материалы диссертации содержатся в годовых отчетах о работе Суперкомпьютерного комплекса МГУ имени М.В. Ломоносова и

Межведомственного суперкомпьютерного центра Российской академии наук.

Благодарности

Автор выражает благодарность:

- своему научному руководителю чл.-корр. РАН Р.А. Ибраеву за внимательное руководство в процессе обучения в магистратуре, аспирантуре и в ходе выполнения научных исследований;

- д.ф.-м.н. К.П. Беляеву за консультации по методам усвоения данных;

- безвременно ушедшему ак. А.С. Саркисяну за переданные научные знания и опыт;

- к.ф.-м.н. К.В. Ушакову и к.ф.-м.н. В.В. Калмыкову за активное участие в доработке модели динамики океана ИВМИО и компактной вычислительной платформы CMF;

- д.ф.-м.н. В.И. Агошкову, д.ф.-м.н. В.Б. Залесному и д.ф.-м.н. М.А. Толстых за сделанные полезные замечания при подготовке диссертации.

- всем сотрудникам ИВМ РАН, многие из которых читали автору различные курсы.

Работа выполнена при поддержке Программы Президиума РАН I.33П «Фундаментальные проблемы математического моделирования» и гранта Российского научного фонда (проект №14-37-00053). Численные эксперименты с моделью океана высокого разрешения были выполнены на суперкомпьютерах «Ломоносов» и «МВС-10П».

Глава 1. Трехмерная математическая модель Мирового океана и компактная вычислительная платформа для моделирования

1.1 Введение

В работе для создания системы усвоения данных наблюдений будет использоваться математическая модель динамики океана ИВМИО [15], [52] и компактная вычислительная платформа для моделирования CMF 3.0 (*eng. Compact Modeling Framework*), которая является программным комплексом для решения задач совместного моделирования Земной системы и ее компонентов высокого разрешения на массивно-параллельных компьютерах [16], [17].

По сравнению с предыдущей версией — моделью гидродинамики внутренних морей (МГВМ) [14], модель динамики океана была переработана так, чтобы с ее помощью можно было решать задачи с высоким пространственным разрешением для области всего Мирового океана. Многие программно-вычислительные аспекты работы модели, такие, как работа с файловой системой или обмен данными между вычислительными ядрами, реализуются средствами компактной вычислительной платформы CMF 3.0. Основной же функцией CMF 3.0 является возможность реализации на массивно-параллельных компьютерах совместных моделей, как единого целого.

Результаты использования CMF 3.0 и ее ранней версии CMF 2.0 (или программный комплекс для совместного моделирования – ПКСМ) в совместных моделях океан-лед [10] и океан-атмосфера [69] демонстрируют высокую вычислительную эффективность и параллельную масштабируемость CMF 2.0 (3.0). Поэтому блок усвоения данных

наблюдений реализован как часть компактной вычислительной платформы для моделирования CMF3.0. Параллельный алгоритм усвоения, базирующийся на CMF3.0, дает возможность удовлетворить высоким требованиям по параллельной и вычислительной эффективности, а также условию абстрактности (минимальная зависимость от используемой модели).

1.2 Модель динамики океана

В работе применяется численная модель динамики океана – разработка Института вычислительной математики РАН и Института океанологии им. П.П. Ширшова РАН, получившая название модель ИВМИО. Модель предназначена для исследования океанских термогидродинамических процессов в широком диапазоне пространственных и временных масштабов.

Модель принадлежит классу 3D-PEM, основана на системе уравнений трехмерной динамики океана в приближениях несжимаемости, Буссинеска и гидростатики. На границе раздела воздух-вода используется нелинейное кинематическое условие свободной поверхности с явным описанием потоков воды, тепла, соли и импульса. Уравнения аппроксимируются методом конечных объемов на сетке типа В с z-координатами по вертикали. Описание баротропной динамики выполняется с помощью двумерной системы уравнений мелкой воды, аппроксимированной по оригинальной быстрой схеме с перекрытиями [17]. Турбулентное перемешивание описывается методом Манка-Андерсона и Смагоринского с применением бигармонических фильтров и мгновенного конвективного приспособления. Обмен теплом, импульсом и влагой с пограничным слоем атмосферы задан балк-формулами CORE [75] или [88]. Для описания

динамики морского льда (наряду с моделью CICE-5.1, см. раздел 1.3) используется термодинамическая модель, основанная на работе [122]. За исключением вертикального турбулентного перемешивания [59] все процессы описаны с помощью явных численных методов. Благодаря этому стало возможным распараллеливание модели с применением метода двумерной декомпозиции области. Тестовые расчеты показали почти линейную зависимость времени работы модели от масштаба разбиения вплоть до 32400 процессорных ядер с пространственным разрешением $(1/10)^\circ$ [22]. Детальное описание модели и численных экспериментов по ее верификации с разрешением от $(1/2)^\circ$ до $(1/10)^\circ$ приведены в работах [15], [50], [51].

В данной работе постановка условий на границе раздела атмосферы и океана соответствует протоколу CORE-II, определяющему среднесуточные радиационные потоки, среднемесячные осадки и речной сток, а также суточный ход параметров атмосферы (температуры, влажности и скорости ветра на высоте 10 м) за 62-летний период 1948-2009 гг. по данным реанализа и спутниковых наблюдений. Во избежание дрейфа среднего уровня океана из глобального (двумерного, заданного только над океаном) поля суммы осадков, испарения и речного стока вычитается среднее (его осредненное по горизонтали значение) и этот результат используется при задании граничных условий для уравнения неразрывности. Модельная топография интерполировалась с данных ETOPO5 [65].

1.2.1 Уравнения термогидродинамических процессов

Модель динамики океана ИВМИО включает в себя трехмерные полные уравнения геофизической термогидродинамики [43], [58], [94]. Исследование особенностей циркуляции вод Северной Атлантики с

использованием модели ИВМИО приведено в работе [52]. Описание разработанных параллельных алгоритмов, используемых в модели ИВМИО, и ее применение для создания совместной модели Земной системы (океан-атмосфера) высокого разрешения изложено в работе [16].

Уравнения модели формулируются в трехмерной системе координат (x, y, z) . При этом поверхность $z = 0$, соответствующая невозмущенной поверхности океана, совпадает с геопотенциальной поверхностью (см. раздел 1.2.3). Задача формулируется и решается в трехмерной области Ω . Границей области является поверхность $G = \partial\Omega$, где $G = G_H \cup G_\zeta \cup G_S$. G_H – нижняя граница, определяемая уравнением $z = H(x, y)$, где $H(x, y)$ – двумерная положительная функция, описывающая топографию дна океана, $0 < H(x, y) \leq H_{\max}$; G_S – боковая твердая граница. Верхняя граница G_ζ подвижная и описывается уравнением $z = -\zeta(x, y, t)$, где $\zeta(x, y, t)$ – отклонение поверхности океана от невозмущенного уровня $z = 0$ (модельная функция уровня).

В модели ИВМИО используется следующая система уравнений, содержащая приближения несжимаемости, Буссинеска и гидростатики:

$$u_t + (\mathbf{v} \cdot \bar{\nabla}) u - f v = -\rho_0^{-1} p_x + (K_m u_z)_z + D^u \quad (1.1)$$

$$v_t + (\mathbf{v} \cdot \bar{\nabla}) v + f u = -\rho_0^{-1} p_y + (K_m v_z)_z + D^v \quad (1.2)$$

$$p_z = \rho g \quad (1.3)$$

$$(\bar{\nabla} \cdot \mathbf{v}) = 0 \quad (1.4)$$

$$T_t + (\mathbf{v} \cdot \bar{\nabla}) T = (K_h T_z)_z + D^T + (c_p \rho)^{-1} Q^H \quad (1.5)$$

$$S_t + (\mathbf{v} \cdot \bar{\nabla}) S = (K_h S_z)_z + D^S \quad (1.6)$$

$$\rho = \rho(T, S, p) \quad (1.7)$$

Эволюция верхней границы океана с учетом потока массы (осадков, таяния льда, испарения) определяется уравнением свободной поверхности океана [23]:

$$\zeta_t + u \zeta_x + v \zeta_y + w = \rho_f^{-1} W \quad (1.8)$$

В уравнениях (1.1) - (1.8) используются обозначения:

$\mathbf{v} = (u, v, w)$ – трехкомпонентный вектор скорости течений;

T, S – температура и соленость морской воды;

ρ – плотность морской воды; ρ_0 – средняя плотность морской воды; ρ_f – плотность пресной воды;

f – параметр Кориолиса;

K_v, K_h – коэффициенты вертикальной турбулентной вязкости и диффузии;

D^u, D^v, D^T, D^S – члены, описывающие горизонтальную турбулентную вязкость и диффузию тепла и соли;

Q^H – источник тепла; c_p – теплоемкость морской воды; W – поток воды на границе атмосфера-океан.

Источник тепла определяется уравнением:

$$Q^H = -I_z \cdot (1 - A) \quad (1.9)$$

В формуле (1.9) I – поток проникающей солнечной радиации; A – сплоченность морского льда.

Поток массы на границе атмосфера-океан равен:

$$W = P + M - E + R \quad (1.10)$$

Тут P, E и M – интенсивности осадков, испарения воды и таяния льда, соответственно; R – речной сток.

1.2.2 Краевые условия

Граничные условия на поверхности океана G_ζ ($z = -\zeta(x,y,t)$, граница атмосфера-океан) с явным учетом потока воды [74] следующие:

$$-K_m(u_z, v_z) + (u, v) \cdot \rho_f^{-1} W = \rho_0^{-1} (1 - A) (\tau^x, \tau^y) \quad (1.11)$$

$$p = p_a \quad (1.12)$$

$$-c_p K_h T_z + c_p (T - T_p) \rho_f^{-1} W = \rho_0^{-1} [Q_h^{aw} (1 - A) + Q_h^{iw} A] \quad (1.13)$$

$$-K_h S_z + S T \rho_f^{-1} W = \rho_0^{-1} S^{iw} M A \quad (1.14)$$

В уравнениях (1.11) - (1.14) приняты следующие обозначения:

$\tau_x(x, y, t)$, $\tau_y(x, y, t)$ – касательное напряжение трения ветра;

T_p – температура осадков;

$Q_h^{aw}(x, y, t)$ и $Q_h^{iw}(x, y, t)$ – поток тепла на границе воздух–вода и на границе лед-вода, соответственно;

S^{iw} – соленость льда;

$S^{iw}M$ – интенсивность выделения соли в океан при таянии льда.

На дне океана ($z = H(x, y)$, граница G_H) задаются условия:

$$-w + u H_x + v H_y = 0 \quad (1.15)$$

$$-\rho_0 K_m \frac{\partial \mathbf{v}_\tau}{\partial \mathbf{n}} = (\tau_B^x, \tau_B^y) \quad (1.16)$$

$$-\rho_0 c_p K_h T_n = 0 \quad (1.17)$$

$$-\rho_0 K_h S_n = 0 \quad (1.18)$$

В уравнении (1.16) $\tau_B^x(x, y, t)$, $\tau_B^y(x, y, t)$ – компоненты напряжения трения о дно.

Граничные условия на твердых боковых границах (G_S):

$$\mathbf{v}_n = 0 \quad (1.19)$$

$$\frac{\partial \mathbf{v}_\tau}{\partial \mathbf{n}} = 0 \quad (1.20)$$

$$-\rho_0 c_p K_h T_n = 0 \quad (1.21)$$

$$-\rho_0 K_h S_n = 0 \quad (1.22)$$

В условиях (1.15) - (1.22) \mathbf{n} , $\boldsymbol{\tau}$ - вектор внешней нормали и касательный вектор к границе соответствующей области.

Для системы уравнений (1.1) – (1.8) формулируются следующие начальные условия:

$$(\mathbf{u}, \mathbf{v}, T, S, \zeta)|_{t=0} = (\mathbf{u}^0, \mathbf{v}^0, T^0, S^0, \zeta^0) \quad (1.23)$$

1.2.3 Система координат. Сетка

Выбор системы координат (СК) для описания динамических процессов в океане обсуждается в работах [8], [74], [132]. В задачах моделирования океана вместо сферической СК (λ, φ, r) обычно используется более удобная СК, в которой вертикальная координата r заменяется на z – расстояние точки от поверхности общеземного эллипсоида вращения, с которым связана географическая СК. Также принимается «сферическое приближение». Для этого вводится сфера S_R радиуса R – «средний радиус Земли» ($R = 6367456$ м), центр которой совпадает с центром масс Земли и центром общеземного эллипсоида. Тогда $z = (R - r)$, т.е. ось z направлена к центру сферы S_R . Дополнительная аппроксимация при написании уравнений гидродинамики океана заключается в замене r на R . Система координат (λ, φ, z) называется географической СК, часто упоминается также название «широтно-долготная СК».

Поверхность $z = 0$, соответствующая невозмущенной поверхности океана, с большой точностью совпадает с геопотенциальной поверхностью. При таком определении оси z вектор силы тяжести, доминирующей силы в уравнениях движения, параллелен оси z и не имеет составляющих перпендикулярных ей. В работе [8] показывается, что уравнения гидродинамики океана в координатах сплющенного сфероида $(\lambda^*, \varphi^*, r^*)$ с большой точностью будут такими же, как и в сферических координатах. В свою очередь сплюснутый сфероид является хорошим приближением геопотенциальной поверхности на уровне моря.

Традиционно уравнения гидродинамики Мирового океана решались в широтно-долготной СК (λ, φ) . Однако в этом случае в области решения оказывается точка Северного полюса. Решению проблемы особой точки широтно-долготной системы координат на полюсе посвящено достаточно много работ, например [99].

В последнее время в глобальных моделях океана наибольшее распространение получили биполярная и трехполярная СК. Обе они относятся к классу ортогональных. Биполярная СК применяется в таких моделях океана, как POP [108], HYCOM [55], INMSOM [11]. Решение уравнений динамики океана в трехполярной системе координат применяется в модели Мирового океана MOM4.0 [73], HIM [76]. Для Мирового океана в модели ИВМИО уравнения решаются в трехполярной системе координат [48].

1.2.4 Алгоритм решения уравнений динамики

Задача моделирования динамики океана заключается в нахождении временной эволюции вектора-состояния модели, состоящего из $(u, v, w, T, S, \rho, p, \zeta)$, при указанных параметризациях турбулентной

вязкости и диффузии и заданных начальных и граничных условиях, в том числе атмосферном воздействии (атмосферном форсинге).

Система уравнений (1.1) – (1.8) включает в себя широкий спектр движений. Принято выделять два основных диапазона характерных скоростей:

- скорости течений и бароклинных гравитационных волн ($1-3$ м/с);
- скорости баротропных гравитационных волн ($30 - 200$ м/с):

$$c = \sqrt{gH}, \quad g \approx 10 \text{ м/с}^2, \quad H \sim 100 - 5000 \text{ м} \Rightarrow c \sim 30 - 200 \text{ м/с} \quad (5.24)$$

Применение в модели ИВМИО явных по времени методов аппроксимации для уравнений (1.1), (1.2), (1.3) вызывает необходимость использовать достаточно малые шаги по времени, чтобы удовлетворить условию Куранта для всего диапазона описываемых системой уравнений волн, включая баротропные гравитационные волны. На сегодняшний день полагают, что наиболее значимые физические процессы для климатических явлений в океане определяются течениями и бароклинными волнами [48].

Поэтому решение динамической подсистемы уравнений (1.1) – (1.8) разбивается на две части – решение трехмерных уравнений для бароклинных движений и решение двумерных уравнений мелкой воды для баротропных движений [119].

В модели ИВМИО применяется явная аппроксимация по времени для решения системы уравнений мелкой воды, рассмотренная в работе [17]. Предполагаемые размеры модельной сетки для Мирового океана составляют порядка $10^4 \times 10^4 \times 10^2$ точек. Наиболее эффективным методом решения уравнений на массивно-параллельных компьютерах с распределенной памятью является метод декомпозиции области.

Разбиение на вертикальные расчетные подобласти нерационально из-за небольшого характерного размера (100 точек). Использование явных методов в модели ИВМИО при аппроксимации горизонтальных пространственных операторов для решения дифференциальных уравнений задачи позволяет эффективно применять метод декомпозиции расчетной области по двум горизонтальным координатам [16].

1.3 Термодинамическая модель льда

Морской лед в совместной модели океан-лед описывается с помощью модели CICE-5.1 Лос-Аламосской национальной лаборатории США [77]. В работе [10] рассматриваются результаты моделирования внутригодовой изменчивости циркуляции вод и льда Северного Ледовитого океана с применением совместной модели, построенной на основе трехмерной модели динамики океана ИВМИО и модели морского льда CICE5.1. Модель морского льда CICE широко используется научными группами по всему миру, в частности применяется и в отечественных работах, например в исследованиях ИВМиМГ СО РАН [9]. Модель описывает состояние льда и снега в терминах функции их распределения $g(t,x,h)$ во времени, географическом пространстве и пространстве толщин, при этом для льда отдельно описываются несколько (по умолчанию 5) категорий толщины. Основными прогностическими переменными для каждой категории являются сплоченность льда, средние по ячейке толщины льда и снега, внутренние энергии льда и снега, соленость льда и температура поверхности [77].

В работе [10] модель CICE используется в термодинамическом режиме, рассчитывающем локальные процессы роста, таяния и взаимного преобразования льда и снежного покрова в соответствии с вертикальными потоками и осадками, диффузионного переноса тепла, распространения

радиации разных диапазонов, турбулентного обмена теплом с атмосферой и океаном, испарения и соответствующего потока скрытого тепла, обмена водой и солью с океаном и другие процессы, без учета пространственного перемещения и механических напряжений. Термодинамика льда описывается в пористом приближении, согласно которому морской лед рассматривается как среда чистого пресного льда, заполненная вкраплениями рассола, задающими распределение солености и энтальпии [71]. Для описания теплопроводности применяется пузырьковая параметризация [112]. Дополнительно к описанию снежного покрова модель воспроизводит эволюцию снежиц в виде трасеров, в том числе просачивание из них талой воды в океан. Поглощение солнечной радиации описывается по схеме, основанной на приближении дельта - Эддингтона [79].

1.4 Компактная вычислительная платформа для моделирования

Развитие программных средств для решения задач моделирования Земных систем происходит в первую очередь за счет эволюции совместных моделей океан-атмосфера. Первые системы объединяли вычислительные модели океана и атмосферы напрямую и они работали как единая программа. По мере усложнения программ-моделей появилась необходимость в отдельном сервисном компоненте — каплере (англ. couple — соединять), который занимался бы интерполяцией модельных полей на соответствующие сетки при передаче данных между компонентами. Изначально эта задача решалась через файловую систему: одна компонента сохраняет файл с данными, потом отдельная программа его интерполировала на нужную модельную сетку, а потом другая модель считывала интерполированный файл. Потом этот алгоритм был выделен в отдельную последовательную программу, представляющую собой аналог

центрального хаба для связи всех моделей. По мере увеличения разрешения сеток моделей последовательные алгоритмы каплера становились неэффективными, и на их место пришла полностью параллельная архитектура [16], [19].

Компактная вычислительная платформа для моделирования CMF (Compact Modeling Framework) - это программный комплекс для совместного моделирования Земной системы и ее компонентов высокого разрешения на массивно-параллельных вычислительных системах. За счет использования абстрактных интерфейсов главная программа и каплер становятся полностью независимыми от количества подключаемых моделей — для работы в совместной системе пользователю достаточно создать производный класс своей модельной компоненты [19].

Версия CMF 2.0 компактной вычислительной платформы для моделирования показала себя пригодной для создания моделей высокого разрешения, она имела несколько направлений для улучшения. Во-первых, хотя чистый MPI подход к коммуникациям имеет высокую скорость работы, он требует явной работы с буферами данных. Во-вторых, разработка региональных подмоделей морей, вложенных в сетку глобальной модели, становились довольно сложными при использовании только MPI-процедур. Результаты работы CMF 2.0 показали, что мы можем пожертвовать частью производительности для выбора более простой (и возможно менее эффективной) абстракции для упрощения коммуникационных алгоритмов.

На программный комплекс совместного моделирования системы океан-лед-атмосфера-почва на массивно-параллельных компьютерах оформлено свидетельство о Государственной регистрации программы для ЭВМ [18].

1.4.1 PGAS-коммуникатор

В версии CMF 3.0 [20],[80] используется библиотека Global Arrays (GA) [104], реализующая парадигму PGAS (Partitional Global Address Space). Библиотека позволяет обращаться к глобальным индексам массива, как будто он весь доступен в локальной памяти.

CMF 3.0 содержит класс Communicator, который инкапсулирует логику работы с библиотекой и предоставляет интерфейс для put/get операций частей глобальных данных различных компонентов. Оказалось, что такой подход позволяет упростить не только взаимодействие вложенных компонентов, но и предоставляет удобную замену системе обмена данными, основанной на прямом MPI-подходе между компонентами и кластером.

В итоге, все обмены между частями системы реализованы с использованием класса Communicator. Он содержит хэш-таблицу для хранения всей информации о массивах, включая их состояние и метаданные. Каждый массив компонента, участвующий в обменах, содержит распределенную копию, хранящуюся в виде виртуального глобального массива GA. Когда процесс должен отправить данные, он заполняет эту копию своими текущими данными. Благодаря тому, что распределение глобального массива полностью повторяет декомпозицию компонента, эта операция происходит локально.

На компактную вычислительную платформу CMF3.0 для моделирования Земной системы на массивно-параллельных компьютерах оформлено свидетельство о Государственной регистрации программы для ЭВМ [21].

1.4.2 Новая архитектура совместной модели

Поскольку сложность совместной системы растет, необходим более удобный способ объединения компонентов. Первоначально появившаяся для веб-приложений СОА (Сервис-Ориентированная Архитектура) предоставляет хороший фундамент для решения подобной задачи.

В CMF 3.0 все модели отправляют свои запросы общего вида в единую очередь сообщений (**Рисунок 1.1**). Сервисные компоненты берут из этой очереди только сообщения, которые могут обработать, забирают данные из виртуальных глобальных массивов и выполняют соответствующие действия. Архитектура позволяет минимизировать связи между физическими и сервисными компонентами и значительно упростить разработку. Более того, поскольку все сервисы наследуют общий базовый класс Service, добавление нового сервиса не представляет сложностей. Сейчас CMF 3.0 содержит следующие независимые параллельные сервисы: CPL (операции интерполяции данных при пересылке между модельными компонентами на различных сетках), IOS/IOF (I/O Fast, I/O Slow — быстрые и медленные устройства работы с файлами), DAS (подробнее сервис усвоения данных наблюдений рассматривается в **Главе 3**).

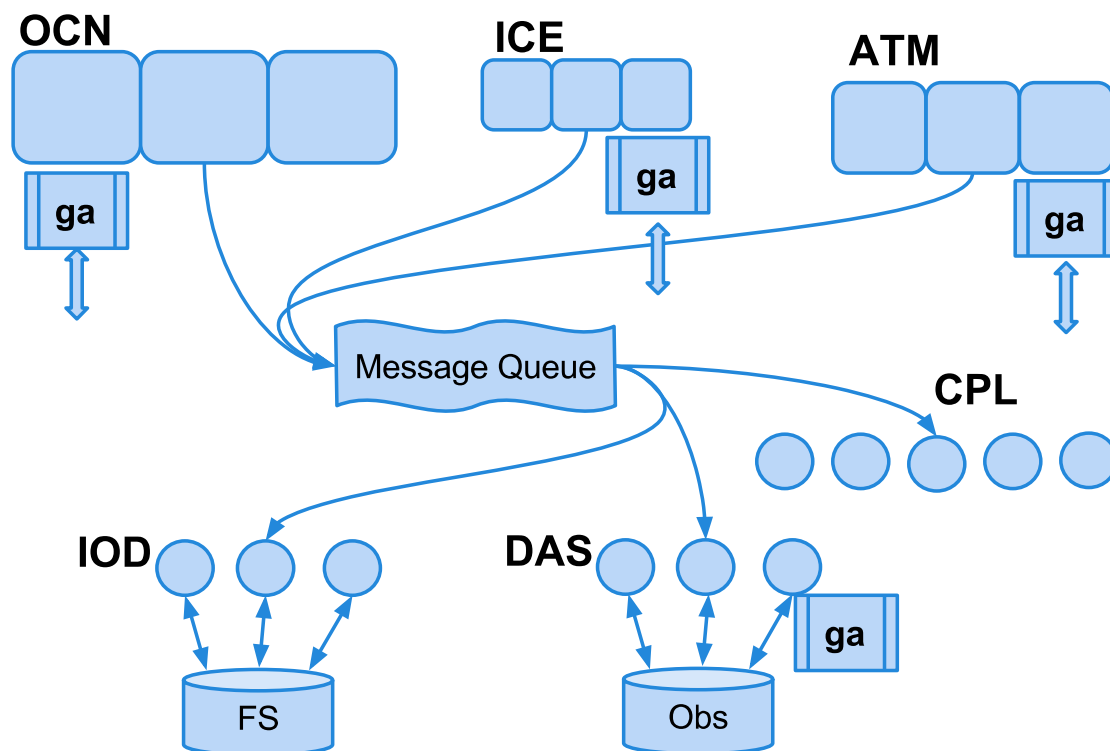


Рисунок 1.1. Сервис усвоения данных DAS в архитектуре компактной вычислительной платформы для моделирования CMF 3.0. В данном примере три компонента (океан, лед, атмосфера) отправляют запросы в очередь сообщений, откуда их извлекают сервисы каплера (CPL), усвоения данных (DAS) и ввода-вывода (IOD). Сами данные передаются через механизм глобальных массивов, который также используется для межпроцессорных обменов в компонентах и в сервисе DAS.

1.4.3 CMF 3.0 : интерполяция

Несмотря на то, что логика работы процедур интерполяции каплера в CMF 3.0 осталась той же, что и в CMF 2.0 [19], абстракция PGAS

позволила сильно упростить код. Теперь все данные, необходимые процессу от соседей, получаются с использованием класса Communicator. Недостатком такого подхода является некоторое падение производительности, связанное с невозможностью использования отложенных MPI-операций и наличия у библиотеки GA собственных издержек. Однако результаты работы SMF 2.0 показали, что мы можем пожертвовать частью производительности для выбора более простой (и возможно менее эффективной) абстракции для упрощения коммуникационных алгоритмов [80].

Поскольку, во-первых, хотя чистый MPI-подход к коммуникациям имеет высокую скорость работы, он требует явной работы с буферами данных, что значительно усложняет разработку и совершенствование ПКСМ. Во-вторых, разработка региональных подмоделей морей, вложенных в сетку глобальной модели, становились довольно сложными при использовании только MPI-процедур.

1.5 Выводы

Создание системы усвоения данных в модели глобального океана высокого разрешения (0.1° и выше, что позволяет воспроизводить мезомасштабные вихри) для решения задач сезонного прогноза и исследования климата вызывает необходимость решить ряд проблем, связанных с реализацией системы на массивно-параллельных компьютерах. Они включают в себя разработку следующих подсистем:

1. Модель динамики океана, аппроксимирующая полную 3-мерную систему уравнений гидротермодинамики океана и эффективно масштабируемая на десятках тысяч процессорных ядер. В качестве

такой модели планируется использовать вихререзрешающую модель динамики океана ИВМИО.

2. Система ввода-вывода и передачи между компонентами Земной системы массивов полей модельного решения и данных наблюдений, включающая алгоритмы их постпроцессинга, анализа, хранения и представления в виде, пригодном для интерпретации конечным пользователем. Для решения этой подзадачи будет использоваться CMF 3.0.
3. Параллельный алгоритм усвоения рационально реализовать на базе CMF3.0, что дает возможность удовлетворить высокие требования по параллельной и вычислительной эффективности, условию абстрактности (минимальной зависимости метода усвоения от модели).

Глава 2. Характеристики данных наблюдений об океане и обзор существующих систем прогноза

2.1 Введение

В настоящее время в мире разработано и функционирует несколько систем глобального и регионального прогноза высокого пространственно-временного разрешения, их обзор приведен в [106]. Физически обусловленным требованием к таким моделям является воспроизведение вихревой динамики океана, для чего модель должна иметь пространственное разрешение 0.25° и выше. В **таблицу 2.1** сведены характеристики основных существующих систем оперативного прогноза состояния океана с указанием используемой модели океана и алгоритма усвоения данных. Из анализа этой таблицы видно, что многие системы используют 3DVar или оптимальную интерполяцию (OI), которые по своей сути являются весьма простыми методами и были предложены более 50 лет назад [7], [120]. Эти методы хорошо распараллеливаются и поэтому вполне применимы для глобальных моделей, в то время как 4DVar метод, в силу вычислительной сложности, на сегодняшний день не используется ни в одной глобальной модели оперативного прогноза океана высокого пространственного разрешения (см. обзоры в [72], [106], [109], [110], [121]). В некоторых системах используются относительно недавно появившиеся ансамблевые методы усвоения данных [67].

Таблица 2.1. Список систем оперативного прогноза состояния океана с усвоением данных наблюдений из работы [121]

Название системы Страна	Модель океана	Область	Горизонтальное разрешение	Верт. уровни	Метод усвоения
BLUElink> Австралия	MOM 4	Глобальная	1 °	47 z-ур.	EnOI
		Региональная	1/10°		
C-NOOFS Канада	NEMO	Сев. Атлантика	1/4°	50 z-ур.	Нет
ECCO-JPL США	MIT OGCM	Глобальная	l'x 0.3°	46 z-ур.	Kalman Filter+ Smoother
FOAM Великобритания	NEMO	Глобальная	1/4°	50 z-ур.	Analysis correction
		Региональная	1/12°		
NCODA США	HYCOM	Глобальная	1/12°	32 гибрид. ур.	Multivariate OI 3DVar
MERCATOR Франция	NEMO	Глобальная	1/4°+1/12°	50 z-ур.	SEEK Filter
		Региональная	1/12°		
MFS Италия	NEMO	Средиземное море	1/16°	71 z-ур.	3DVar
Move/MRI Япония	MRI.com	Глобальная	1 °	50 z-ур.	3DVar
		Региональная	1/2°+1/10°		
NCOM США	POM	Глобальная	1/8°	42 гибрид. ур.	2D-OI+3DVar
NMEFC Китай	Lap/Cas	Тихий ок.	2°x 1°	14 z- уровней	3DVar
RTOFS США	HYCOM	Сев. Атлантика	4-18 km	26 гибрид. ур.	3DVar
TOPAZ Норвегия	HYCOM	Сев. Атлантика + Арктика	11-16 km	22 гибрид. ур.	EnKF 100 members
REMO Бразилия	HYCOM	Сев. Атлантика	1/12°	21 гибрид. ур.	EnOI
ECMWF Европа	NEMO	Глобальная	1°	42 z-ур.	NEMOVAR

Точность ансамблевых методов на основе EnKF согласно результатам, опубликованным в статьях [82], [135], находится на уровне 4DVar, но при этом такие методы значительно лучше архитектурно подходят для

массивно-параллельных компьютеров с распределенной памятью. Вышесказанное обусловлено тем, что ансамбль прогнозов в методе EnKF может считаться параллельно, и нет необходимости в последовательном интегрировании модели для поиска минимума функционала ошибки с учётом эволюции решения внутри окна усвоения (*eng. assimilation window*). Результаты, также подтверждающие схожую и для некоторых данных превосходящую эффективность EnKF по сравнению с методом 4DVar, приведены в работе [96] на примере модели прогноза погоды JMA [78]. В работе [133] приводятся результаты, показывающие значительное преимущество EnKF над 3DVar на примере модели атмосферы NCEP. На основании исследования [111], где показано существенное уменьшение ошибки прогноза в ансамблевом методе усвоения данных наблюдений по сравнению с 3DVar, был выполнен переход с 3DVar на гибридный метод усвоения данных, основанный на EnKF в системе GODAS [54].

В работе [118] на примере модели MOM4 разрешения 0.1° для Тасманского моря приводится экспериментальное сравнение методов EnOI и EnKF. На **рисунке 2.1** из этой работы видно, что графики MAD (*eng. mean absolute deviation*), полученные по данным спутниковой ТПО, которые наряду с данными T и S с дрейфтеров ARGO и данными спутниковой альтиметрии усваиваются в модели этими двумя методами, отличаются на 10 %⁴. Похожая ситуация наблюдается для остальных усваиваемых данных наблюдений, так что для данных T, S и SLA разница в ошибке MAD не превышает 17 %. Тот факт, что модель ИВМИО очень похожа на модель MOM4 (явные разностные схемы, В-сетка по горизонтали, z-координата по вертикали, одинаковое пространственно-временное разрешение) дает нам право предполагать, что эффективность

4 Для данных ТПО $\frac{\text{MAD}(\text{EnKF}) - \text{MAD}(\text{EnOI})}{\text{MAD}(\text{EnKF})} \approx 10\%$

метода EnOI также не будет сильно отличаться от метода EnKF, по крайней мере разница в ошибке не оправдывает возрастающую в N раз (количество элементов ансамбля) вычислительную стоимость такой системы усвоения данных.

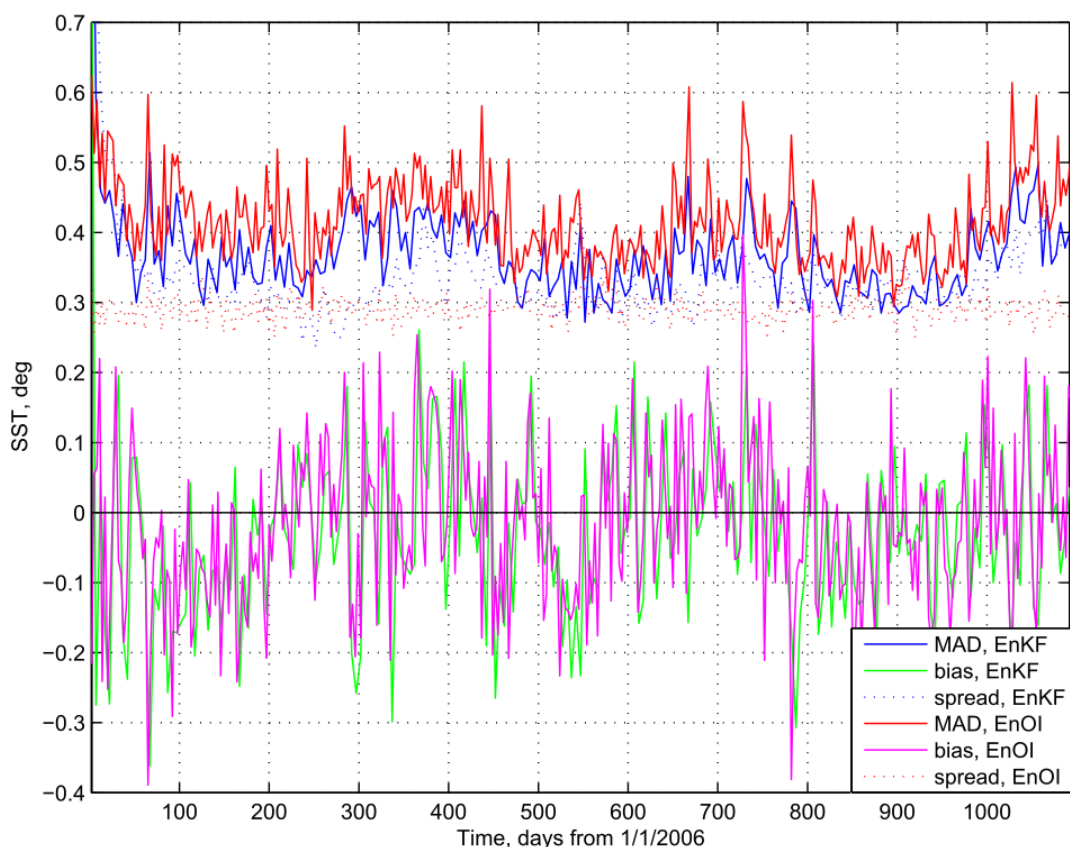


Рисунок 2.1. Статистика ошибки суточных прогнозов ТПО методом EnKF и EnOI в модели MOM4 разрешения 0.1° для Тасманского моря. Условные обозначения: MAD – средняя абсолютная ошибка ; bias – средняя ошибка. Рисунок взят из работы [118].

Модель ИВМИО непрерывно развивается [14], [15], [50], [52], в частности добавляются новые параметризации, например параметризации вертикального перемешивания, вертикальной диффузии и вязкости, изменяются используемые разностные схемы, подключается модель льда CICE5.1 [10]. При этом в модели за исключением вертикального турбулентного перемешивания все процессы описаны с помощью явных

численных методов, как следствие нет необходимости в вычислении и хранении в явном виде оператора модели. В силу вышесказанного, вывод и поддержание в актуальном состоянии системы прямой и сопряжённой задачи, которые необходимы для использования метода 4DVar в модели ИВМИО становится очень трудным вопросом, который даже в случае его решения будет требовать постоянного внимания, вследствие развития модели.

Метод 4DVar все же используется совместно с некоторыми моделями динамики Мирового океана, например в системе Ocean-S4 [97], состоящей из модели NEMO [90] и системы усвоения данных NEMOVAR [98], и используется в организации ECMWF. Система внедрена в 2011 году, и основной целью работы Ocean-S4 является получение данных реанализа для всего Мирового океана. Эти данные могут использоваться для задания начальных условий для сезонных прогнозов, которые выполняются с использованием совместных моделей динамики океана и атмосферы, а также для изучения изменчивости климата и проверки климатических моделей. Нужно отметить, что система имеет невысокое по современным меркам горизонтальное разрешение - порядка 1° , что не позволяет явным образом воспроизводить вихревые структуры течений в океане и использовать данную систему для оперативного прогноза его состояния.

В институте ИВМ РАН созданы информационно-вычислительные системы для оперативного прогноза динамики Черного [2] и Балтийского моря [1] с использованием процедуры четырехмерной вариационной ассимиляции (4DVar) данных наблюдений (спутниковых, корабельных и других). Небольшой размер расчетной области (порядка 10^6 расчетных узлов) позволяет использовать разработанную систему даже на персональном компьютере.

Еще одно преимущество ансамблевых методов усвоения данных — это возможность их использования в совместных моделях океана и атмосферы с построением кросковариационных матриц океанических и атмосферных данных наблюдений, чтобы данные наблюдений полученные для одной из компонент могли напрямую влиять на состояние другой компоненты при получении анализа. В статье [124] демонстрируется преимущество такого подхода по сравнению с отдельным усвоением данных в каждой из компонент, работающих в рамках общей модели Земной системы (океан-атмосфера). Схема работы такого подхода на примере совместной модели Земной системы состоящей из модели океана (NEMO [90]) и атмосферы (SPEEDY[124]) приведена на **рисунке 2.2**. В качестве метода усвоения используется локальный ансамблевый фильтр Калмана с преобразованием ансамбля (*eng. Local Ensemble Transform Kalman Filter, LETKF*). В данном методе используется локализация наблюдений: для анализа в каждой точке модельной сетки используются только данные наблюдения, расположенные в локальной области вокруг этой точки. Это позволяет частично решить проблему ложных ковариаций между удаленными (далеко расположенными) точками, возникающую из-за оценки матрицы ковариаций ошибок в пространстве существенно меньшей размерности (ограниченной размером ансамбля) [107].

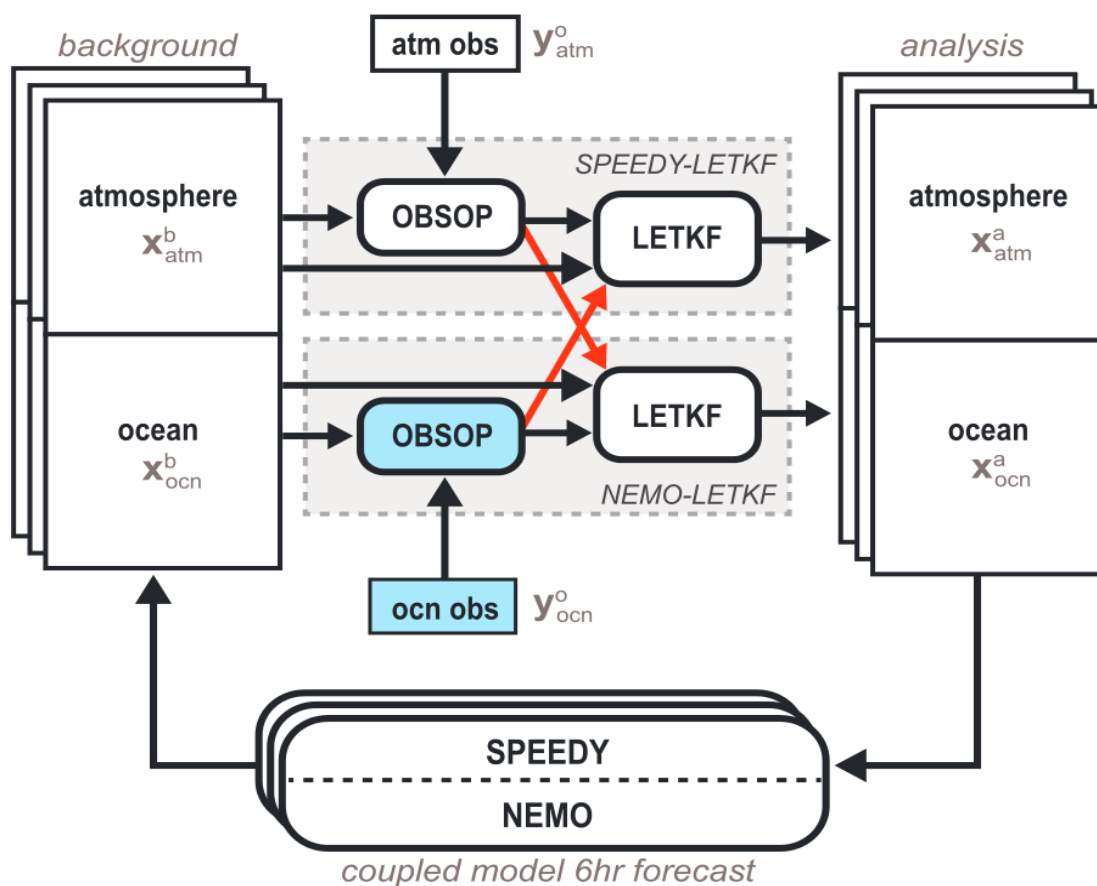


Рисунок 2.2. Схема работы метода LETKF для совместной модели Земной системы, состоящей из модели океана (NEMO) и атмосферы (SPEEDY). Красными стрелками показана частичная передача данных наблюдений между компонентами для их совместного усвоения методом LETKF. Рисунок взят из работы [124].

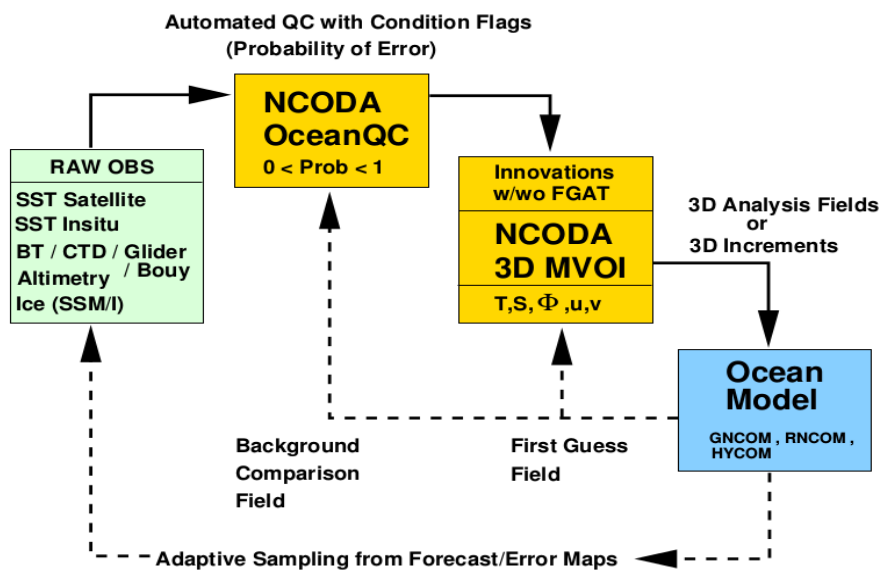


Рисунок 2.3. Схема функционирования системы усвоения NCODA совместно с глобальной моделью океана. Рисунок взят из [89].

Схематически процесс усвоения различных данных наблюдений системой NCODA в глобальную модель океана показан на **рисунке 2.3**. Различные исходные данные наблюдений обрабатываются NCODA OceanQC (Quality Control), и оценивается вероятность ошибки в каждом значении (POE - “probability of error”). Данные вместе со своим POE передаются в NCODA для анализа. В результате анализа значения с низким POE отбрасываются, и генерируются 3-мерные (3D) и 2-мерные (2D) корректирующие поля соответствующих физических величин. Модель океана усваивает эти поля и продолжает работу. Пример корректирующего поля для ТПО показан на **рисунке 2.4**.

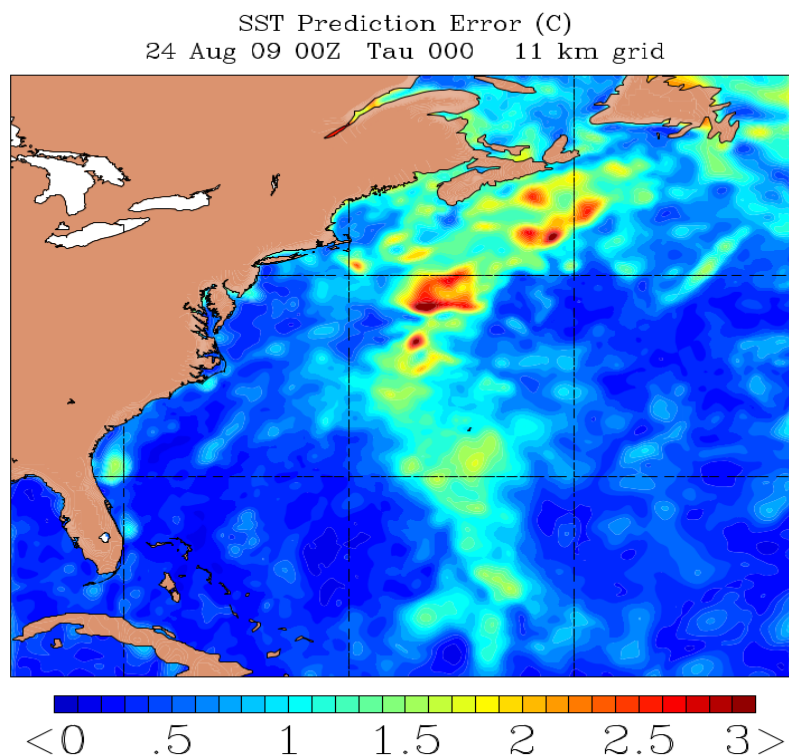


Рисунок 2.4. Пример поля для коррекции ТПО, вычисленного системой NCODA с помощью спутниковых данных наблюдений. Рисунок взят из [89].

Параллельная реализация для компьютеров с распределенной памятью (см. таблицу 2.2) позволяет работать данной системе совместно с глобальной моделью NUCOM с разрешением $(1/12)^\circ$. В то же время из таблицы видно, что масштабируемость NCODA далеко не идеальная.

Таблица 2.2. Оценка масштабируемости системы NCODA [89]

Область	Увеличение количества процессоров	Ускорение	Параллелизм
Глобальная	С 64 до 128 ядер	1.62	38 %
Глобальная	С 128 до 256 ядер	1.26	21 %

2.2 Данные для усвоения

2.2.1 Данные ARGO

ARGO - это система наблюдения за температурой, соленостью и течениями в Мировом океане, действующая с начала 2000-х годов [70]. Дрифтеры⁵ ARGO разбрасываются с исследовательских судов практически по всему Мировому океану. Данные ARGO усваиваются в модели динамики океана, разрабатываемые участниками проекта GODAE OceanView и другими прогностическими центрами, что позволяет существенно повысить качество моделирования океана. В рамках проекта ARGO систематически измеряется физическое состояние верхнего слоя океана, при этом данные выдаются почти в режиме реального времени с задержкой всего несколько часов.

Во второй половине XX века большая часть измерений выполнялась вдоль основных маршрутов судоходства, которые могли меняться в зависимости от сезона или в ходе выполнения специальных научно-исследовательских программ, например «ПОЛИМОДЕ» [35] и «РАЗРЕЗЫ»

⁵ Термином «дрифтер» будем обозначать измерительное устройство дрейфующее по океану, которое периодически погружается для измерения геофизических параметров (проект ARGO); «буй» - это прибор, плавающий и производящий измерения на поверхности океана (как в проекте PIRATA или TOGA-TAO)

[42], [44], которые, несмотря на свою значимость и трудоёмкость, носили единичный характер. Проект ARGO позволил уйти от этого ограничения. Академик А.С. Саркисян, непосредственный участник программ «ПОЛИМОДЕ» и "РАЗРЕЗЫ", в своих последних работах [36], [45] отмечал крайнюю важность развития наблюдательной сети ARGO для всей океанологии в целом.

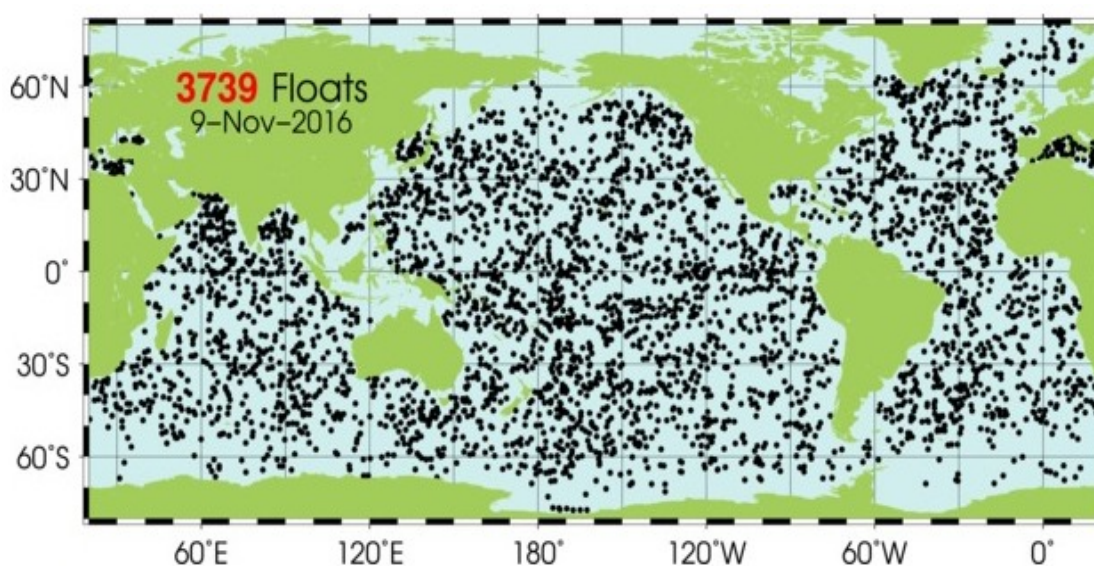


Рисунок 2.5. Расположение дрейфтеров ARGO, которые предоставляли данные за 30-дневный временной интервал (<http://www.argo.ucsd.edu>)

Глобальный массив из более 3500 дрейфтеров (**рисунок 2.5**) функционирует практически по всему Мировому океану (среднее расстояние между дрейфтерами 300 км) и обеспечивает свыше 100 000 профилей температуры, солёности (T / S) и измерений скорости в год. Сеть дрейфтеров ARGO является единственным глобальным источником физических данных о глубинном слое океана (до 2000 м) и поэтому используется во всех моделях, ассимилирующих океанические данные наблюдений. Каждый дрейфтер ARGO погружается на глубину до 2000

метров и передает данные каждые 10 дней. Срок службы каждого такого дрейфтера около 5 лет. Все данные, собранные в рамках проекта, доступны через глобальную сеть Интернет после автоматизированного контроля качества (QC).

2.2.2 Спутниковые данные по температуре поверхности океана (ТПО)

Существуют различные способы измерения ТПО: прямые измерения «in situ» с кораблей и буев; непрямые измерения средствами дистанционного зондирования Земли из космоса. Каждый из способов измерения ТПО имеет свои достоинства и недостатки [114].

Прямые контактные измерения позволяют определять ТПО с высокой точностью и частотой, в любых погодных условиях, но не обладают достаточным пространственным охватом. Эти данные целесообразно использовать для калибровки и оценки спутниковых данных.

Космические методы позволяют получать информацию о ТПО в глобальном масштабе с использованием инфракрасных и микроволновых радиометров, установленных на различных спутниковых системах [95].

Исходными данными для формирования файлов, пригодных для усвоения в модели динамики Мирового океана, являются стандартные картированные изображения (SMI), представляющие собой продукт третьего уровня обработки (L3) космических изображений с мультиспектральной аппаратуры MODIS, размещенной на спутниках TERRA и AQUA.

С помощью инфракрасных датчиков, работающих в диапазонах 3,5-3,9 мкм и 10-12,5 мкм (AVHRR, VIIRS, MODIS и др.), определяется

температура «скин-слоя» поверхности океана. Такие датчики позволяют проводить измерения с высокой точностью и пространственным разрешением 1-10 км, но чувствительны к влиянию суточного цикла, содержанию аэрозолей в атмосфере и не позволяют проводить измерения в районах со сплошной облачностью.

Микроволновые датчики, работающие в полосе частот от 6 до 11 ГГц (AMSR-E, WindSat и др.), позволяют определить температуру океана на глубине 1 м. Они предоставляют возможность проводить круглосуточные всепогодные измерения (за исключением областей с интенсивными осадками), однако, их разрешение не превышает 50 км, и они также подвержены влиянию внешних параметров (ветер, телевизионное вещание).

Данные о ТПО предоставляются различными ведомствами и компаниями (NOAA, NASA, MetOffice EU, U.S. Navy и др.) в различном качестве и уровнях обработки (L2, L3, L4).

2.2.3 Данные спутниковой альтиметрии

Измерение спутниковой альтиметрии - важная часть космического мониторинга Земли. При помощи анализа и усвоения данных спутниковой альтиметрии возможно обнаружение волн цунами, значительное изменение солёности и плотности морской воды. Эффективное решение задачи измерения альтиметрии требует привлечения оперативных и прецизионных средств измерения высоты, устанавливаемых на ИСЗ (Искусственный спутник Земли) [12]. При этом время поставки альтиметрических данных потребителю исчисляться единицами часов, а суммарная погрешность измерения высоты радио- и лидарными

высотомерами, такими как TOPEX/Poseidon, Jason-1,2, ERS-1, ERS-2, LITE, SLA, GLAS, составляет 10-12 см [115]. В качестве примера данных спутниковой альтиметрии на **рисунке 2.6** показаны данные аномалии уровня океана со спутника Jason-2. Данные спутниковой альтиметрии наряду с данными о температуре и солености с систем плавучих буев и данными спутниковой ТПО используются практически во всех ранее упомянутых центрах, занимающихся оперативным прогнозом состояния океана, таких как NOAA, MetOffice, U.S. Navy, MERCATOR [72].

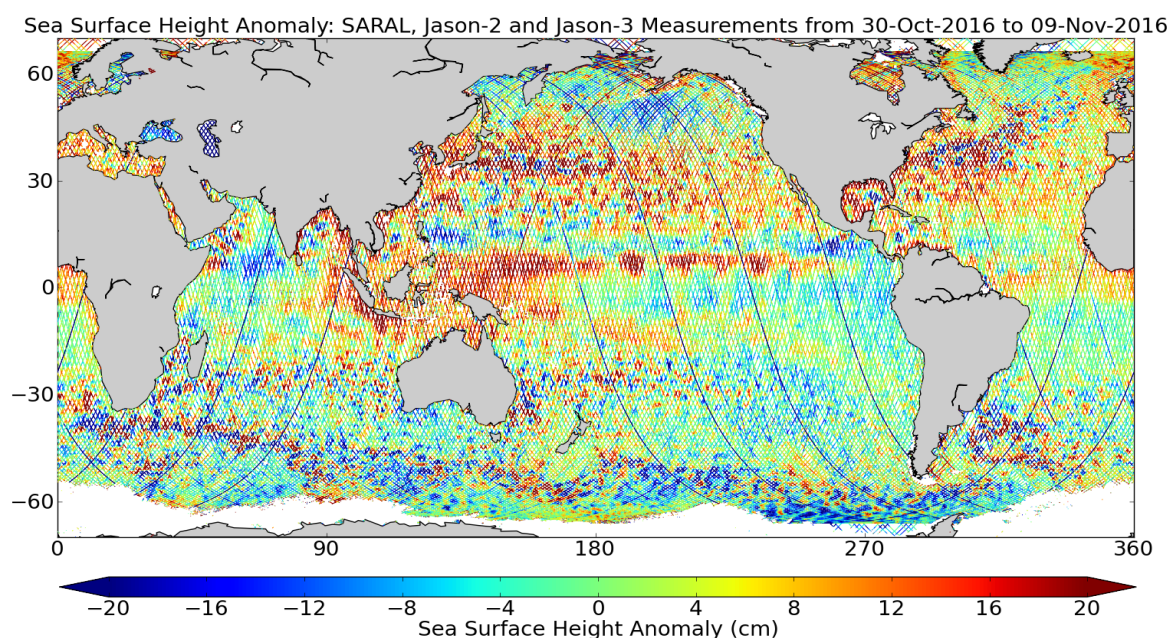


Рисунок 2.6. Показана суточная выборка измерений аномалии уровня океана (SLA), полученная в режиме приближенного к реальному времени (near-real time) вдоль трека спутника Jason-2 [100].

Нужно отметить, что в рамках проекта AVISO используются модели приливов [113]. Полученные по этим моделям данные уровня приливов интерполируются в точки измерений высотометров вдоль треков спутников (пространство-временная интерполяция), а затем вычитаются из данных измерений уровня. Таким образом пользователи получают данные

альтиметрии, которые могут использоваться в моделях динамики океана не учитывающие действие приливных сил.

2.3 Выводы

Выводы на основе анализа источников данных наблюдений:

- Дрифтеры обеспечивают данными о температуре, солености и других параметрах океана для верхнего слоя океана (до 2 км). За сутки поступает около 350 профилей на весь Мировой океан. Среднее расстояние между ними порядка 300 км. Объем данных измерений – порядка 10^4 точек в сутки для верхнего 2 км слоя океана.
- Спутниковые данные ТПО и альтиметрии доступны практически равномерно по всему Мировому океану между 66 с.ш. и 66 ю.ш. Объем данных измерений - порядка 10^5 – 10^6 точек в сутки.
- Спутниковые данные ТПО и альтиметрии вблизи полюсов носят пока экспериментальный характер, т.е. почти нет данных севернее 66° с.ш..
- Данные измерений для некоторых областей океана носят единичный характер. В частности, практически нет (или их очень мало) данных измерений для Северного Ледовитого океана и глубинных слоев Мирового океана (глубже 2 км).

Следует сказать, что появилось много так называемых GRID продуктов, когда различные нерегулярные данные океанических наблюдений интерполируются по времени и пространству на регулярную модельную сетку требуемого разрешения (1° , 0.25° и так далее). Такой подход применим при моделировании климата, но не подходит для

решения задач сезонного и оперативного прогноза, поскольку не позволяет реагировать на резкое изменение состояния океана (например, возникновение шторма или цунами). В конечном итоге такая экстраполяция относительно малочисленных данных наблюдений приводит к искусственному увеличению объема этих данных, не внося дополнительную информацию об окружающей среде.

Глава 3. Параллельный алгоритм метода оптимальной интерполяции

3.1 Введение

В главе 3 рассматривается параллельная реализация одного из простых методов усвоения данных - метода многомерной оптимальной интерполяции (*eng. multivariate optimum interpolation, MVOI*). С применением MVOI решается задача усвоения данных о температуре и солености (профили с дрейфтеров ARGO) [24] в модели динамики океана ИВМИО.

Затем рассматривается более сложный метод ансамблевой оптимальной интерполяции (*eng. Ensemble Optimal Interpolation, EnOI*) [25]. Для него разработан параллельный алгоритм, учитывающий недостатки параллельного подхода на базе MVOI, и описывается его реализация в виде программного сервиса DAS (*eng. Data Assimilation Service*) на базе платформы CMF 3.0 [84]. Разработанный алгоритм на основе EnOI рассчитан на усвоение данных с различных измерителей (дрейфтеры, спутники) с возможностью построения кросс-ковариационных матриц (между различными модельными величинами) для корректировки модельных полей тех величин, измерения для которых отсутствуют. Разработанный и реализованный метод на базе EnOI планируется использовать для всех данных наблюдений, обзор которых приведен в главе 2. Результаты численных экспериментов для разработанных параллельных методов усвоения данных приведены в главе 4.

Далее будем пользоваться следующим правилами:

- Матрицы обозначаются прописными буквами.
- Векторы обозначаются строчными буквами и выделяются **жирным**.

- Размер матрицы или вектора указывается в круглых скобках после названия.

3.2 Методы усвоения данных

Основные уравнения, используемые при усвоении динамико-стохастическими методами [66], следующие:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y}_{\text{obs}} - \mathbf{H}\mathbf{x}_b) \quad (3.1)$$

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \quad (3.2)$$

В формулах (3.1), (3.2):

\mathbf{x}_a , \mathbf{x}_b – векторы модельной величины после и до усвоения (анализа и прогноза, *eng. analysis and background*) размера n , где n - количество точек модельной сетки, умноженное на число корректируемых модельных величин (температура, соленость, уровень, скорости и т.д.), имеет порядок 10^8 для модели Мирового океана с разрешением 0.1 градуса;

\mathbf{y}_{obs} – вектор наблюдений размера m , где m - число точек наблюдения, имеющее порядок 10^4 . Может также содержать различные данные наблюдений, такие как температура, соленость, уровень и т.д.;

$\mathbf{K}(n \times m)$ – весовая матрица (*eng. Kalman gain matrix*);

$\mathbf{R}(m \times m)$ - ковариационная матрица инструментальных ошибок наблюдений, которая имеет диагональный вид, так как предполагается, что инструментальные ошибки не коррелированы;

$\mathbf{H}(m \times n)$ - матрица проектирования значений модели в пространство наблюдений;

\mathbf{B} - ковариационная матрица состояния модели. Ее строгое определение обычно не дается, вместо этого описывается метод ее расчета.

3.3 Метод оптимальной интерполяции (MVOI)

Метод оптимальной интерполяции, впервые предложенный в работе [7], широко используется в системах прогнозирования погоды [63] и океана [126]. Полный вывод и описание метода представлены в [61]. В методе ковариационная матрица B аппроксимируется функцией от координат. С применением MVOI будет решаться задача усвоения данных о температуре и солёности (профили с дрейфтеров ARGO). В работе [24] используется общая ковариационная матрица для температуры и солёности, элементы которой задаются функцией расстояния между точками. Поскольку эта корреляционная зависимость строится одновременно для обеих величин, то это позволяет назвать этот общий метод - методом многомерной оптимальной интерполяции.

Ковариация между узлами сетки с индексами i и j будет вычисляться по формуле, предложенной в работе [61]:

$$B_{ij} = \begin{cases} (1 + r_{ij}) \exp(-r_{ij}), r_{ij} \leq r_{inf} \\ 0, r_{ij} > r_{inf} \end{cases} \quad (3.3)$$

В формуле (3.3) r_{ij} – расстояние в градусах между двумя точками сетки, а r_{inf} – радиус отсечения т.е. на расстоянии больше r_{inf} ковариация между точками полагается равной нулю.

Матрица (NBH^T) , задаваемая вышеприведенной формулой, обладает свойствами:

- а) Симметричности
- б) На диагонали стоят единицы, а вне диагональные элементы практически экспоненциально убывают как функция расстояния.

Для используемых данных (данные о температуре и солёности с дрейфтеров ARGO, которые расположены достаточно редко) матрица

обладает свойством диагонального доминирования, что является достаточным условием ее положительной определенности. Вертикальная корреляция не учитывается, поскольку усваиваются профили дрейфтеров ARGO температуры и солёности на глубине до 1500 м и практически для каждого модельного горизонта доступны свои данные измерений.

3.4 Параллельный алгоритм для метода MVOI

Основная задача при реализации алгоритма MVOI заключается в вычислении матриц B и K . Элемент матрицы B_{ij} показывает ковариацию между модельными величинами в i и j узлах модельной сетки. Предположим, что ковариация между достаточно удалёнными друг от друга точками сетки пренебрежимо мала. В работе [62] дается общее обоснование такого предположения. При выборе конкретного значения радиуса отсечения (то есть предельного расстояния, на котором ковариация между характеристиками модели отлична от нуля) сделаем следующую оценку: пусть максимальная скорость в модели составляет U_m . Если усвоение делается раз в сутки, т.е. за $\sim 10^5$ секунд, то за это время взаимовлияние модельных характеристик может сказаться на расстоянии $\sim U_m \times 10^5$ метров или $U_m \times 100$ км. Таким образом, выбор радиуса отсечения будет пропорционален этому значению. В работе [24] он выбран равным $r_{inf} = 500$ км.

В модели Мирового океана ИВМИО с пространственным разрешением 0.1° размеры двумерных процессорных подобластей метода декомпозиции [17], на которые разбивается вся область, варьируются в пределах 500 — 5 000 км в зависимости от числа используемых вычислительных ядер. Следовательно, можно считать, что $B_{ij} \neq 0$ только, если точки (i, j) находятся в одной или соседних процессорных

подобластях (т.е. матрица B сильно разреженная). Это позволяет вычислять матрицу K локально для каждой подобласти при условии, что мы получим от 8-ми соседних подобластей их локальные матрицы проекции $H(m' \times n)$ и векторы проекции модельных данных в точки наблюдения $\mathbf{H}_b(m')$, где m' - количество точек наблюдения в локальной подобласти и 8-ми соседних.

Обращение матрицы $(\mathbf{H}\mathbf{B}^T + \mathbf{R})$ размера $m' \times m'$ будем выполнять с помощью процедур `potri` (вычисление обратной матрицы через факторизацию Холецкого для симметричных положительноопределенных матриц) из пакета Intel MKL. Это вычислительно дорогой, прямой метод, требующий порядка $(m')^3$ операций с плавающей точкой, но матрица $(\mathbf{H}\mathbf{B}^T + \mathbf{R})$ уже не является разреженной, при этом имеет небольшой размер (m' не более 1000) и уменьшается по мере увеличения количества MPI процессов, поэтому выбор метода не является критичным. Численные эксперименты, выполненные в работе [24], показали хорошую масштабируемость такого параллельного алгоритма усвоения данных, но выявили и ряд недостатков. В частности, неравномерность данных наблюдений приводит к падению ускорения при увеличении числа процессорных ядер, т.к. в одних процессорных подобластях данных наблюдений оказывается существенно меньше, чем в других. Другой же недостаток — это значительный объем памяти, занимаемый матрицами K и $(\mathbf{H}\mathbf{B}^T + \mathbf{R})$. Также были выявлены некоторые архитектурно-программные проблемы, связанные с необходимостью чтения данных наблюдений всеми процессорными ядрами, на которых выполняется программный код модели океана ИВМИО, и, как следствие такой концепции, должно выполняться усвоение данных. В силу описанных причин было принято

решение вынести усвоение данных в отдельный программный сервис (раздел 3.8).

3.5 Метод ансамблевой оптимальной интерполяции (EnOI)

Основная идея ансамблевых методов (EnKF и EnOI) заключается в том, что ковариационная матрица B не задается в явном виде или в виде функции, как в методе MVOI [24], а получается из ансамбля векторов состояния модели (выборки).

Пусть

$$A_b = [x_b^1 \dots x_b^{en}] - [\bar{x}_b \dots \bar{x}_b] \quad (3.4)$$

A_b — матрица размера $n \times N$, где N — количество элементов ансамбля (обычно не более 100), столбцы которой равны значениям состояния модели минус среднее по ансамблю. При этом ансамбль состояний модели строится по модельным расчетам с разными начальными условиями. В методе EnOI для ансамбля состояний выбираются модельные расчеты за разные даты, сделанные при долговременном интегрировании модели [25]. Тогда матрица ковариации модели:

$$B^{en} = \frac{1}{N-1} A_b (A_b)^T \quad (3.5)$$

Вопрос оптимального построения ансамбля элементов-состояний модели остается открытым, и ему посвящены ряд работ, например [81], [134]. В настоящий момент используется вычислительно более дешевый метод EnOI, описанный в [105], когда в качестве элементов ансамбля выступают состояния модели за предыдущие несколько лет. В работе (проект TOPAZ) [117] используется более сложный подход: модель

стартует со ста (количество элементов в ансамбле) различных начальных состояний, но считается с реальным атмосферным форсингом за текущую дату, затем на основании полученных ста прогнозов уже строится ковариационная матрица V . Легко видеть, что при таком подходе вычислительные затраты пропорциональны количеству элементов в ансамбле.

В статье [118] проводится сравнение EnKF и EnOI для модели океана MOM4 [73] разрешения 0.1° и при значительно возрастающих (пропорционально количеству элементов в ансамбле) вычислительных затратах, точность прогноза с использованием EnKF выше лишь на 9-21 %.

3.6 Математическая формулировка метода EnOI

Ансамблевый метод усвоения данных (EnKF) был предложен в работе [67], а в работе [68] приведены алгоритмы его практического применения.

Пусть n – количество точек модельной сетки, умноженное на число корректируемых модельных величин (температура, соленость, уровень, скорости и т.д.), m – число точек данных наблюдения, N – количество элементов ансамбля. Далее будем предполагать, что соотношение между n , m и N для рассматриваемой модели ИВМИО разрешения 0.1° следующее:

$$N \ll m \ll n, \quad N \sim 10^2, \quad m \sim 10^4, \quad n \sim 10^8 \quad (3.6)$$

$x_a, x_b (n)$ – векторы модельной величины после и до усвоения (анализа и прогноза, *eng. analysis and background*)

$y_{obs} (m)$ – вектор данных наблюдений;

$K (n \times m)$ – весовая матрица (*eng. Kalman gain matrix*);

$R (m \times m)$ – ковариационная матрица инструментальных ошибок данных наблюдений;

H ($m \times n$) – матрица проектирования значений модели в пространство наблюдений;

V ($m \times n$) – ковариационная матрица состояния модели;

Если обозначить за A_b ($n \times N$) матрицу ансамбля, где

$$A_b = [\mathbf{x}_b^1 \dots \mathbf{x}_b^{en}] - [\bar{\mathbf{x}}_b \dots \bar{\mathbf{x}}_b] \quad (3.7)$$

Тогда матрица ковариации модели в методе EnKF и EnOI строится как:

$$V^{en} = \frac{1}{N-1} A_b (A_b)^T \quad (3.8)$$

Введем матрицу S , в столбцах которой содержится проекция элементов ансамбля в точки данных наблюдений:

$$S = H A_b, \quad (m \times N) \quad (3.9)$$

И вектор инновации данных наблюдений:

$$\mathbf{d} = \mathbf{y}_{obs} - H \mathbf{x}_b, \quad (m) \quad (3.10)$$

Также введем матрицу E ($m \times N$) – ансамблевое представление матрицы R . E – псевдослучайная матрица, которая строится на основе требуемого уровня дисперсии и радиуса ковариации для конкретных данных наблюдений [68].

$$R = \frac{EE^T}{N-1} \quad (3.11)$$

Тогда из уравнений (3.1) и (3.2) получаем:

$$\mathbf{x}_a = \mathbf{x}_b + A_b S^T (SS^T + EE^T)^{-1} \mathbf{d} \quad (3.12)$$

Ниже для формулы (3.12) приведены размеры соответствующих матриц и векторов:

$$n = (n) + (n \times N)(N \times m)(m \times m)(m) \quad (3.13)$$

Для обращения симметричной матрицы $(SS^T + EE^T)$ можно использовать ее спектральное разложение, но для его вычисления требуется порядка m^3 операций с плавающей точкой. Количество требуемых операций с плавающей точкой для вычисления будем указывать после расчетной формулы и размера матрицы (или вектора).

$$SS^T + EE^T = Z\Lambda Z^T, (m \times m), O(m^3) \quad (3.14)$$

Такой подход неприемлем для большого количества данных наблюдений. Поэтому для больших m используется другой подход.

$$NA_b A_b^T N^T + EE^T = (NA_b + E)(NA_b + E)^T \quad (3.15)$$

В (3.15) учтено, что ансамбли модельных прогнозов и ошибок измерений не коррелируют, что эквивалентно предположению о том, что ошибки модельного прогноза не связаны с ошибками данных наблюдений:

$$NA_b E^T \equiv 0$$

Вычислим сингулярное разложение матрицы $(S+E)$ $(m \times N)$ и левую матрицу собственных векторов U . Вычислительная стоимость такой процедуры составляет $O(mN^2)$ операций для $m > N$.

$$S+E = U\Sigma V^T, m \times N, O(mN^2) \text{ – операций} \quad (3.16)$$

Из (3.15) и (3.16) получаем:

$$SS^T + EE^T = U\Sigma V^T V\Sigma^T U^T = U\Sigma\Sigma^T U^T \quad (3.17)$$

Здесь $\Sigma\Sigma^T$ будет идентично левому верхнему квадрату $(N \times N)$ матрицы $\Lambda \sim \Sigma\Sigma^T$, $(N \times N)$, что соответствует N ненулевым собственным значениям. Кроме того, N сингулярных векторов, содержащихся в U , эквивалентны первым собственным векторам в Z .

Таким образом из (3.12), вектор анализа можно получить из вектора прогноза по следующей формуле:

$$\mathbf{x}_a = \mathbf{x}_b + A_b S^T U \Lambda^{-1} U^T \mathbf{d} \quad (3.18)$$

Ниже приведен алгоритм вычисления вектора анализа в несколько шагов, с указанием вычислительной стоимости каждого шага.

$$G_1 = \Lambda^{-1} U^T, (N \times m), O(mN) \text{ – операций} \quad (3.19)$$

$$G_2 = G_1 \cdot \mathbf{d}, (N), O(mN) \text{ – операций} \quad (3.20)$$

$$G_3 = U G_2, (m), O(mN) \text{ – операций} \quad (3.21)$$

$$G_4 = S^T G_3, (N), O(mN^2) \text{ – операций} \quad (3.22)$$

$$\mathbf{x}_a = \mathbf{x}_b + A_b G_4, (n), O(nN^2) \text{ – операций} \quad (3.23)$$

Основным достоинством такого подхода является то, что его вычислительная стоимость растет линейно от размера задачи: числа узлов модельной сетки и количества данных наблюдений (параметров n и m , соответственно):

$$O((m+n)N^2) \text{ – операций} \quad (3.24)$$

3.7 Параллельный алгоритм метода EnOI

В рамках изложенного в **разделе 3.5** метода EnOI был предложен следующий параллельный алгоритм усвоения данных наблюдений для параллельных компьютеров с распределенной памятью. Как упоминалось в **разделе 1.2** и **приложении 1** модель океана ИВМИО имеет параллельную реализацию для компьютеров с распределенной памятью на основе принципа декомпозиции области (**рисунок 3.1a**), где каждую подобласть обчисляет отдельное процессорное ядро. Но такой подход оказался не совсем подходящий для усвоения данных (см. **раздел 3.5**), и было принято решение при усвоении данных наблюдений использовать другую процессорную декомпозицию области и отдельные от модели океана процессорные ядра.

В описании предлагаемого параллельного метода цветом будет обозначаться максимальный размер объекта (матрицы или вектора).

$$n \sim 10^8; \quad m \sim 10^4; \quad N \sim 10^2 \quad (3.25)$$

В параллельной реализации будут использоваться три декомпозиции области для компьютеров с распределенной памятью. Двумерная процессорная декомпозиция области, используемая для модели океана (**рисунок 3.1a**, обозначается верхним левым индексом $^{2d}_x$, данные хранятся на вычислительных ядрах океана). Две одномерные (обозначается – $^{1d}_x$) процессорные декомпозиции: равномерная по точкам модельной сетки (показана черным пунктиром на **рисунке 3.1c**) или равномерные по данным наблюдений (показана красным пунктиром на **рисунке 3.1c**). Данные хранятся на вычислительных ядрах, занимающихся усвоением. В рамках комплекса CMF 3.0 блоку усвоения данных дано название DAS (Data Assimilation Service).

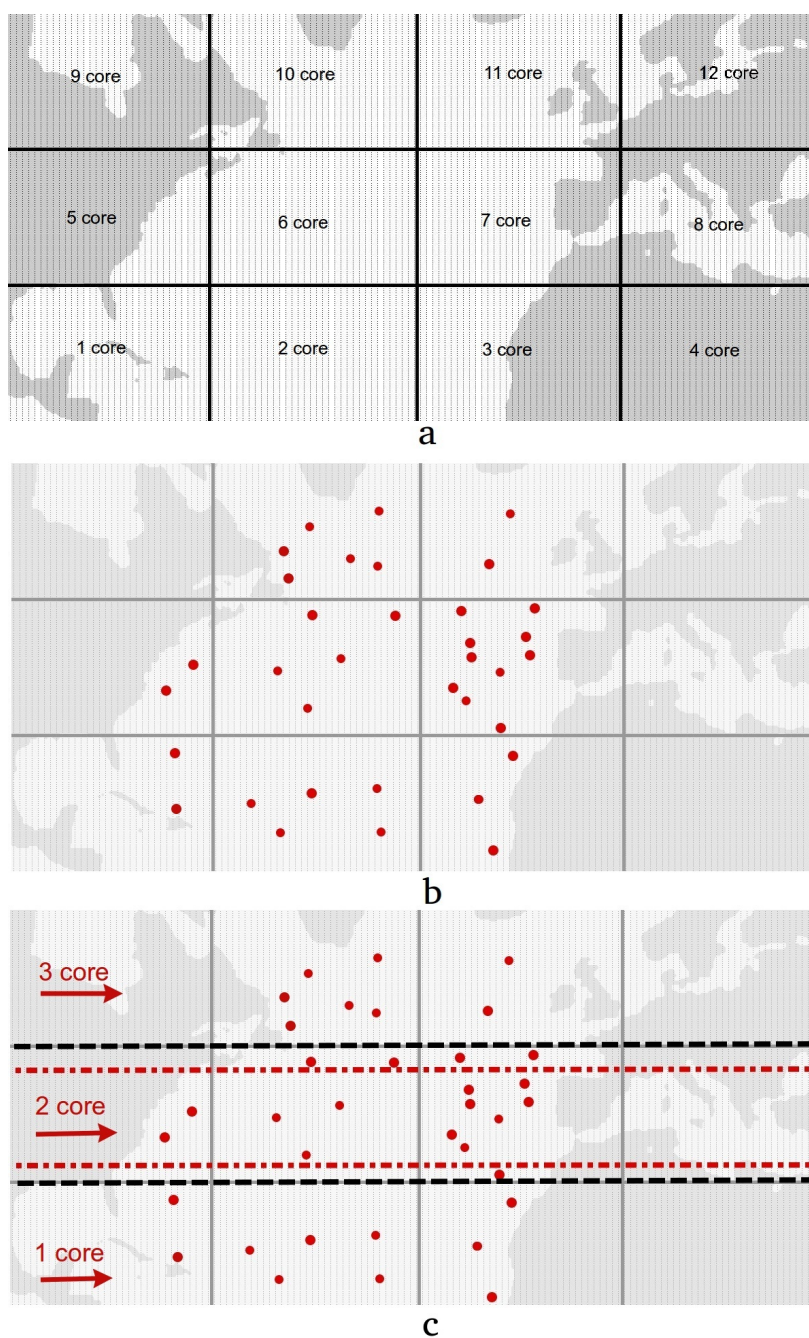


Рисунок 3.1. (а) Двумерная процессорная декомпозиция области, используемая в модели Мирового океана ИВМИО (показана серыми прямоугольниками); (б) красными точками показаны данные наблюдения; (с) одномерная процессорная декомпозиция, используемая в сервисе DAS, показана красным и черным пунктиром.

Пусть изначально у нас есть вектор модельного решения и вектор данных наблюдений:

$${}^2d\mathbf{x}_b \quad y_o \quad (3.26)$$

Вектор модельного решения пересылается в DAS.

$${}^2d\mathbf{x}_b \rightarrow {}^1d\mathbf{x}_b \quad (3.27)$$

Вычисляется инновация данных наблюдений.

$${}^1d\mathbf{y}_o \rightarrow {}^1d\mathbf{d} = {}^1d\mathbf{y}_o - \mathbf{H}\mathbf{x}_b \quad (3.28)$$

Вычисляется матрица \mathbf{S} по формуле (3.9).

$${}^1d\mathbf{A}_b \rightarrow {}^1d\mathbf{S} \quad (3.29)$$

Далее вычислим сингулярное разложение матрицы $(\mathbf{S}+\mathbf{E})$ и левую матрицу собственных векторов \mathbf{U} . Благодаря тому, что данные наблюдений (и, как следствие, блоки матрицы \mathbf{S} и \mathbf{E}) распределены равномерно по процессорным ядрам, нагрузка на эти ядра распределяется также равномерно.

$${}^1d(\mathbf{E}+\mathbf{S}) = {}^1d\mathbf{U}\Sigma\mathbf{V}^T \quad (3.30)$$

$$\Sigma\Sigma^T \rightarrow \Lambda^{-1} \quad (3.31)$$

Далее ядрами сервиса DAS параллельно выполняются операции, описанные формулами (3.19) – (3.23).

$${}^1d\mathbf{G}_1 = \Lambda^{-1} {}^1d\mathbf{U}^T \quad (3.32)$$

$$\mathbf{G}_2 = {}^1d\mathbf{G}_1 {}^1d\mathbf{d} \quad (3.33)$$

$${}^1d\mathbf{G}_3 = {}^1d\mathbf{U}\mathbf{G}_2 \quad (3.34)$$

$$\mathbf{G}_4 = {}^1d\mathbf{S}^T {}^1d\mathbf{G}_3 \quad (3.35)$$

Отметим, что лишь в следующей формуле для оптимизации нагрузки на вычислительные ядра требуется использовать равномерную по точкам сетки декомпозицию.

$${}^{1d}x_a = {}^{1d}x_b + {}^{1d}A G_4 \quad (3.36)$$

Вектор x_a отправляется в модель океана и раскладывается на ее двумерную процессорную декомпозицию:

$${}^{1d}x_a \rightarrow {}^{2d}x_a \quad (3.37)$$

3.8 Программная реализация.

Как всякий сервис программного комплекса CMF 3.0 усвоение данных выполняется на отдельных вычислительных ядрах [30]. Это позволяет лучше структурировать систему совместного моделирования, где каждая программная компонента решает свою задачу. При этом модель океана в усвоении практически не участвует, а используются лишь модельные результаты расчётов в виде векторов-элементов ансамбля, на основе которых аппроксимируется ковариационная матрица модели В. Данные от модели океана поступают в сервис (обычно один раз в модельные сутки) без обращения к файловой системе, что важно, поскольку размер трехмерных массивов состояния для модели океана с разрешением 0.1° составляет несколько гигабайт.

Проблемы и причины вынесения процедуры усвоения данных в отдельный сервис следующие:

1. Данные наблюдений (спутниковые или данные дрейфтеров) практически всегда распределены крайне неравномерно по расчётной области океана, поэтому если для их обсчета использовать ту же

двумерную процессорную декомпозицию океана (**Рисунок 3.1а**), то нагрузка на вычислительные ядра будет также распределена неравномерно.

2. Ковариационные матрицы занимают значительный объем памяти, и их лучше хранить на отдельных от модельных компонент вычислительных ядрах.

Алгоритм усвоения EnOI в рамках DAS

1. На основе поступивших данных наблюдений y_b (**Рисунок 3.1b**) строится новая одномерная процессорная декомпозиция расчётной области, равномерная по данным наблюдения и имеющая вид горизонтальных полосок (**Рисунок 3.1с**).
2. Каждая «полоска» получает только свои данные наблюдения y_{obs} и только свою часть глобального модельного массива x_b и строится проекция модельного решения в точки наблюдения Hx_b .
3. Вычисляется вектор инновации наблюдений $(y_{obs}-Hx_b)$, также локально для каждой «полоски».
4. Асинхронно читается ансамбль состояний A_b , распределяется по ядрам, строится проекция в точки наблюдения Na_b .
5. По формулам (3.30) – (3.36) вычисляется вектор анализа x_a , который затем отправляется в модель океана и используется как начальное условие для дальнейшего интегрирования модели.

Замечания

- Матрица H (используется метод билинейной интерполяции) строится каждый раз при вызове сервиса, поскольку данные наблюдений каждый раз доступны в новых точках. Как следствие каждый раз строится новая процессорная декомпозиция, равномерная по точкам наблюдений и неравномерная по точкам сетки.
- Для реализации алгоритма, описанного выше, используются вызовы

функции из пакетов BLAS и LAPACK (используя Intel MKL) через API библиотеки Global Arrays, что существенно упростило написание межпроцессорных обменов.

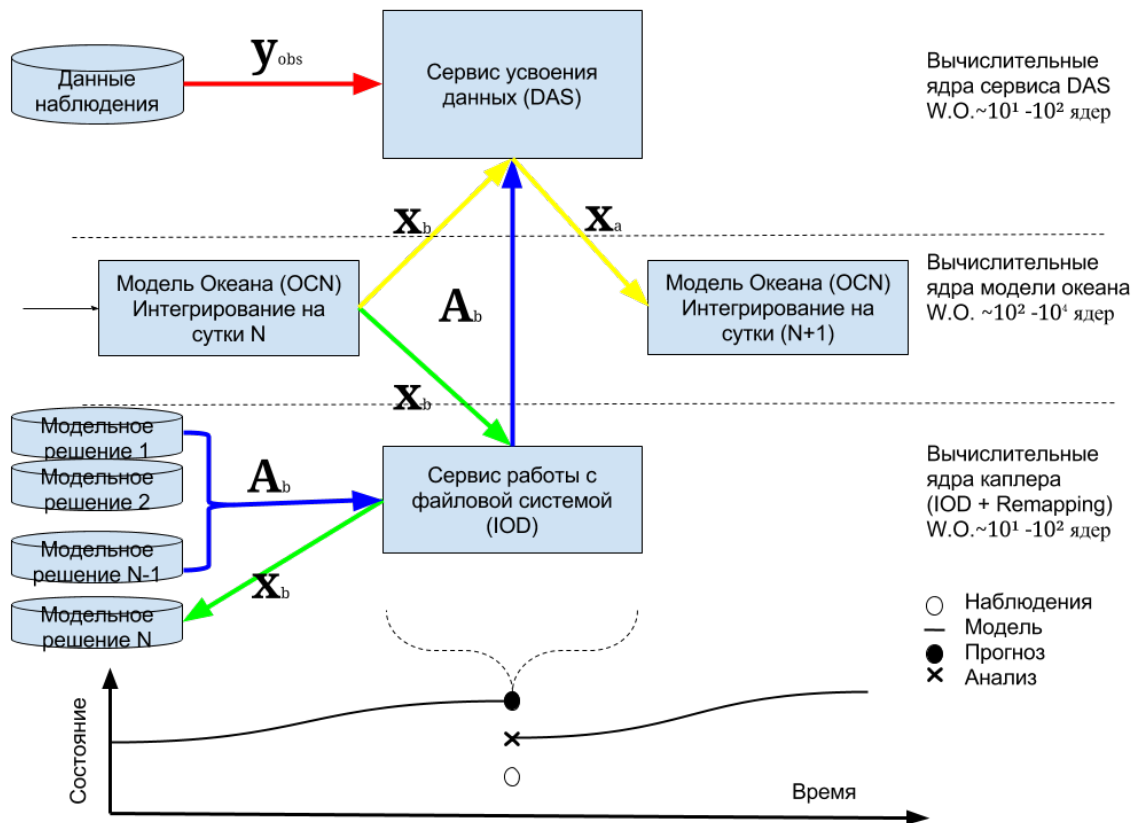


Рисунок 3.2. Принципиальная схема передачи данных между моделью океана (OCN), сервисом усвоения данных (DAS), сервисом работы с файловой системой (IOD) в рамках CMF 3.0. Снизу показана временная шкала изменения вектора состояния модели при интегрировании модели и усвоении данных.

Принципиальная схема передачи данных между моделью океана (OCN), сервисом усвоения данных (DAS), сервисом работы с файловой системой (IOD) в рамках программного комплекса совместного

моделирования CMF 3.0 показана на **рисунке 3.2**. Также нужно отметить необходимость сохранять вектор модельного решения на файловой системе для построения матрицы ансамбля в дальнейшем. На рисунке также приведены оценки требуемого количества вычислительных ядер для модели Мирового океана (W.O.).

3.9 Выводы

В **главе 3** описаны разработанные параллельные методы усвоения данных наблюдений на базе MVOI и EnOI, а также особенности их программной реализации.

Научные результаты:

1. Разработан параллельный алгоритм усвоения данных наблюдений MVOI в модели динамики океана ИВМИО
2. Создана программная архитектура системы усвоения данных наблюдений
3. Разработан параллельный алгоритм на основе метода EnOI, равномерно масштабируемый для нерегулярных данных наблюдений
4. Написана программная реализация блока усвоения данных в качестве программного сервиса DAS вычислительной платформы CMF 3.0.

На описанные программные продукты оформлены свидетельства о Государственной регистрации программ для ЭВМ [26], [27].

Глава 4. Численные эксперименты по усвоению данных в модели Северной Атлантики

4.1 Введение

Атлантический океан является одним из наиболее изученных и богатых регионов Мирового океана на различные системы наблюдения и прогноза (в частности системы TOPAZ, REMO, C-NOOFS созданы и функционируют именно для этого региона). Качественное воспроизведение Гольфстрима очень важно для тех океанических моделей, в рамках которых требуется провести анализ и прогноз климатического состояния Земли, так как он является составляющей частью глобального океанского конвейера. Течение Гольфстрим играет ключевую роль в балансе тепла в Атлантике и, кроме того, в отличие от многих других течений, обладает рядом устойчивых характеристик, и поэтому они могут быть ориентирами при проверке океанических моделей в целом.

Успешное воспроизведение Гольфстрима, вместе с его вихревыми структурами и последующего Североатлантического течения, в океанических моделях является результатом многих факторов, и усвоение данных наблюдения может помочь справиться с этой задачей, которая является краеугольным камнем для моделей динамики океана высокого пространственно-временного разрешения.

Итак, регион Северной Атлантики выбран для тестирования созданной системы усвоения данных по следующим причинам:

- Большое количество данных наблюдений, как спутниковых данных ТПО и альтиметрии, так и несколько систем плавучих буев.
- Относительно большой размер исследуемого региона (1/5 по количеству узлов модельной сетки от модели Мирового океана).

- Много научных коллективов занимаются моделированием динамики океана в Североатлантическом регионе в оперативном режиме с усвоением всех доступных данных наблюдения, что позволяет объективно оценить качество выдаваемого модельного прогноза.
- В Атлантике присутствуют как области открытого океана вдали от интенсивных течений, так и динамически активный район в области течения Гольфстрим, что позволит всесторонне оценить эффективность созданной системы моделирования.

4.2 Постановка численных экспериментов

Для тестирования параллельного алгоритма усвоения EnOI была взята область Северной Атлантики с границами от 33° ю.ш. до 67° с.ш, на которых задано условие твердой стенки⁶. Перенос через Гибралтарский пролив не учитывался.

Порядок проведения численных экспериментов. Модель Северной Атлантики интегрировалась за период с 01.01.1948 по 01.05.2008 с нулевых начальных скоростей и климатических распределений температуры и солёности, взятых из архива WOA 2009 [88]. При этом использовался атмосферный форсинг CORE-IAFv2 [85] за соответствующий период времени.

После разгона (в английской литературе Spin Up) модели продолжительностью 5 лет, в процессе ее интегрирования каждые 10 дней сохранялись контрольные точки — полные состояния модели, которые использовались в качестве элементов ансамбля. В дальнейшем для

6) Такие граничные условия не являются физически верными, но их аккуратная постановка не является целью данной работы.

построения ансамбля использовались поля за тот же календарный месяц, где выполняется усвоение, но за предыдущие годы.

Были проведены четыре основных численных эксперимента с атмосферным воздействием, соответствующим периоду с 01.05.2008 по 29.06.2008 (60 дней). Ниже проводится анализ и сравнение результатов этих экспериментов. Сводная информация по этим экспериментам приведена в **таблице 4.1**.

Эксперимент A01 — контрольный эксперимент. Усвоение данных не выполняется.

Эксперимент A02 — эксперимент с усвоением данных дрейфтеров ARGO методом MVOI.

Эксперимент A03 — эксперимент с усвоением данных дрейфтеров ARGO методом EnOI.

Эксперимент A04 — эксперимент с усвоением данных альтиметрии AVISO методом EnOI.

Во всех экспериментах модельное решение ежедневно сравнивается со спутниковыми данными альтиметрии AVISO и абсолютно независимыми от них данными дрейфтеров ARGO о температуре и солености.

В экспериментах A02, A03 в качестве x_a , x_b (см. **раздел 3.2**) используются векторы модельного решения, составленные из трехмерных полей температуры и солености; y_{obs} - вектор данных наблюдений состоит из данных дрейфтеров ARGO поступивших за модельные сутки (порядка 10^2 профилей в сутки). Также в экспериментах A02, A03 в матрице B не будет учитываться вертикальная корреляция, так как данные наблюдений

ARGO доступны для всех горизонтов на глубине до 1500 метров. В методе MVOI матрица В будет вычисляться по аналитической формуле (3.3).

Таблица 4.1. Описание экспериментов по усвоению данных наблюдений ARGO и AVISO в северной Атлантике.

Эксперимент	Усвоение T и S (данные ARGO)	Усвоение уровня (данные AVISO)	Сравнение с ARGO	Сравнение с AVISO
A01	-	-	+	+
A02	MVOI	-	+	-
A03	EnOI	-	+	-
A04	-	EnOI	+	+

В эксперименте A04 в качестве x_a , x_b используются векторы модельного решения, составленные из трехмерных полей температуры и солености и двумерного поля уровня океана (размер вектора $1.2 \cdot 10^8$). y_{obs} - вектор данных наблюдений, состоящий из данных аномалии уровня (*eng. sea level anomaly, SLA*) AVISO со спутника Jason-1 (размер переменный, порядка 10^3). Из спутниковых и модельных данных уровня океана перед усвоением вычитается их среднее, чтобы модельные и спутниковые данные имели одинаковое (равное нулю) среднее для их корректного усвоения. После усвоения данных модельное среднее прибавлялось обратно к аномалии уровня океана. В вычисляемой таким образом ковариационной матрице В будет учитываться корреляция (связь) между различными модельными величинами: аномалией уровня океана, температурой и соленостью на различных модельных горизонтах. Таким

образом, усвоение одной из величин (в данном случае уровня океана) будет корректировать весь вектор модельного решения.

Как отмечалось в разделе 2.2.3 в используемых данных альтиметрии уже учтено влияние приливов и эти данные могут использоваться в моделях океана, не учитывающих влияние приливообразующих сил. В модели [3] влияние приливообразующих сил учитывается в самой модели динамики океана.

В численных экспериментах A03 и A04 - число элементов ансамбля - выбрано равным пятидесяти (50). В разделе 4.8 дано экспериментальное обоснование такого выбора.

4.3 Проверка метода на синтетических данных

Для проверки корректности работы реализованного метода проведем эксперимент по усвоению синтетических данных наблюдений, полученных на основе известного состояния океана, которое будем считать «истинным». На основе контрольного эксперимента A01 продолжительностью 20 дней (с 2008-05-01 по 2008-05-21) были получены малочисленные ежесуточные профили температуры и солёности, распределённые по области как данные ARGO (порядка 50–90 профилей в сутки). Обозначим эти данные как Y_{SOD} (synthetic observation data). Данные Y_{SOD} получались билинейной интерполяцией модельного решения в точки расположения дрейфтеров ARGO.

Затем проведены два эксперимента - T01 и T02, в каждом из которых используется начальное условие, соответствующее модельному решению за 1 мая другого модельного года (2000-05-01), но атмосферный форсинг используется за май 2008 г., таким образом имитируется возмущение начального модельного поля. В эксперименте T02 будем ежедневно

усваивать малочисленные модельные профили Y_{SO_2} . Таким образом, модельное поле эксперимента A01 будет выступать в качестве «истинного» состояния системы, а усвоение малочисленных данных в эксперименте T02 должно приблизить модельное поле к A01. Более наглядно описание экспериментов приведено в **таблице 4.2**.

Таблица 4.2 Описание экспериментов с использованием синтетических данных наблюдений Y_{SO_2} .

Эксп.	Начальное условие	Использование данных наблюдений Y_{SO_2}	Атмосферный форсинг
A01	2008-05-01	Ежесуточное сохранение данных Y_{SO_2}	COREII
T01	2000-05-01	Не используются	2008-05-01
T02	2000-05-01	Ежесуточное усвоение данных Y_{SO_2} методом EnOI	- 2008-05-21

На **рисунке 4.1** приведены графики среднеквадратичной и средней ошибок модельного поля ТПО в экспериментах T01 и T02 по отношению к «истинным» данным A01, посчитанных по всем расчетным узлам модельной сетки. Из графика видно, что усвоение данных Y_{SO_2} методом EnOI значительно приближает модельное поле в эксперименте T02 к модельному полю A01.

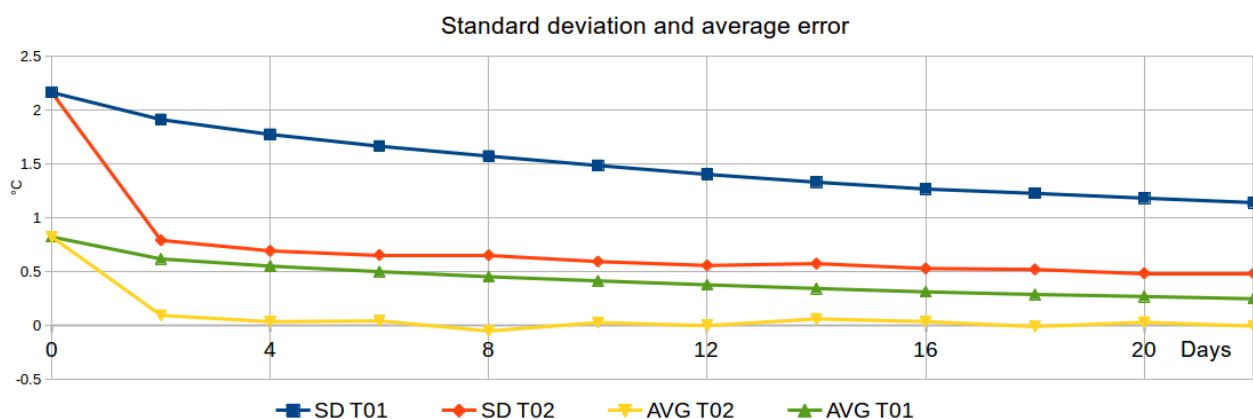


Рисунок 4.1. Среднеквадратичная и средняя ошибки модельного поля ТПО в экспериментах T01 и T02 по отношению к «истинным» данным A01, посчитанные по всем расчетным узлам модельной сетки.

Также для визуализации эффективности усвоения малочисленных данных наблюдений на **рисунке 4.2** приведены разницы между модельными полями ТПО в экспериментах T01 и T02 и контрольным A01 за 2008-05-21 модельное число. При сравнении **рисунков 4.2a** и **4.2b** наглядно видно, как усвоение исправляет решение во всей расчетной области, а не только в точках, где находятся данные Y_{SOD} . Это подтверждает корректную работу реализованного метода усвоения данных, использование которого на малочисленных данных наблюдений способно приблизить модельное решение к неизвестному «истинному» состоянию океана.

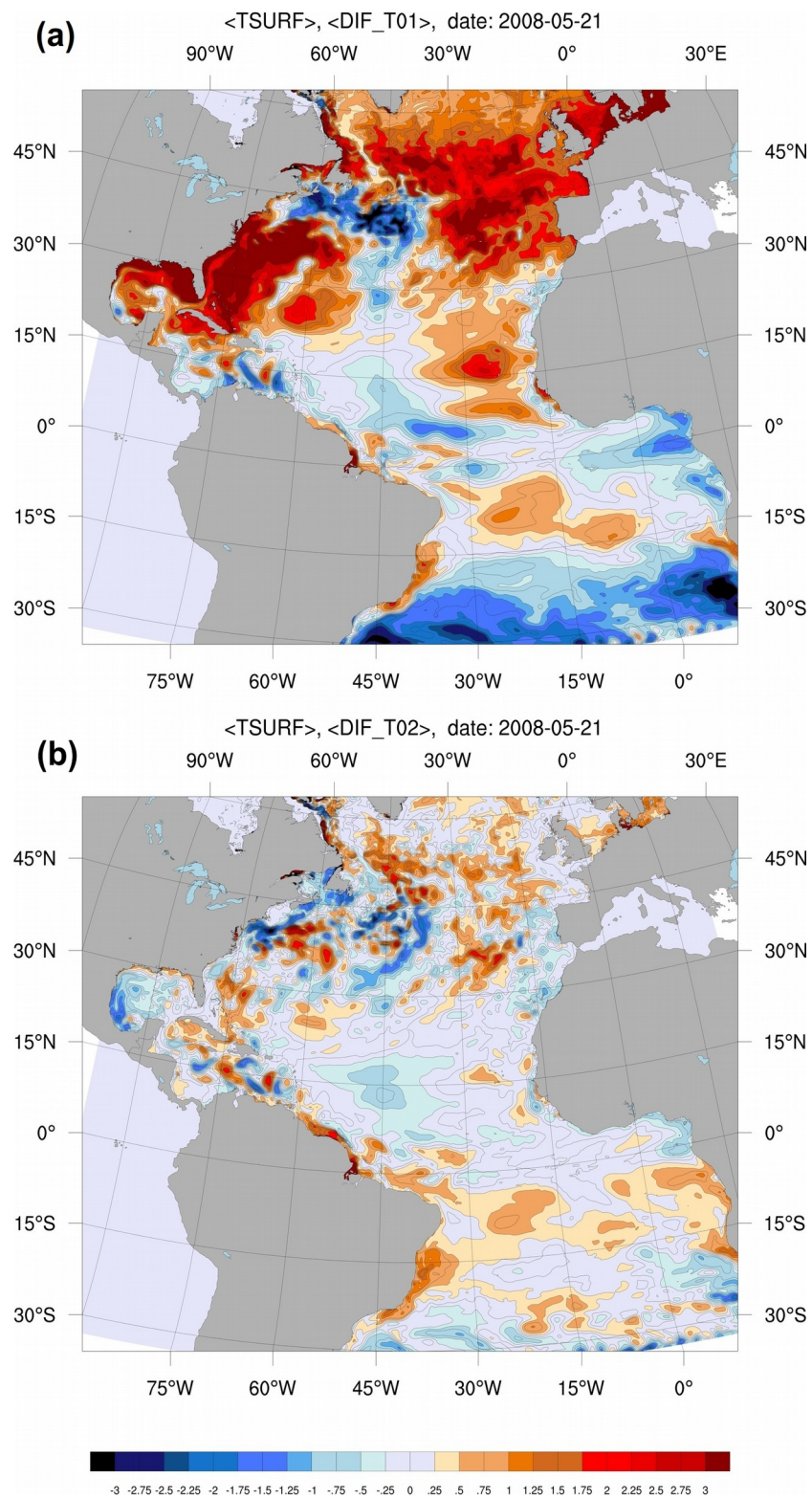


Рисунок 4.2. Разница между модельными полями ТПО за 2008-05-21

(a) (ТПО_{A01} – ТПО_{T01}) **(b)** (ТПО_{A01} – ТПО_{T02})

4.4 Качественный анализ модельного решения

Результаты экспериментов представлены на **рисунке 4.3**. Показаны температура поверхности океана (ТПО) в базовом эксперименте (A01) и эксперименте с усвоением методом EnOI (A03). Кругами показано расположение дрейфтеров ARGO, профили температуры и солёности с которых поступили на 2008-06-29. Размер кругов пропорционален разнице между температурой с дрейфтеров и модельной температурой. Крестик в круге означает, что модельная температура ниже температуры с дрейфтера, точка в круге - модельная температура выше. Обратим внимание, что усвоение проводилось в районе от 10° ю.ш. до 55° с.ш., вдали от границ модельной области. Поэтому наиболее заметные различия в результатах двух экспериментов наблюдаются именно в этом районе. В эксперименте A01 хорошо воспроизводится струйный и фронтальный характер течения Гольфстрим, его меандрирование, которое прослеживается севернее мыса Гаттерас. В решении с усвоением данных не нарушается характер течения Гольфстрим, а именно струйный (в южной части) и фронтальный (в северной). В то же время заметно усилились пространственная неоднородность и меандрирование течения.

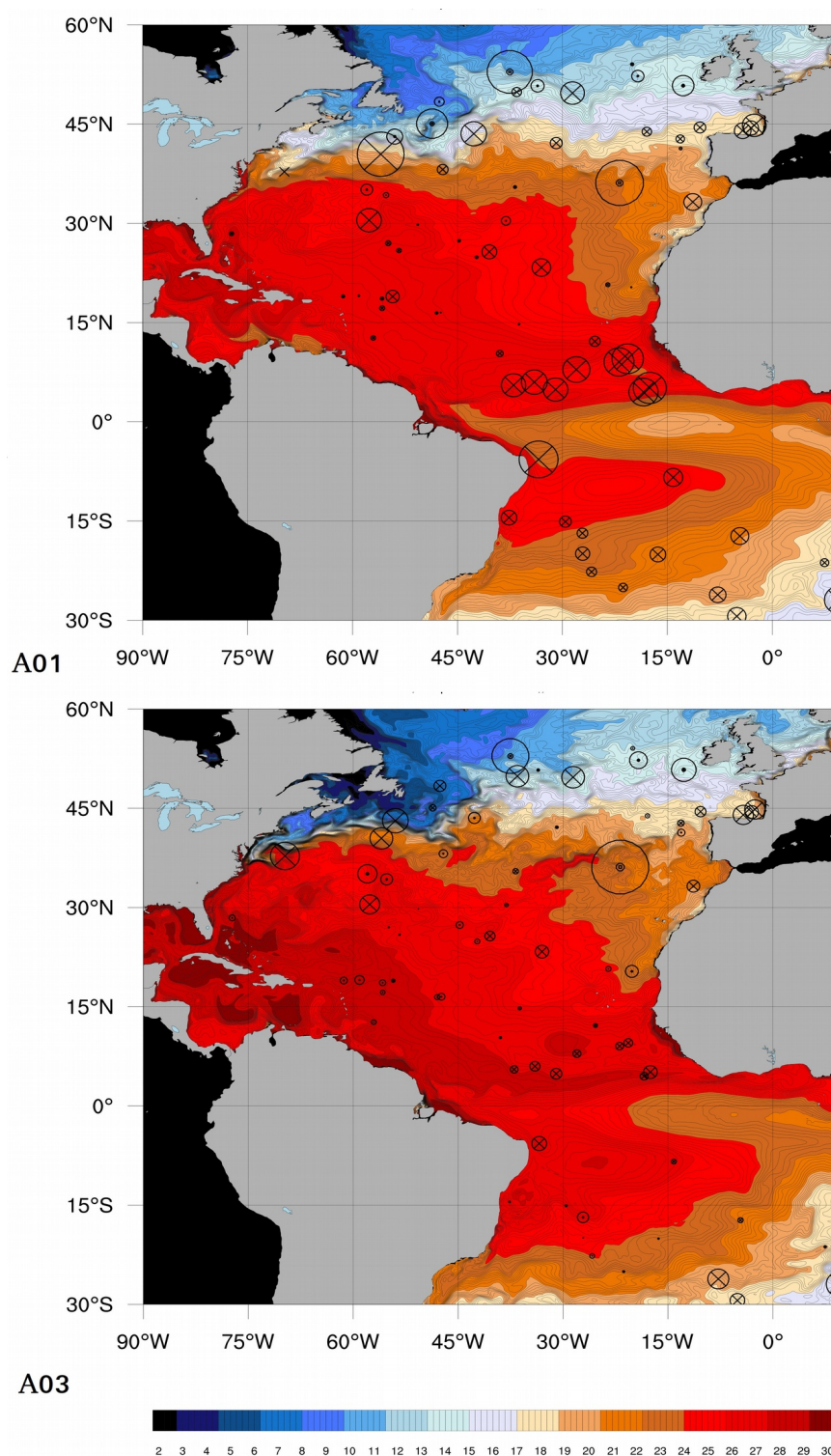


Рисунок 4.3 Температура поверхности океана в °С в модели Северной Атлантики за 2008-06-29 в базовом эксперименте (A01) и эксперименте с усвоением методом EnOI (A03).

Значительные различия в результатах экспериментов наблюдаются в открытом океане. В эксперименте A03 распределение температуры достаточно регулярное, без пространственных неоднородностей, в полосе 15-45° с.ш. наблюдается большое количество вихрей различного пространственного масштаба. При этом анализ эволюции вихрей показывает, что однажды возникнув как следствие усвоения данных, они продолжают существовать в модельном решении, заметно не диссипируя. При повторном «попадании» усваиваемых данных на ранее сгенерированный вихрь, его характеристики (температура, соленость, скорости течений) меняются незначительно.

Отсутствие вихрей открытого океана показывает, что в модели отсутствует механизм поддержки существования вихрей на протяжении длительного (более 2-3 месяцев) времени. Очевидно, что при этом механизм генерации вихрей посредством меандрирования и отрыва вихрей от течения Гольфстрим в модели воспроизводится.

4.5 Анализ результатов усвоения данных с дрейфтеров ARGO

Для количественной оценки эффекта от усвоения данных с дрейфтеров ARGO использована функция ошибки – среднеквадратичное отклонение модельного решения от данных наблюдения, посчитанное по всем дрейфтерам за конкретный день.

На **рисунке 4.4** показано изменение этой функции с течением модельного времени для различной глубины (3, 105, 310 метров). Ошибка прогноза $cost(\mathbf{H}\mathbf{x}_b - \mathbf{y}_o)$ считается каждый день в момент до усвоения данных методом EnOI (A03) или методом MVOI (A02). Ошибка контроля $cost(\mathbf{H}\mathbf{x}_b - \mathbf{y}_o)$ без усвоения считается каждый день относительно данных наблюдения, но их усвоение не происходит (A01).

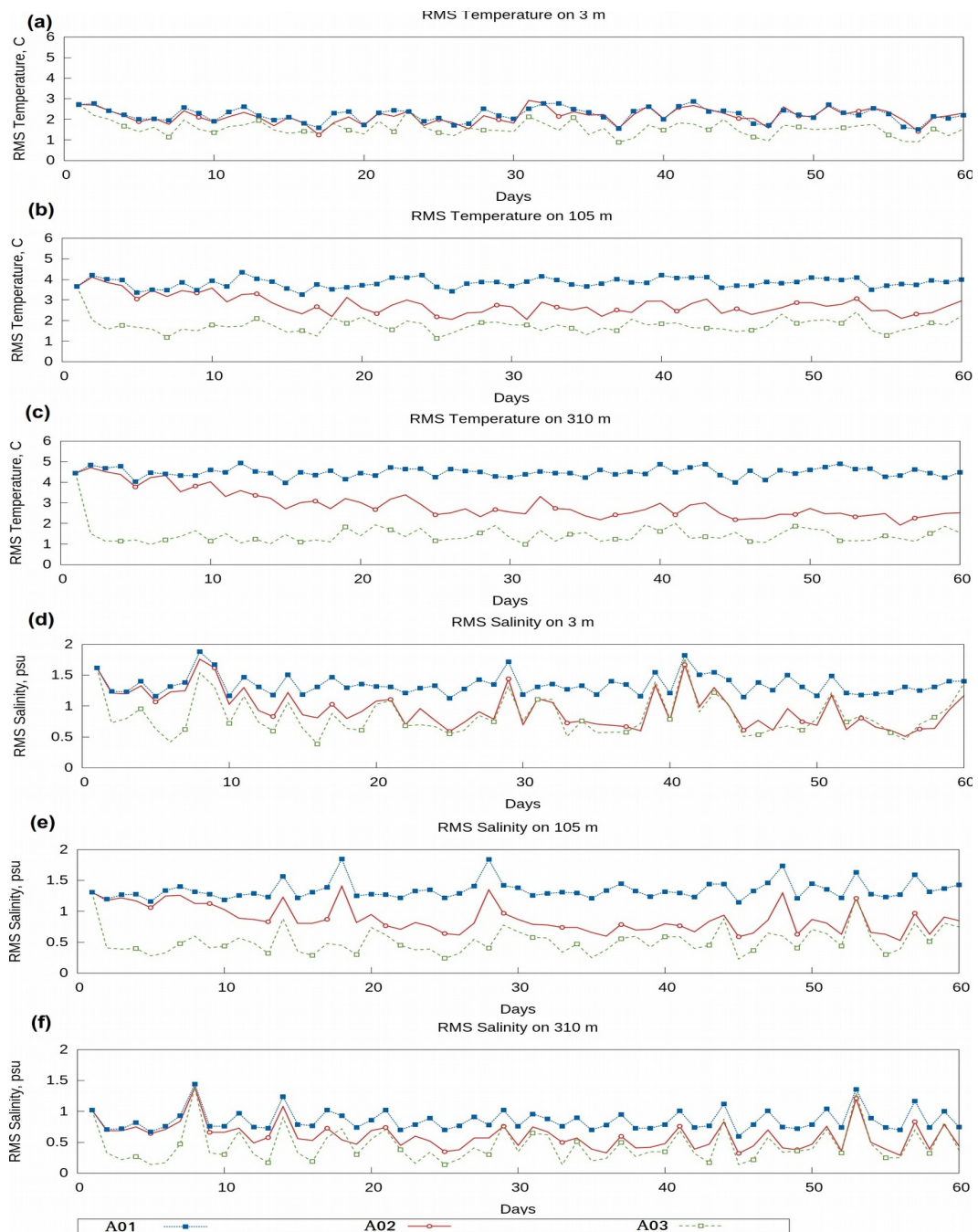


Рисунок 4.4. Среднеквадратичная ошибка контроля для контрольного эксперимента (A01), ошибка прогноза для метода MVOI (A02) и ошибки прогноза для метода EnOI (A03) для поля температуры в °C на глубине 3 (a), 105 (b) и 310 (c) метров и для поля солёности в psu на глубине 3 (d), 105 (e) и 310 (f) метров. По оси X — дни, по оси Y — ошибка температуры в °C или ошибка солёности в psu.

При сравнении графиков ошибки контроля для базового эксперимента (A01) и ошибки прогноза для экспериментов с усвоением (A02, A03) можно сделать вывод, что усвоение дает правильный знак коррекции на всех глубинах. Разница между ошибкой прогноза и контроля увеличивается с глубиной, при этом эффективность метода EnOI выше, чем у MVOI. Следует особо отметить, что близко к поверхности океана, на горизонте 5 метров, усвоение методом MVOI не эффективно из-за быстро меняющейся со временем температуры поверхности океана (ТПО) и сильного влияния атмосферного форсинга.

К концу эксперимента A02 (метод EnOI) среднеквадратичная ошибка прогноза температуры становится порядка 1.5°C , в то время как в эксперименте A03 (метод MVOI) она составляет порядка 2.5°C . Усвоение солёности также приносит пользу на всех горизонтах.

4.6 Анализ результатов усвоения данных аномалии уровня AVISO

Модельное поле уровня для Северной Атлантики с усвоением методом EnOI (A04) за 2008-06-29 и без усвоения данных (A01) показана на **рисунке 4.5**. Из анализа размера окружностей, который пропорционален разнице между модельным уровнем поверхности океана и данными измерений, можно сделать выводы об эффективности усвоения. Хорошо заметна большая разница этих окружностей в районе Северо-Атлантического течения, где динамика океана особенно выражена.

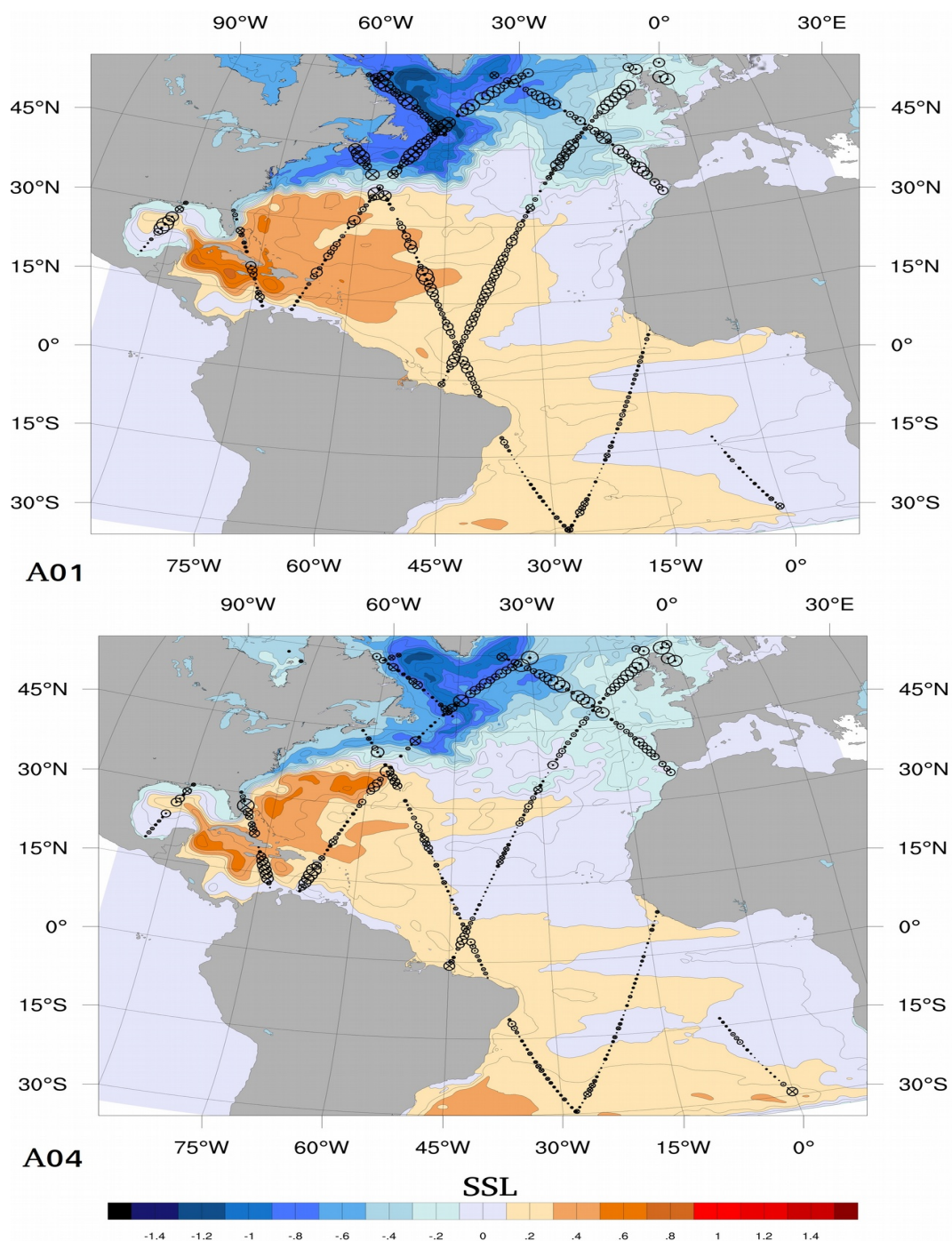


Рисунок 4.5. Функция уровня в метрах в модели Северной Атлантики за 2008-06-29 в эксперименте с усвоением уровня методом EnOI (A04) и базовом эксперименте (A01). Кругами показаны точки, для которых поступили данные уровня на 2008-06-29 со спутника Jason-1.

На **рисунке 4.6** показано изменение ошибки уровня океана относительно данных AVISO в зависимости от модельного времени. Ошибка прогноза $\text{cost}(\text{Hxb-yo})$ считается каждый день в момент до усвоения данных AVISO методом EnOI (A04), а ошибка контроля $\text{cost}(\text{Hxb-yo})$ получена без усвоения данных (эксперимент A01). На графике видно, что через 60 дней усвоения ошибка в уровне океана уменьшается примерно на 25 % (с 0.16 до 0.12 метров).

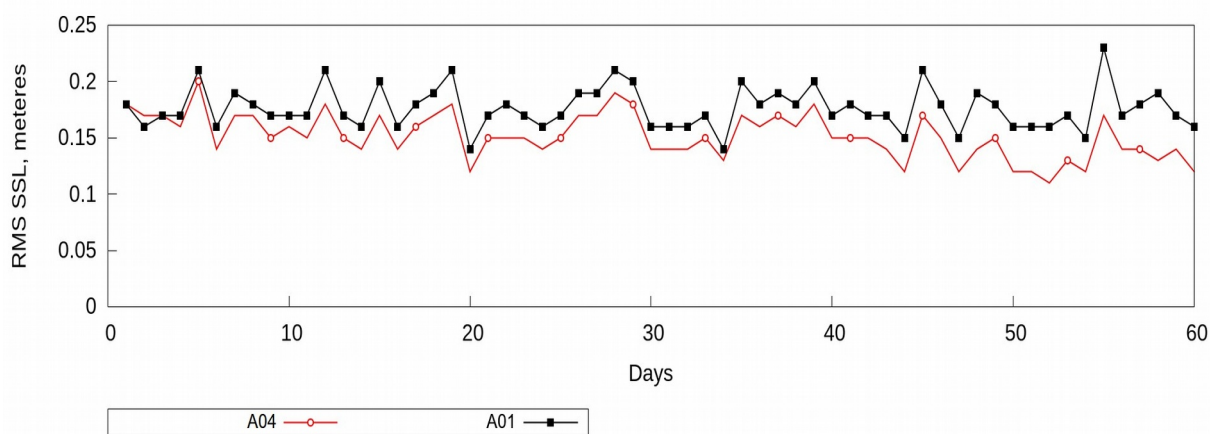


Рисунок 4.6. Среднеквадратичная ошибка контроля для эксперимента без усвоения (A01) и ошибка прогноза при усвоении данных AVISO методом EnOI (A04) для функции уровня. По оси X — дни, по оси Y — ошибка уровня в метрах.

На **рисунке 4.7** показано изменение функции ошибки прогноза (A04) в момент до усвоения данных альтиметрии и ошибки контроля (A01) для температуры и солёности по сравнению с независимыми данными ARGO по всем профилям, поступившим за текущие сутки для расчетной области. Ошибка считается каждый день для различной глубины (3, 105, 310 метров). Усвоение данных ARGO не производится.

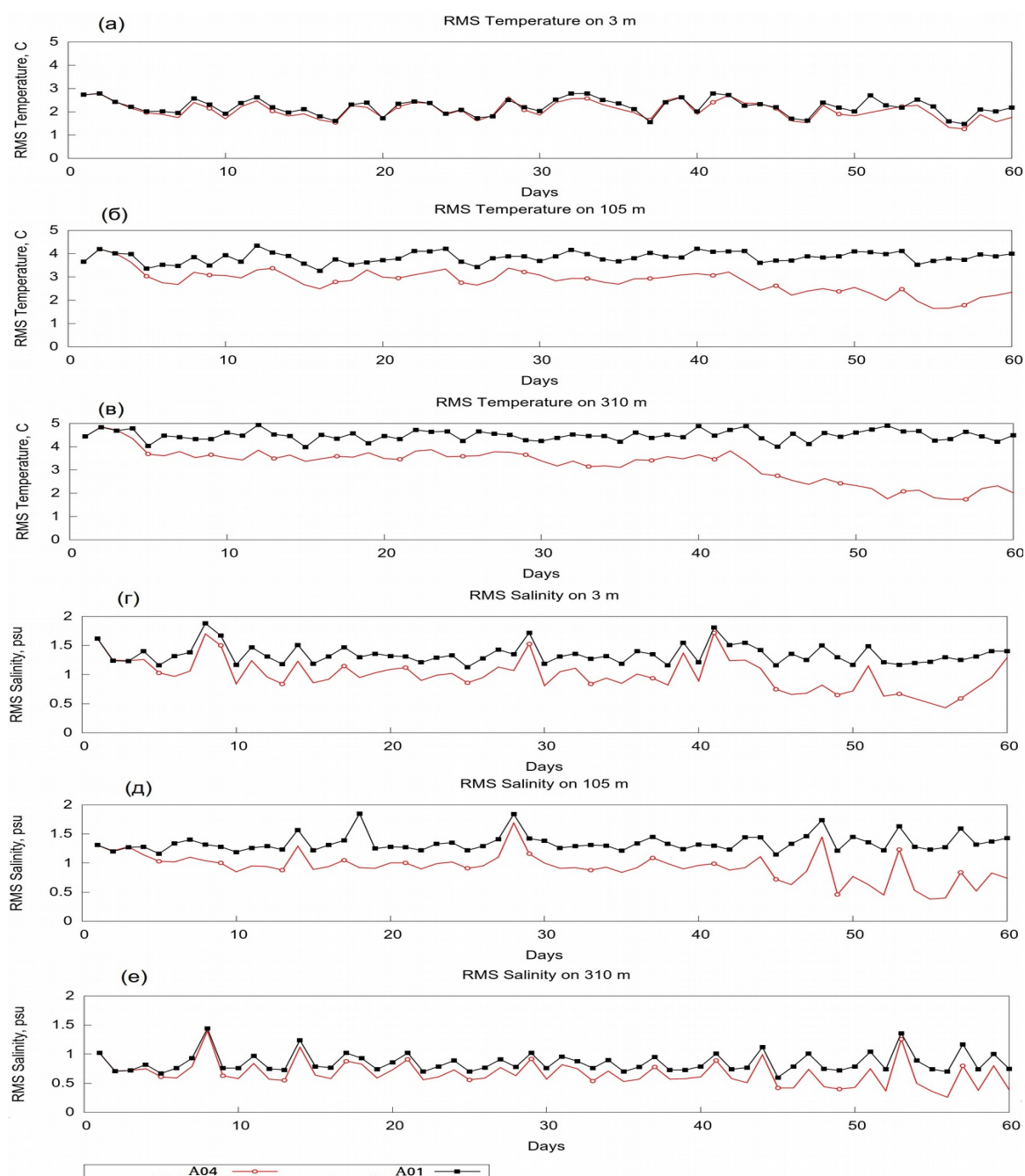


Рисунок 4.7. Среднеквадратичная ошибка контроля для эксперимента без усвоения (A01) и ошибка прогноза при усвоении данных альтиметрии методом EnOI (A04) для поля температуры в °C на глубине 3 (а), 105 (б) и 310 (в) метров и для поля солености в psu на глубине 3 (г), 105 (д) и 310 (е) метров при сравнении с данными дрейферов ARGO. По оси X — дни, по оси Y — ошибка температуры в °C или ошибка солености в psu.

При сравнении графиков ошибки контрольного решения (A01) и ошибки прогноза для экспериментов с усвоением уровня океана (A04) можно сделать вывод, что усвоение дает правильный знак коррекции на всех глубинах для температуры, солёности и для уровня океана, несмотря на то, что используются только данные альтиметрии. Разница между ошибками в прогнозе и в контрольном эксперименте для температуры увеличивается с глубиной: от 0.5°C на 3 метрах до 3°C на 310 метрах. Это связано с тем, что близко к поверхности океана (на модельном горизонте 3 метра) усвоение дает не столь значительный эффект из-за сильного влияния атмосферного форсинга. Как известно, поле плотности в океане зависит от поля солёности сильнее чем от поля температуры (это выражается в уравнении состояния), а от поля плотности напрямую зависит модельная функция уровня. Таким образом, усвоение данных аномалии уровня сильнее корректируют модельное поле солёности (нежели поле температуры), уменьшает ошибку приблизительно на 0.5 psu на глубине от 3 до 1500 метров.

На **рисунке 4.8** приведены профили температуры и солёности для двух дрейфтеров ARGO (D1 и D2) и модельные профили для экспериментов A01 (контрольный расчет) и A04 (усвоение данных AVISO), интерполированных в точки расположения дрейфтеров (данные за 2008-06-29). Дрейфтер D1 на дату 2008-06-29 располагался в точке с координатами 30.4° с.ш., 38.1° з.д., дрейфтер D2 имел координаты 40.5° с.ш., 56.0° з.д. Дрейфтеры выбраны таким образом, чтобы D1 был в открытом океане, а D2 – в области течения Гольфстрим. Более наглядно их расположение показано на **рисунке 4.9**.

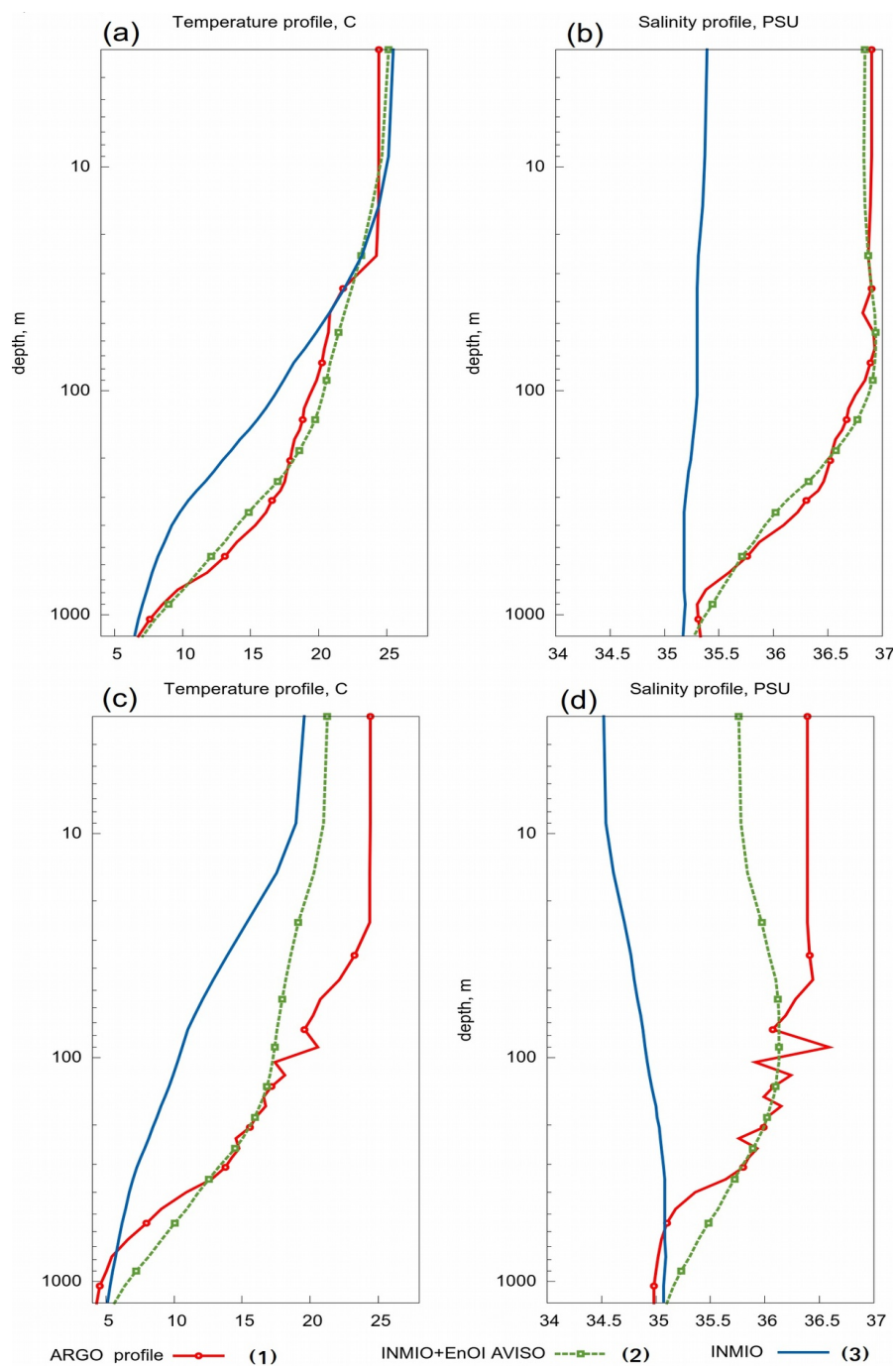


Рисунок 4.8. Профили температуры и солёности

- (1) двух дрейфтеров ARGO D1 (a,b) и D2 (c,d);
 (2) модельные из эксперимента A04 с усвоением альтиметрии AVISO;
 (3) модельные из контрольного эксперимента A01;

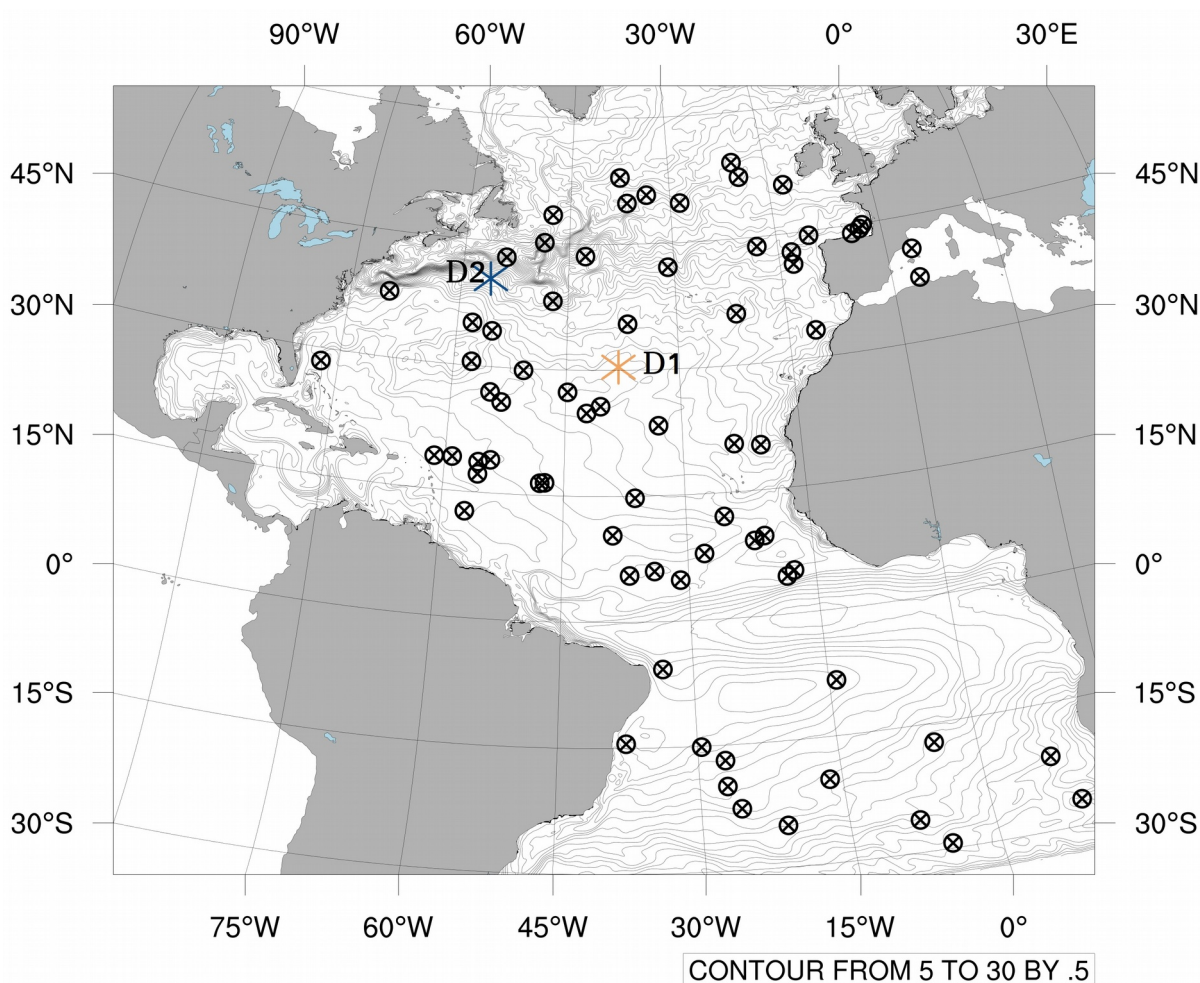


Рисунок 4.9. Поле поверхностной температуры в эксперименте A01. Звездочками показано расположение дрейфтеров D1 (30.4° с.ш., 38.1° з.д.) и D2 (40.5° с.ш., 56.0° з.д.). Черными кругами показаны остальные дрейфтеры ARGO, данные профилей температуры и солёности с которых поступили 2008-06-29.

Видно, что на глубине от 3 до 1000 метров в эксперименте с усвоением данных альтиметрии (эксперимент A04) значительно лучше воспроизводится вертикальный профиль температуры и солёности, представленный данными ARGO по сравнению с экспериментом без усвоения (A01), как для области открытого океана вдали от интенсивных

течений, так и для области течения Гольфстрим, в которой происходят активные динамические процессы. В открытом океане в точке дрейфера D1 модель с усвоением практически идеально воспроизводит профиль данных дрейфера: ошибка для температуры не превышает 1°C и 0.2 psu для солёности на всей глубине. Для динамически активного района в точке расположения дрейфера D2 результат в эксперименте A04 хуже, но все равно ошибка для температуры и солёности уменьшается практически в два раза по сравнению с экспериментом A01. Здесь, скорее всего, сказывается довольно простая схема параметризации вертикального перемешивания методом Манка-Андерсона, используемая в модели океана.

4.7 Сравнение с результатами, опубликованными в других работах

В работе [46], в которой проводились эксперименты с тем же методом усвоения, но для модели NUCOM показано, что усвоение аномалий уровня океана уменьшает ошибку прогноза для независимых наблюдений, в частности, для буев PIRATA по температуре на поверхности на 0.3 градуса в среднем, причем уменьшение ошибки имеет место и для глубоких слоев, до 500 м включительно. Аналогичный эффект имеет место и для солёности, где величина коррекции ошибки прогноза достигает 0.5 psu .

В статье [105], где усвоение данных также выполняется методом EnOI в системе BODAS, ошибка температуры на глубине от 0 до 1500 метров по сравнению с данными дрейферов ARGO находится в пределах 1.5 градусов, что меньше на два градуса, чем для модели без усвоения, при этом в качестве данных для усвоения используются спутниковая ТПО и уровень океана, а также данные самих дрейферов ARGO. Похожие результаты были также получены в работе [6]. Таким образом, уменьшение ошибки модельного прогноза за счет усвоения данных спутниковой

альтиметрии, полученное в данной работе методом EnOI, соответствует мировому уровню.

4.8 Выбор параметров метода EnOI

На **рисунке 4.10** приведены результаты экспериментального исследования чувствительности функции ошибки к количеству элементов в ансамбле (параметр N) и коэффициенту на диагонали матрицы R (далее R_c) при усвоении данных ARGO. Разница между ошибкой для экспериментов с ансамблем из 20 и 40 элементов довольно существенная, в то время как разница между ансамблем из 40 и 60 элементов незначительна. При этом для хранения матрицы ансамбля уже недостаточно памяти одного вычислительного узла суперкомпьютера «МВС-10П» (16 процессорных ядер Intel Xeon и 32 гигабайта памяти), хотя скорость усвоения данных на одном узле достаточная. Поэтому для усвоения данных в модели океана ИВМИО размер ансамбля полагается равным 50. Отметим, что теоретически метод Монте-Карло (расчет статистических характеристик по сформированной в результате экспериментов выборке) имеет погрешность порядка $N^{-1/2}$, поэтому дальнейшее увеличение точности в зависимости от числа членов ансамбля требует существенных вычислительных затрат.

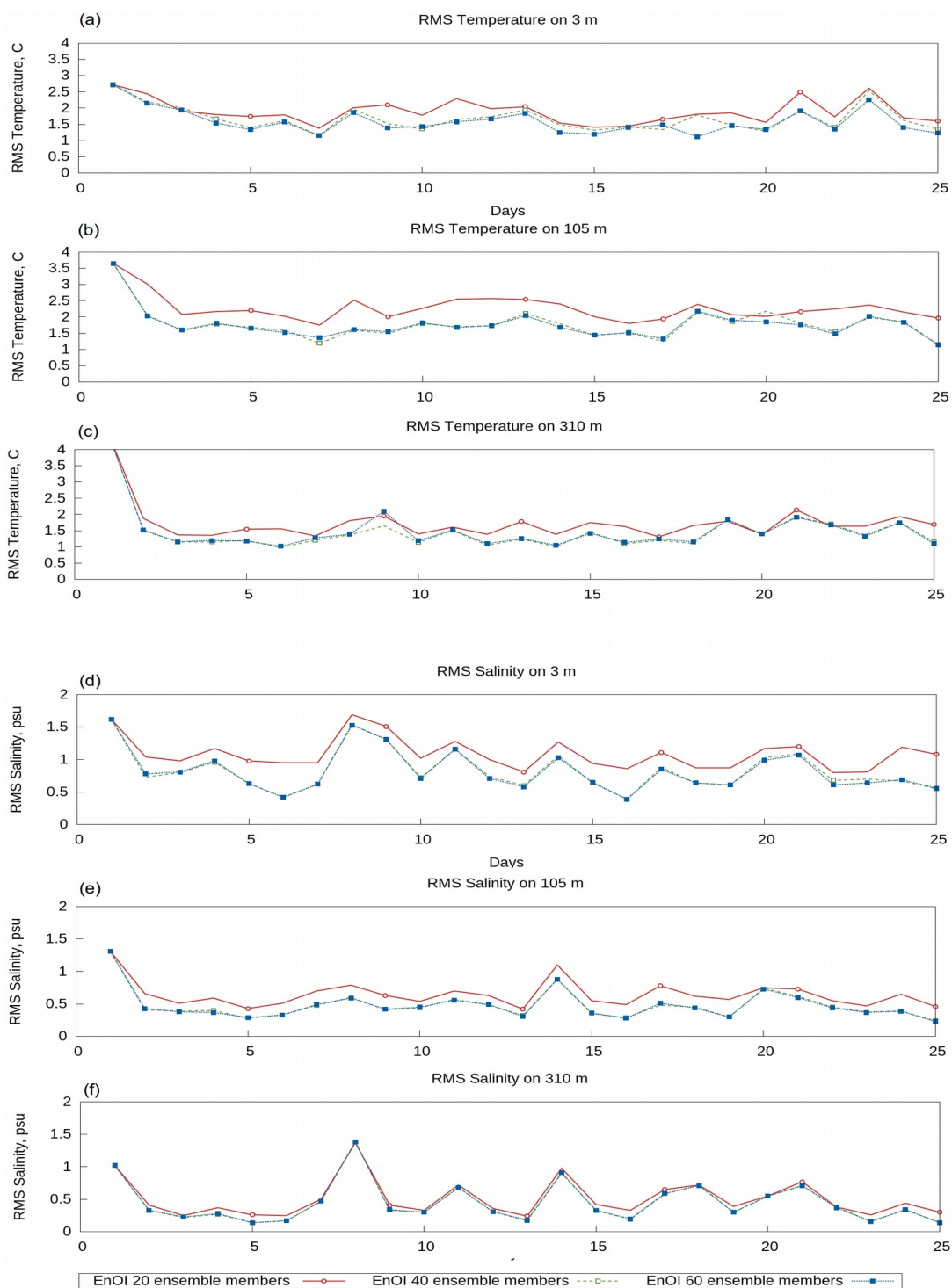


Рисунок 4.10. Среднеквадратичная ошибка контроля при усвоении методом EnOI при использовании ансамбля из 20 (EnOI 20), 40 (EnOI 40) и 60 (EnOI 60) элементов для поля температуры в °C на глубине 3(a), 105(b)

и 310 (с) метров и для поля солености в psu на глубине 3(d), 105(e) и 310(f) метров.

На рисунке 4.11 приведены результаты экспериментального исследования чувствительности функции ошибки к коэффициенту на диагонали матрицы R (далее R_c) при усвоении данных ARGO. По оси X — дни, по оси Y — ошибка температуры в градусах или ошибка солености в psu. Из рисунка видно, что разница между графиками ошибки для экспериментов с R_c изменяющимся в диапазоне от 0.02 до 5 не является критичной или носит хаотический характер (для поверхностной температуры). В тоже время стоит отметить, что при $R_c=100$, ошибка становится существенно выше, а при $R_c=0$ при усвоении данных происходит резкое нарушение теплового баланса — температура модельного поля в некоторых точках становится ниже -2 °C, а лед при этом не успел образоваться, что приводит к нефизическому решению (голубой график обрывается на 3 дне интегрирования модели). Таким образом, выбор коэффициента R_c — это компромисс между силой, с которой модельное решение притягивается к данным наблюдения и не нарушением геофизического баланса в модели. На основе численного эксперимента коэффициент R_c выбирается равным 1.5.

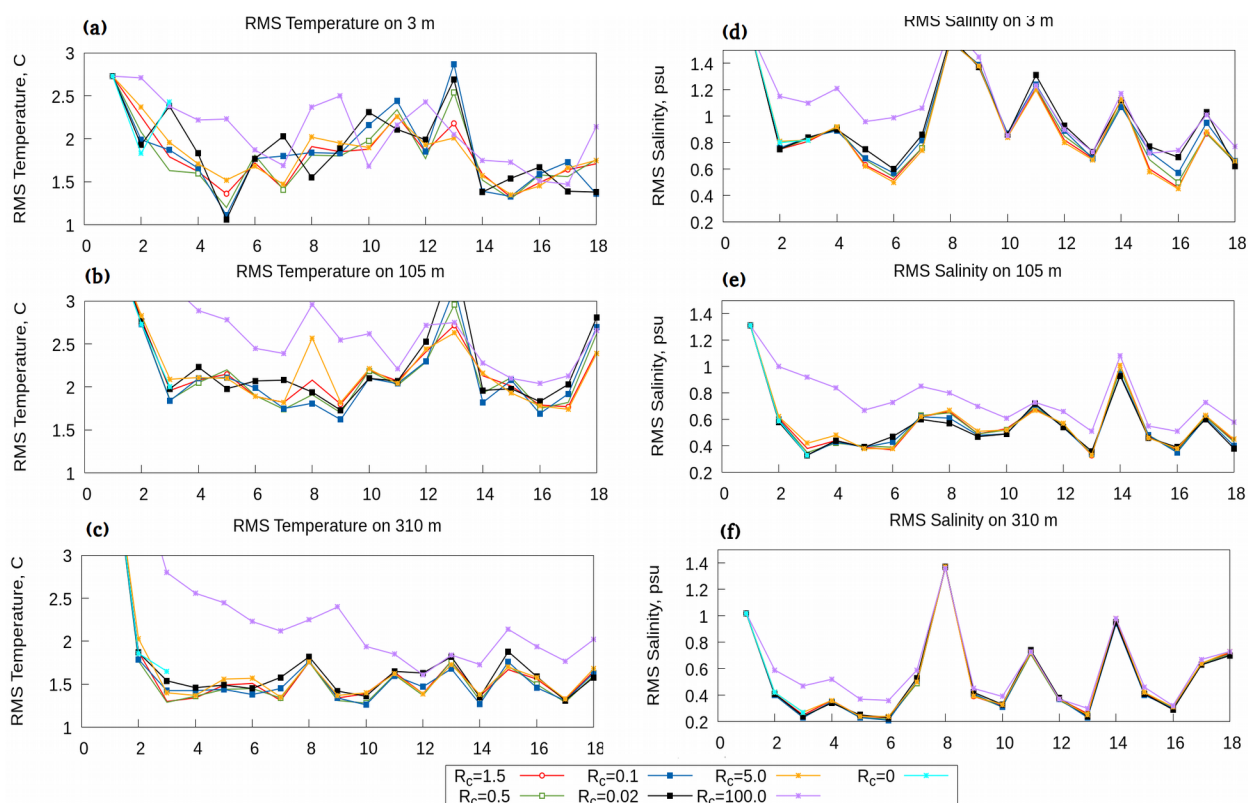


Рисунок 4.11. Среднеквадратичная ошибка контроля при усвоении данных ARGO методом EnOI при коэффициентах на диагонали матрицы R_c равным 1.5 (красный), 0.5 (зеленый), 0.1(синий), 0.02 (черный), 5.0(желтый), 100 (фиолетовый), 0 (голубой) для поля температуры в °C на глубине 3 (a), 105 (b) и 310 (c) метров и для поля солености в psu на глубине 3 (d), 105 (e) и 310 (f) метров.

4.9 Сравнение с независимыми спутниковыми данными ТПО

Обратим внимание, что ТПО в районе Мексиканского залива в эксперименте A03 (рисунок 4.12b) превышает значение ТПО контрольного эксперимента A01 (рисунок 4.12a), при этом разница в среднем составляет около 2° C. При сравнении с независимыми спутниковыми данными ТПО (ARMOR 3d, рисунок 4.12c) [86] видно, что

расчетные поля A03 заметно ближе к наблюдениям, разница в среднем составляет меньше 1° C. Это свидетельствует о том, что схема усвоения совместно с моделью корректно приближает поля модели к измерениям, при этом сохраняя имеющийся в реальных физических полях баланс тепла и массы. Стоит дополнительно отметить, что за 2 месяца эксперимента в Мексиканском заливе не наблюдалось ни одного дрефтера ARGO, а данные ТПО ARMOR 3d никак не использовались в процедурах усвоения. Это подтверждает тот факт, что используемый метод EnOI эффективно корректирует модельное решение в области, где нет данных наблюдений для усвоения (в данном случае дрефтеров ARGO). Важность этого свойства отмечалась в **разделе 2.3**.

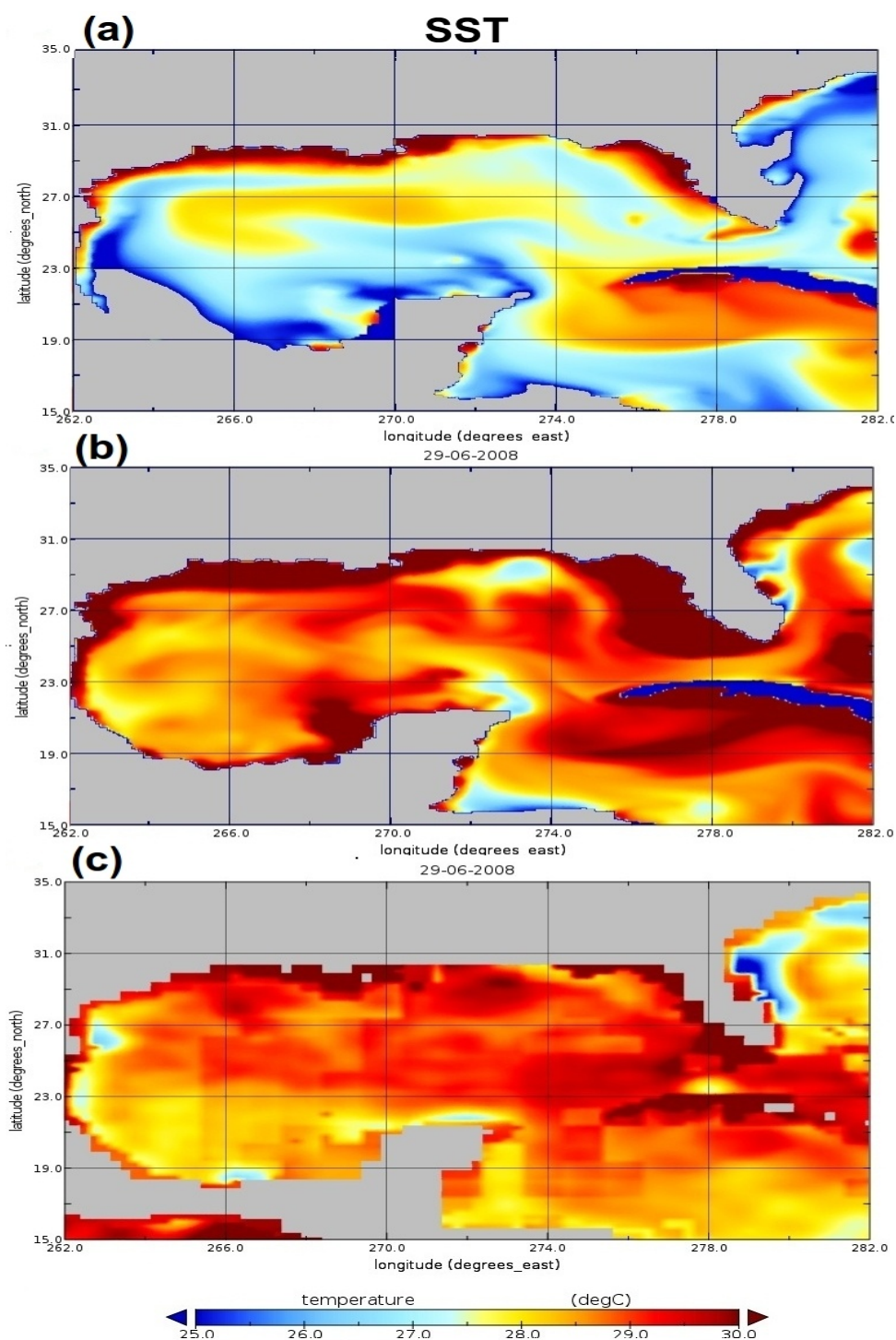


Рисунок 4.12. ТПО в °С в модели Северной Атлантики для Мексиканского залива за 2008-06-29; **(a)** в базовом эксперименте (A01); **(b)** в эксперименте с усвоением методом EnOI (A03) **(c)** спутниковые данные наблюдений Armor 3d

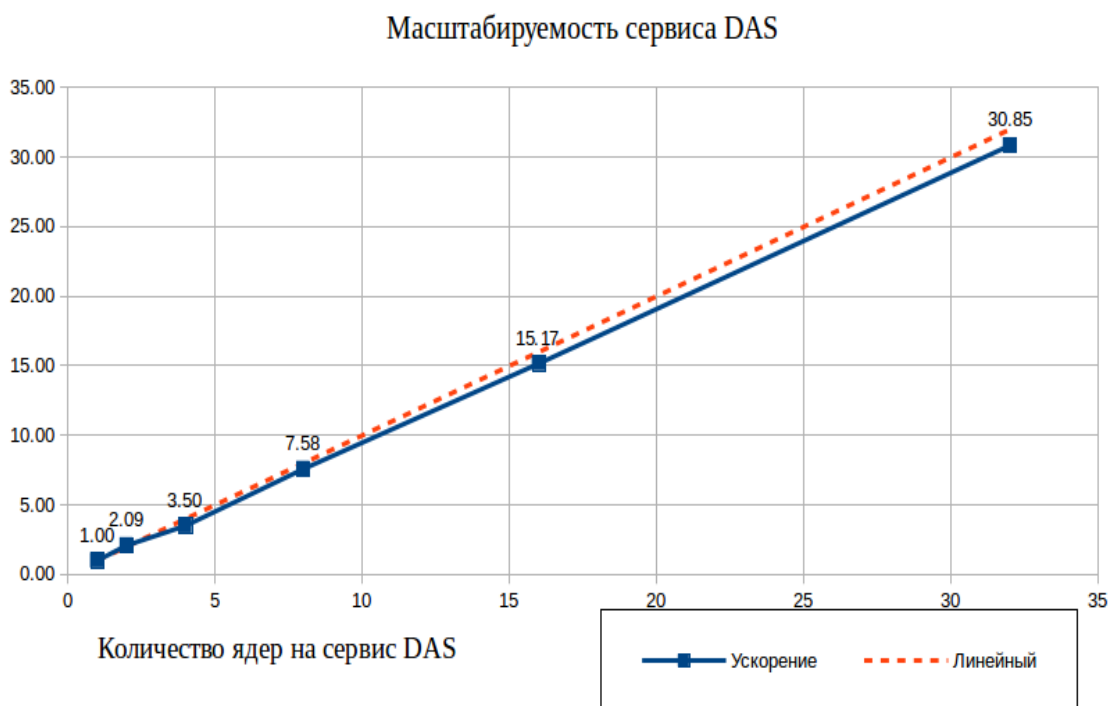


Рисунок 4.13. Масштабируемость метода EnOI в рамках сервиса DAS при усвоении 10^4 точек на суперкомпьютере «Ломоносов»

4.10 Параллельная эффективность

Благодаря эффективной реализации метода EnOI в качестве параллельного программного сервиса DAS, решение задачи усвоения данных масштабируется практически линейно (**рисунок 4.13**). Так усвоение 10^4 точек наблюдений на 16 процессорных ядрах занимает около 20 секунд вместо 5 минут на одном ядре, что было бы сравнимо со временем, затрачиваемым на суточный прогноз модели Атлантического океана на 200 ядрах.

На **рисунке 4.14** приведены результаты масштабируемости модели океана ИВМИО на массивно-параллельных компьютерах серии BlueGene, полученные в работе [16], как подтверждение правомерности высоких

требований, предъявляемых к параллельному алгоритму усвоения данных, работающему совместно с этой моделью океана.

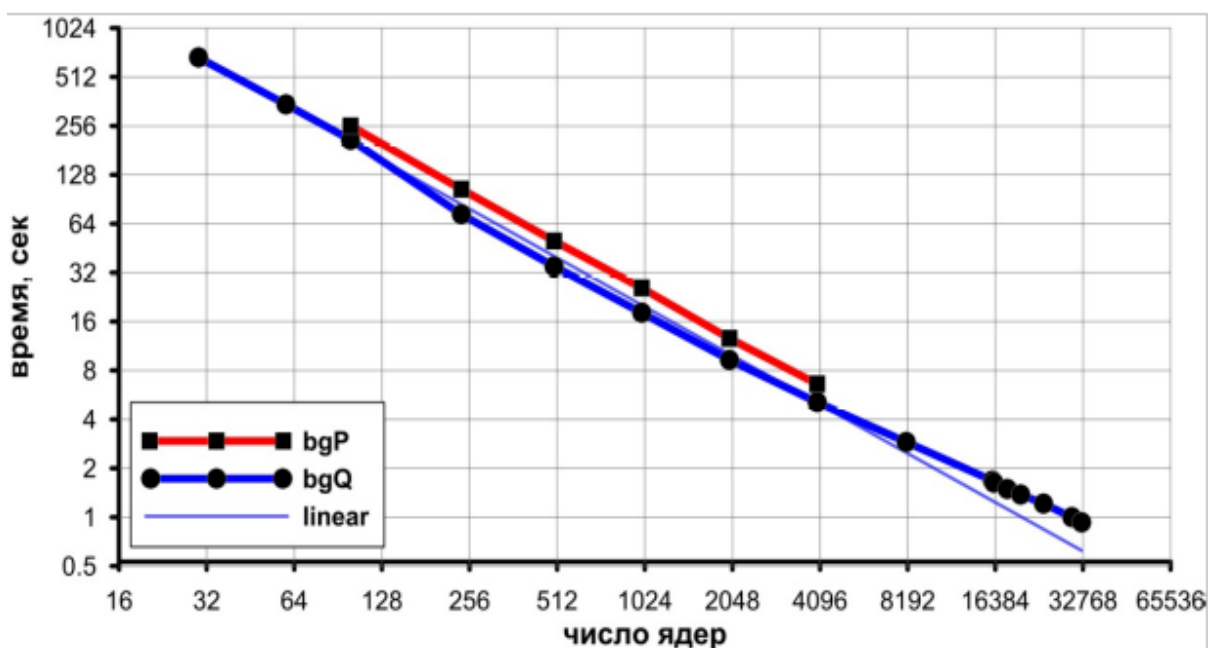


Рисунок 4.14. Время работы 10 шагов модели Мирового океана ИВМИО в зависимости от числа вычислительных ядер на суперкомпьютерах «BlueGeneP» (МГУ им. Ломоносова) и «BlueGeneQ» (IBM Research Center Thomas J. Watson). Рисунок взят из [16].

4.11 Выводы

В главе 4 на примере региона Северной Атлантики проведено комплексное тестированное разработанной системы усвоения данных наблюдений на базе математической модели динамики океана ИВМИО.

Показана хорошая масштабируемость алгоритма усвоения, которая позволит использовать его в модели Мирового океана с высоким пространственным разрешением и с большим количеством усваиваемых данных.

Разработанная система проверена на данных спутниковой альтиметрии AVISO и данных о температуре и солености с дрейфтеров ARGO.

На примере данных ARGO подтверждена и оценена количественно эффективность реализованного метода ансамблевой оптимальной интерполяции (EnOI) перед используемой ранее многомерной оптимальной интерполяции (MVOI), в частности показано, что ансамблевый метод имеет меньшую до 30% ошибку прогноза, чем метод MVOI, при этом разность этих ошибок растет с глубиной. Результаты расчетов модельного поля поверхностной температуры океана были проверены независимыми спутниковыми данными ТПО (ARMOR 3d), и на примере Мексиканского залива наглядно показано, что метод EnOI уменьшает ошибку в среднем на 1.5°C . Показано, что прогноз характеристик после усвоения дает меньшую в 2 и более раз погрешность по сравнению с контрольным расчетом. Оценена также зависимость ошибки прогноза от числа элементов ансамбля и коэффициента на диагонали матрицы R , в частности показано, что значимый эффект в точности достигается при увеличении числа членов ансамбля до 40 и при $R_c=1.5$.

Показано, что ассимиляция данных альтиметрии AVISO методом EnOI в модели океана ИВМИО дает значительное уменьшение ошибок в трехмерных полях температуры и солености по сравнению с экспериментами без ассимиляции. Ошибки прогноза после усвоения по сравнению с контрольным расчетом уменьшаются почти в два раза, при этом ошибка уменьшается с глубиной. Показано также как конкретно усвоение уровня влияет на те параметры, которые не усваивались непосредственно - температуру и соленость. Полученные модельные поля физических характеристик, в частности поля температуры, физически

достоверны и соответствуют непосредственно наблюдаемым.

Все полученные результаты по усвоению данных наблюдений в модели океана ИВМИО соответствуют результатам других научных коллективов из ведущих прогностических центров (см. **таблица 2.3**).

При проведении численных экспериментов, описанных в этой главе, использовались ресурсы суперкомпьютерного комплекса МГУ М.В. Ломоносова и Межведомственного суперкомпьютерного центра РАН.

Заключение

Основной результат диссертационной работы – разработан параллельный алгоритм и создана программная архитектура системы усвоения спутниковых и дрейфтерных данных наблюдений на основе динамико-стохастического подхода (MVOI, EnOI) в модели динамики океана высокого пространственного разрешения.

Получены следующие частные результаты:

1. Разработан параллельный алгоритм многомерной оптимальной интерполяции (MVOI) для усвоения данных наблюдений ARGO в модели динамики океана ИВМИО.
2. Создана программная архитектура и разработан параллельный алгоритм ансамблевой оптимальной интерполяции (EnOI), масштабируемый для нерегулярных по пространству спутниковых и дрейфтерных данных наблюдений.
3. Написана программная реализация параллельного алгоритма усвоения данных EnOI в виде программного сервиса вычислительной платформы совместного моделирования CMF 3.0. Вычислительная эффективность и параллельная масштабируемость разработанной системы подтверждены тестами на различных современных суперкомпьютерах массивно-параллельной архитектуры («Ломоносов», «МВС-10П», «РСК Торнадо»).
4. На основе серии численных экспериментов для модели динамики Северной Атлантики высокого разрешения ИВМИО с усвоением данных наблюдений о температуре и солености с дрейфтеров ARGO и данных спутниковой альтиметрии AVISO получены качественные и количественные оценки точности решения на основе независимых

данных наблюдений, которые соответствуют результатам ведущих научных коллективов.

Список сокращений

3DVar	-	variational three-dimensional assimilation
4DVar	-	variational four-dimensional assimilation
AVHRR	-	Advanced Very High Resolution Radiometer
CMF	-	Compact Modeling Framework
CORE	-	Coordinated Ocean-ice Reference Experiments
DAS	-	Data Assimilation Service
DEnKF	-	Deterministic Ensemble Kalman filter
ECMWF	-	European Center for Medium-Range Weather Forecasts
EKF	-	Extended Kalman Filter
EnKF	-	Ensemble Kalman Filter
EnOI	-	Ensemble Optimal Interpolation
GA	-	Global Arrays
GODAE	-	Global Ocean Data Assimilation Experiment
GODAS	-	Global Ocean Data Assimilation System
HYCOM	-	HYbrid Coordinate Ocean Model
INM	-	Institute of Numerical Mathematics – ИБМ РАН
JMA	-	Japan Meteorological Agency
LETKF	-	Local Ensemble Transform Kalman Filter
MFC	-	Monitoring & Forecasting Centers
MODAS	-	Modular Ocean Data Assimilation System
MODIS	-	Moderate Resolution Imaging Spectroradiometer
MOM	-	Modular Ocean Model
MOVE	-	Multivariate Ocean Variational Estimation
MPI	-	Message Passing Interface
MRI.COM	-	Meteorological Research Institute Community Ocean Model
MVOI	-	MultiVariate Optimum Interpolation

NCEI	-	National Centers for Environmental Information
NCEP	-	National Centers for Environmental Prediction
NCODA	-	Navy Coupled Ocean Data Assimilation
NCOM	-	Navy Coastal Ocean Model
NEMO	-	Nucleus for European Modeling of the Ocean
NLOM	-	NRL Layered Ocean Model
NOAA	-	National Oceanic and Atmospheric Administration
NRL	-	Naval Research Laboratory
PGAS	-	Partitioned Global Address Space
POM	-	Princeton Ocean Model
POP	-	Parallel Ocean Program
RTOFS	-	Real-Time Ocean Forecast System
SLA	-	Sea Level Anomaly
SMI	-	Standard Mapped Image
SREnKF	-	Square Root Ensemble Kalman filter
SSHA	-	Sea Surface Height Anomaly
SST	-	Sea Surface Temperature
VIIRS	-	Visible Infrared Imaging Radiometer Suite
ГМЦ	-	Гидрометцентр России
ИСЗ	-	Искусственный Спутник Земли
ПКСМ	-	Программный Комплекс Совместного Моделирования
ТПО	-	Температура Поверхности Океана

Приложение

В приложении приведено краткое описание разработанных в рамках данной диссертации программ на которые получены «Свидетельства о государственной регистрации».

Свидетельство 1. Программа для усвоения данных температуры, солености и спутниковой альтиметрии в модель Мирового океана методом оптимальной интерполяции

Авторы: Кауркин Максим Николаевич, Ибраев Рашит Ахметзиевич

Правообладатель:

Федеральное государственное бюджетное учреждение «Гидрометеорологический научно-исследовательский центр Российской Федерации»

Свидетельство о государственной регистрации программы для ЭВМ №2015661002.

Дата регистрации: 14.10.2015

Аннотация:

Программа предназначена для автоматизированного получения, интерполяции, предварительной диагностики и усвоения данных о температуре, солености и уровне в модель Мирового океана, с целью повышения качества модельного прогноза. Алгоритм метода оптимальной интерполяции адаптирован для работы на массивно-параллельных компьютерах с распределенной памятью (с использованием технологии MPI) на регулярных модельных сетках с высоким пространственным разрешением. В программе реализован интерфейс чтения данных в формате

NETCDF и распределения между вычислительными узлами данных наблюдений системы дрейфующих буев АРГО (температуры, солености) и данных спутниковой альтиметрии АВИЗО (аномалии уровня океана). Написан блок для контроля качества входных данных и выполнения их первичной обработки. Написана подпрограмма, производящая оценку практической пользы от использования усвоения данных наблюдений с точки зрения уменьшения среднеквадратичной ошибки прогноза и анализа модельных величин (температуры, солености, аномалии уровня океана).

Язык: FORTRAN, bash, MPI

ОС: UNIX/LINUX

Свидетельство 2. Программа усвоения данных наблюдений методом EnOI для модели океана высокого пространственного разрешения на базе платформы CMF3.0 для моделирования на массивно-параллельных вычислительных системах

Авторы: Кауркин Максим Николаевич, Ибраев Рашит Ахметзиевич

Правообладатель:

Федеральное государственное бюджетное учреждение
«Гидрометеорологический научно-исследовательский центр Российской Федерации»

*Свидетельство о государственной регистрации программы для ЭВМ
№2017610805*

Дата регистрации: 18.01.2017

Аннотация:

Программа предназначена для интерполяции, предварительной диагностики и усвоения данных наблюдений о температуре, солености и уровне в модель Мирового океана методом EnOI, с целью повышения качества модельных расчетов. Алгоритм метода ансамблевой оптимальной интерполяции (EnOI) адаптирован для работы на массивно-параллельных компьютерах с распределенной памятью (с использованием технологии MPI, Global Arrays и CMF3.0) на отдельных от модели океана вычислительных ядрах. В программе реализован интерфейс параллельного чтения данных наблюдений и ансамбля состояний модели в формате NETCDF и распределения данных между вычислительными ядрами. Реализован блок для контроля качества входных данных и выполнения их первичной обработки для данных с дрейфтеров АРГО (температуры и солености) и данных спутниковой альтиметрии АВИЗО (аномалии уровня океана).

Язык: FORTRAN, bash, MPI, Global Arrays

ОС: UNIX/LINUX

Литература

1. Агошков В.И., Асеев Н.А., Захарова Н.Б. и др. Информационно-вычислительная система "ИВМ РАН - Балтийское море". Москва: ИВМ РАН. 2016. 139 с.
2. Агошков В.И., Асеев Н.А., Гиниатулин С.В. и др. Информационно-вычислительная система "ИВМ РАН - Черное море". Москва: ИВМ РАН, 2016. 136 с.
3. Агошков В.И., Ассовский М.В. Математическое моделирование динамики Мирового океана с учетом приливообразующих сил. 2016. 122 с.
4. Агошков В.И., Ипатова В.М., Залесный В.Б., Пармузин Е.И., Шутяев В.П. Задачи вариационной ассимиляции данных наблюдений для моделей общей циркуляции океана и методы их решения // Известия РАН. Физика атмосферы и океана. 2010. Т. 46. № 6. С. 734–770.
5. Агошков В.И., Пармузин Е.И., Шутяев В.П. Ассимиляция данных наблюдений в задаче циркуляции Черного моря и анализ чувствительности ее решения // Известия РАН. Физика атмосферы и океана. 2013. Т. 49. № 6. С. 643–654.
6. Беляев К.П., Танажура К.А.С., Тучкова Н.П. Сравнительный анализ экспериментов с усвоением данных дрейфтеров АРГО // Океанология. 2012. Т. 52(5). С. 643–653.
7. Гандин Л.С. Объективный анализ метеорологических полей. — Л., Гидрометеоиздат. 1963. 118 с.
8. Гилл А. Динамика атмосферы и океана. Т.1-2. Москва, Мир. 1986. 397 с.
9. Голубева Е.Н., Платов Г.А., Якшина Д.Ф. Численное моделирование современного состояния вод и морского льда Северного Ледовитого океана // Лед и Снег. 2015. Т. 55(2). С. 81–92.

10. Громов И.В., Коромыслов А.Ю., Ушаков К.В., Кауркин М.Н., Ибраев Р.А. Совместная модель внутригодовой изменчивости циркуляции вод и льда северного ледовитого океана // Труды Гидрометцентра России. 2016. № 361. С. 29-46.
11. Дианский Н.А., Багно А.В., Залесный В.Б. Сигма-модель глобальной циркуляции океана и ее чувствительность к вариациям напряжения трения ветра // Известия РАН. Физика атмосферы и океана. 2002. Т. 38. № 4. С. 537–556.
12. Егоров В.В. Космическая радиовысотометрия. Современное состояние, актуальные проблемы и пути их решения // Труды ВНИИЭМ. Вопросы электромеханики. 2011. Т. 120. № 1. С. 17-26.
13. Залесный В.Б., Марчук Г.И. Моделирование циркуляции мирового океана с четырехмерной вариационной ассимиляцией полей температуры и солености // Известия РАН. Физика атмосферы и океана. 2012. Т. 48. № 1. С. 21–36.
14. Ибраев Р.А. Математическое моделирование термогидродинамических процессов в Каспийском море. М. ГЕОС. 2008. 128 с.
15. Ибраев Р.А., Хабеев Р.Н., Ушаков К.В. Вихреразрешающая $1/10^\circ$ модель Мирового океана // Известия РАН. Физика атмосферы и океана. 2012. Т. 48. № 1. С. 45–55.
16. Калмыков В.В. Программный комплекс численного моделирования совместной системы океан-атмосфера на массивно-параллельных компьютерах: дис. канд. физ.- мат. наук: 05.13.18 — М., 2013. 139 с.
17. Калмыков В.В., Ибраев Р.А. Быстрый алгоритм решения системы уравнения мелкой воды на компьютерах с распределенной памятью // Вестник УГАТУ. 2013. Т. 17. №5(58). С. 252–259.

18. Калмыков В.В., Ибраев Р.А. Программный комплекс совместного моделирования системы океан-лед-атмосфера-почва на массивно-параллельных компьютерах // ИВМ РАН. Свидетельство о государственной регистрации программы для ЭВМ №2013619320. Дата регистрации: 01.10.2013.
19. Калмыков В.В., Ибраев Р.А. Программный комплекс совместного моделирования системы океан-лед-атмосфера-почва на массивно-параллельных компьютерах // Вычислительные методы и программирование. 2013. № 14. С. 88–95.
20. Калмыков В.В., Ибраев Р.А. SMF - фреймворк для модели Земной системы высокого разрешения. В сборнике: Суперкомпьютерные дни в России. Труды международной конференции. Суперкомпьютерный консорциум университетов России, Федеральное агентство научных организаций России. Москва, 2015. С. 34–40.
21. Калмыков В.В., Ибраев Р.А., Ушаков К.В. Компактная вычислительная платформа SMF3.0 для моделирования Земной системы на массивно-параллельных компьютерах // ФГБУ "ГИДРОМЕТЦЕНТР РОССИИ". Свидетельство о государственной регистрации программы для ЭВМ №2017610785. Дата регистрации: 18.01.2017.
22. Калмыков В.В., Р.А. Ибраев, К.В. Ушаков. Проблемы и вызовы при создании модели Земной системы высокого разрешения // Суперкомпьютерные технологии в науке, образовании и промышленности: Альманах под ред. Садовниченко В.А., Савина Г.И., Воеводина В.В. М.: Издательство Московского университета. 2014. С. 14–22.
23. Каменкович В.М. Основы динамики океана. Ленинград. Гидрометиздат. 1973. 240 с.
24. Кауркин М.Н., Ибраев Р.А., Беляев К.П. Усвоение данных наблюдений в модели динамики океана высокого пространственного

разрешения с применением методов параллельного программирования // Метеорология и гидрология. – 2016. №7. С. 47–57.

25. Кауркин М.Н., Ибраев Р.А., Беляев К.П. Усвоение данных наблюдений в модели динамики океана высокого пространственного разрешения с применением методов параллельного программирования // Океанология. 2016. Т. 56. №6. С. 252–260.

26. Кауркин М.Н., Ибраев Р.А. Программа усвоения данных наблюдений методом EnOI для модели океана высокого пространственного разрешения на базе платформы CMF3.0 для моделирования на массивно-параллельных вычислительных системах // ФГБУ "ГИДРОМЕТЦЕНТР РОССИИ". Свидетельство о государственной регистрации программы для ЭВМ №2017610805. Дата регистрации: 18.01.2017.

27. Кауркин М.Н., Ибраев Р.А. Программа усвоения данных температуры, солености и спутниковой альтиметрии в модель мирового океана методом оптимальной интерполяции // ФГБУ "ГИДРОМЕТЦЕНТР РОССИИ". Свидетельство о государственной регистрации программы для ЭВМ №2015661002. Дата регистрации: 14.10.2015.

28. Кауркин М.Н., Ибраев Р.А. Разработка технологии усвоения данных наблюдений на основе ансамблевых фильтров Калмана в модель океана сверхвысокого пространственного разрешения // Материалы молодежной научной конференции «Комплексные исследования морей России: оперативная океанография и экспедиционные исследования». 2016. С. 76–81.

29. Кауркин М.Н., Ибраев Р.А., Беляев К.П., Тучкова Н.П. Параллельный сервис усвоения данных наблюдений в гидродинамическую модель высокого разрешения // В сборнике: Научный сервис в сети Интернет труды XVIII Всероссийской научной конференции. ИПМ им. М.В. Келдыша РАН. 2016. С. 188–193.

30. Кауркин М.Н., Ибраев Р.А., Коромыслов А.Ю. Технология усвоения данных аэрокосмического мониторинга и измерений буев Арго методом EnOI в модели глобального океана высокого пространственного разрешения с использованием платформы CMF // Суперкомпьютерные дни в России: Труды международной конференции (26 – 27 сентября 2016 года, г. Москва) – М.: Изд-во МГУ. 2016. С.732–740.
31. Кауркин М.Н., Тучкова Н.П., Беляев К.П., Михайлов Г.М., Ибраев Р.А., Сальников А.Н. Параллельное усвоение данных наблюдений в гидродинамических моделях высокого пространственного разрешения // «Научный сервис в сети Интернет: многообразие суперкомпьютерных миров» Труды Международной суперкомпьютерной конференции. 2014. С. 236–239
32. Кныш В.В., Коротаев Г.К., Лишаев П.Н. Методика использования трехмерных полей температуры и солёности Черного моря, восстановленных по малочисленным данным измерений и альтиметрии, в оперативной прогностической модели // Морской гидрофизический журнал. 2016. Т. 188. № 2. С. 53–69.
33. Кныш В.В., Коротаев Г.К., Мизюк А.И., Саркисян А.С. Усвоение гидрологических наблюдений для расчета течений в морях и океанах // Известия РАН. Физика атмосферы и океана. 2012. Т. 48. № 1. С. 67–85.
34. Кныш В.В., Ибраев Р.А., Коротаев Г.К., Инюшина Н.В.. Сезонная изменчивость климатических течений Каспийского моря, восстановленная ассимиляцией климатической температуры и солёности в модели циркуляции вод // Известия РАН. Физика атмосферы и океана. 2008. Т. 44 № 2. С. 251–265.
35. Кныш В.В., Ярин В.Д. Некоторые гидродинамические характеристики и статистическая структура полей температуры и плотности на полигоне // Экспериментальные исследования по

международной программе “ПОЛИМОДЕ” (результаты 17го рейса НИС “Академик Вернадский” и 33го рейса НИС “Михаил Ломоносов”). Севастополь: МГИ АН УССР. 1978. С. 74–81.

36. Лебедев К.В., Саркисян А.С., Никитин О.П. Сравнительный анализ поверхностной циркуляции северной Атлантики, воспроизведенной тремя различными методами // Известия РАН. Физика атмосферы и океана. 2016. Т. 52. № 4. С. 465–474.

37. Марчук Г.И. Сопряженные уравнения и анализ сложных систем. М.: Наука. 1992. 336 с.

38. Марчук Г.И. Численное решение задач динамики атмосферы и океана. Л.: Гидрометеиздат. 1974. 303 с.

39. Пененко А.В., Пененко В.В. Прямой метод вариационного усвоения данных для моделей конвекции-диффузии на основе схемы расщепления // Вычислительные технологии. 2014. Т. 19. № 4. С. 69–83.

40. Пененко В.В. Вариационные методы усвоения данных и обратные задачи для изучения атмосферы, океана и окружающей среды // Сибирский журнал вычислительной математики. 2009. Т. 12. № 4. С. 421–434.

41. Пененко В.В., Образцов Н.Н. Вариационный метод согласования полей метеорологических элементов // Метеорология и гидрология. 1976. №11. С. 1–11.

42. Марчука Г.И. Проект научной программы по исследованию роли энергоактивных зон океана (ЭАЗО) в колебаниях климата ("Разрезы"). Под редакцией академика Марчука Г.И.. Москва. Гидрометеиздат. 1989. 110 с.

43. Саркисян А.С. Численный анализ и прогноз морских течений // Л.: Гидрометеиздат. 1977. 182 с.

44. Саркисян А.С., Кныш В.В., Демышев С.Г. и др. Многоэлементный четырехмерный анализ гидрофизических полей на основе динамико-

- стохастических моделей // Итоги науки и техники. Атмосфера, океан, космос – программа “Разрезы”. 1987. Т. 9. М.: ВИНТИ. С. 5–64.
45. Саркисян А.С., Никитин О.П., Лебедев К.В. Физические характеристики Гольфстрима как индикатор качества моделирования циркуляции Мирового океана // Доклады Академии наук. 2016. Т. 471. № 5. С. 595–598.
46. Танажура К.А.С., Лима Л.Н., Беляев К.П. Усвоение аномалий уровня океана в гидродинамической модели НУСОМ по данным наблюдений со спутников в Атлантике // Океанология. 2015. Т. 55. № 5. С. 738–750.
47. Тимченко И.Е. Динамико-стохастические модели состояния океана. Наукова Думка, 1981. 192 с.
48. Толстых А.М., Ибраев Р.А. и др. Модели глобальной атмосферы и Мирового океана: алгоритмы и суперкомпьютерные технологии. Учеб. пособие. М.: Издательство Московского университета. 2013. 144 с.
49. Толстых М.А., Мизяк В.Г. Параллельная версия полулагранжевой модели ПЛАВ с горизонтальным разрешением порядка 20 км // Труды Гидрометеорологического научно-исследовательского центра Российской Федерации. 2011. вып. 346. С. 181–190.
50. Ушаков К.В., Гранкина Т.Б., Ибраев Р.А. Моделирование циркуляции вод Северной Атлантики в условиях эксперимента CORE-II // Известия РАН. Физика атмосферы и океана. 2016. Т. 52. № 5.
51. 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.
52. Хабеев Р.Н. Особенности циркуляции вод Северной Атлантики в трехмерной вихререзающей модели Мирового океана: дис. канд. физ.-мат. наук: 01.02.05 — М., 2013. 110 с.

53. Behringer D.W. The Global Ocean Data Assimilation System (GODAS) at NCEP. 11th Symp. on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface, San Antonio, TX. 2007. Amer. Meteor. Soc. URL: <https://ams.confex.com/ams/pdfpapers/119541.pdf>.
54. Belyaev K., Tanajura C.A.S., O'Brien J.J. Application of the Fokker-Planck Equation to Data assimilation into Hydrodynamical Models // J. Math. Sciences. 2000. V. 99. N. 4. P.1393–1402.
55. Bleck R. An oceanic general circulation model framed in hybrid isopycnic Cartesian coordinates // Ocean Model. 2002. V. 4. P. 55–88.
56. Bonavita M., Hólm E., Isaksen L., Fisher M. The evolution of the ECMWF hybrid data assimilation system // Quart. J. Roy. Meteor. Soc. 2016. V. 142. P. 287–303.
57. Boukas L.A., Mimikou N.T., Missirlis N.M. et al. The parallelization of the Princeton Ocean Model // Lecture Notes in Computer Sci., Amestoy et al. (Eds.). Springer. 1999. V. 1685. P. 1395–1402.
58. Bryan K. A numerical method for the study of the circulation of the World ocean // J. Comp. Physics. 1969. V. 4. P. 54–72.
59. Burchard H., Bolding K. Comparative Analysis of Four Second-Moment Turbulence Closure Models for the Oceanic Mixed Layer // J. Phys. Oceanography. 2001. V. 31. P. 1943–1968.
60. Chang Y.S., Garraffo Z.D., Peters H., Özgökmen T.M. Pathways of Nordic Overflows from climate model scale and eddy resolving simulations. // Ocean modeling. 2009. V. 29. P. 66–84.
61. Cummings J. Operational multivariate ocean data assimilation // Quart. J. Royal Met. Soc. 2005. V. 131. No. 613. P. 3583–3604.
62. Cummings J., Bertino L., Brasseur P. et al. Ocean Data Assimilation Systems for GODAE // Oceanography. 2009. V. 22. No. 3. P. 96–109.

63. Douville H., Viterbo P., Mahfouf J.-F., Beljaars A.C.M.. Evaluation of the Optimum Interpolation and Nudging Techniques for Soil Moisture Analysis Using FIFE Data // *Mon. Wea. Rev.* V. 128. P. 1733–1756.
64. Ensemble Kalman Filter Code. URL: <http://enkf.nersc.no/Code> .
65. ETOPO5. Data Announcement 88-MGG-02, Digital relief of the Surface of the Earth // NOAA, National Geophysical Data Center, Boulder, Colorado. 1988.
66. Evensen G. *Data Assimilation, The Ensemble Kalman Filter*, 2nd ed. Berlin: Springer. 2009. 307 p.
67. Evensen G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics // *J. Geophys. Res.* 1994. V. 99 P. 10143–10162.
68. Evensen G. The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation // *Ocean Dynamics*. 2003. V. 53(4). P. 343–367.
69. Fadeev R., Ushakov K., Kalmykov V. et al. Coupled atmosphere–ocean model SLAV–INMIO: implementation and first results // *Russian Journal of Numerical Analysis and Mathematical Modelling*. 2016. V. 31(6). P. 329–337.
70. Feder T. Argo begins systematic global probing of the upper oceans // *Physics Today*. 2000. V. 53. P. 50–51.
71. Feltham D.L., Untersteiner N., Wettlaufer J.S., Worster M.G. Sea ice is a mushy layer // *Geophys. Res. Lett.* 2006. V. 33. L14501.
72. GODAE OceanView Science Team. Work Plan 2009 – 2013. URL: <http://www.godae-oceanview.org>
73. Griffies S., Harrison M.J., Pacanowski R.C., Rosati A.. *A Technical Guide to MOM4*. GFDL Ocean Group Technical Report No. 5. Princeton, NJ: NOAA , Geophysical Fluid Dynamics Laboratory. 2004. 342 p.
74. Griffies S.M. 2004. *Fundamentals of Ocean Climate Models*. Princeton University Press, Princeton, USA. 518 p.

75. Griffies S.M., Biastoch A., Böning C. et al. Coordinated ocean-ice reference experiments (COREs) // *Ocean modelling*. 2009. V. 26. N. 1–2. P. 1–46.
76. Hallberg R.W. Stable split time stepping schemes for large-scale ocean modeling // *J. Comput. Phys.*, 1997. V. 135. P. 54–65.
77. Hunke E.C., Lipscomb W. H., Turner A.K., Jeffery N., Elliott S.. CICE: the Los Alamos Sea Ice Model Documentation and Software User’s Manual Version 5.1. 2015. URL: <http://oceans11.lanl.gov>
78. JMA: Outline of the operational numerical weather prediction at the Japan Meteorological Agency. Japan Meteorological Agency. 2007. 194 p.
79. Joseph J.H., Wiscombe W.J., Weinman J.A. The Delta-Eddington approximation for radiative flux transfer // *J. Atmos. Sci.* 1976. V. 33. P. 2452–2459.
80. Kalmykov V., Ibrayev R. CMF – framework for high-resolution Earth system modeling // *CEUR Workshop Proceedings*. V. 1482. Proceedings of the 1st Russian Conference on Supercomputing (RuSCDays 2015). Moscow, Russia. September 28–29, 2015. Moscow, MSU publishers. 2015. P. 34–40.
81. Kalnay E. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge: Cambridge University Press. 2003. 341 p
82. Kalnay E., Li H., Miyoshi T. et al. 4-D-Var or ensemble Kalman filter? // *Tellus A*. 2007. V. 59. No. 5. P. 758– 773.
83. Karl T.R. Key challenges for environmental data and information as viewed from NCEI // 3rd Meeting of Department of Commerce Data Advisory Council, 29 – 30 October 2015. Boulder, CO, USA.
84. 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*. Eds. Voevodin V., Sobolev S. Series: Communications in Computer and Information Science. Springer. 2016. V. 687. P. 57–66.

85. Large W., Yeager S. The global climatology of an interannually varying air–sea flux data set // *Clim. Dyn.* 2009. V. 33. No. 2–3. P. 341–364.
86. Larnicol G., Guinehut S., Rio M.-H. et al. The Global Observed Ocean Products of the French Mercator project // *Proceedings of 15 Years of progress in Radar Altimetry conference, 2006.* SP-614.
87. Launiainen J., Vihma T. Derivation of turbulent surface fluxes - an iterative flux-profile method allowing arbitrary observing heights // *Environmental Software.* 1990. V. 5(3). P. 113–124.
88. Levitus S. (Ed.) *World Ocean Atlas 2009 (WOA09).* NOAA Atlas NESDIS. URL: www.nodc.noaa.gov/OC5/WOA09/pr_woa09.html
89. Lunde, B.N., Coelho E.F. Implementations of the navy coupled ocean data assimilation system at the naval oceanographic office // *IEEE/MTS Oceans.* 2009.
90. Madec G. NEMO reference manual, ocean dynamics component: NEMO-OPA. Preliminary version. Note du Pole de modelisation 27, Institut Pierre-Simon Laplace (IPSL), France. 2008.
91. Malone R.C., Smith R.D., Maltrud M.E., Hecht M.W. Eddy-Resolving Ocean Modeling // *Los Alamos Science.* 2003. V. 28. P. 223–231.
92. Maltrud M.E., McClean J.L. An eddy resolving global 1/10° ocean simulation // *Ocean Modelling.* 2005. V. 8. P. 31–54.
93. Marchuk G.I., Penenko V.V. Application of optimization methods to the problem of mathematical simulation of atmospheric processes and environment // *Modelling and Optimization of Complex Systems/Ed. G.I. Marchuk. Proc. Of the IFIP-TC7 Working conf.* New York: Springer. 1978. P. 240–252.
94. Marchuk G.I., Sarkisyan A.S. *Mathematical modelling of ocean circulation.* Springer-Verlag. 1988. 292 p.

95. Measuring sea surface temperature. URL: <http://hrdds.ifremer.fr/component/content/article/14-sample-data-articles/80-measuring-sea-surface-temperature-data-products-properties>
96. Miyoshi T., Sato Y., Kadowaki T. Ensemble Kalman Filter and 4D-Var Intercomparison with the Japanese Operational Global Analysis and Prediction System. // *Mon. Wea. Rev.* 2010. V. 138. P. 2846–2866.
97. Mogensen K, Balmaseda M.A., Weaver A. The NEMOVAR ocean data assimilation system as implemented in the ECMWF ocean analysis for System 4. European Centre for Medium-Range Weather Forecasts. 2012.
98. Mogensen K, Balmaseda MA, Weaver AT, Martin M, Vidard A. NEMOVAR: A variational data assimilation system for the NEMO model // *ECMWF Newsletter*. 2009. V. 120: 17–22.
99. Murray R.J. Explicit Generation of Orthogonal Grids for Ocean Models // *J. Comp. Physics*. 1996. V. 126. P. 251–273.
100. NASA Along-Track Near Real-Time Sea Surface Height Anomaly Data , URL: <http://sealevel.jpl.nasa.gov/Science/datasources/ssha>
101. NASA Aqua satellite. URL: http://podaac.jpl.nasa.gov/dataset/JPL-L2P-MODIS_A?
102. Nerger L., Hiller W. Software for Ensemble-based Data Assimilation Systems - Implementation Strategies and Scalability // *Computers and Geosciences*. 2013. V. 55. P. 110–118.
103. Nerger, L., Hiller, W., Schröter, J. PDAF - The Parallel Data Assimilation Framework: Experiences with Kalman Filtering // *Use of high performance computing in meteorology : proceedings of the Eleventh ECMWF Workshop on the Use of High Performance Computing in Meteorology*. Eds.: Walter Zwiefelhofer; George Mozdzyński, Singapore: World Scientific. 2005. P. 63–83.

104. Nieplocha J., Palmer B., Tipparaju V., Krishnan M., Trease H., Apra E. Advances, applications and performance of the global arrays shared memory programming toolkit. *IJHPCA*. V. 20(2). 2006. P. 203–231.
105. Oke P.R., Brassington G.B., Griffin D.A., Schiller A. Ocean data assimilation: a case for ensemble optimal interpolation // *AMOJ*. 2010. V. 59. P. 67–76.
106. Oke P.R., Larnicol G., Fujii Y., Smith G.C., Lea D.J., Guinehut S., Remy E., Alonso Balmaseda M., Rykova T., Surcel-Colan D., Martin M.J., Sellar A.A., Mulet S., Turpin V.. Assessing the impact of observations on ocean forecasts and reanalyses: Part 1, Global studies // *J. Of Operational Oceanography*. 2015. V. 8 . Sup 1. P.49–62.
107. Oke P.R., Sakov P.V., Corney S.P. Impacts of localisation in the EnKF and EnOI: experiments with a small model // *Ocean Dynamics*. 2007. V. 57. P. 32–45.
108. Parallel Ocean Program (POP) User Guide, Version 2.0, Los Alamos National Laboratory, LACC 99-18. 2003. 78 p.
109. Park S.K., Liang Xu. Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications. Edited by Seon K. Park and Liang Xu. Berlin, Heidelberg: Springer Berlin Heidelberg. 2009.
110. Park S.K., Liang Xu. Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II). Edited by Seon Ki Park and Liang Xu. Vol. II. Berlin, Heidelberg: Springer Berlin Heidelberg. 2013.
111. Penny S.G., Behringer D.W., Carton J.A., Kalnay E. A Hybrid Global Ocean Data Assimilation System at NCEP // *Mon. Wea. Rev.* 2015. V. 143. P. 4660–4677.
112. Pringle D.J., Eicken H., Trodahl H.J., Backstrom L.G.E. Thermal conductivity of landfast Antarctic and Arctic sea ice // *J. Geophys. Res.* 2007. V. 112.

113. Ray R.D., Luthcke S.B., Boy J.P. Qualitative comparisons of global ocean tide models by analysis of intersatellite ranging data // *J. Geophys. Res.* 2009. V. 114. C09.
114. Rayner N.A., Brohan P., Parker D.E., Folland C.K., Kennedy J.J., Vanicek M., Ansell T.J., Tett B. Improved analyses of changes and uncertainties in sea surface temperature measured in situ since the mid-nineteenth century: The HadSST2 dataset. // *J. Climate.* 2006. V. 19. P. 446–469.
115. Rodriguez E., Pollard B.D. Centimetric Sea Surface Height Accuracy Using the Wide-Swath Ocean Altimeter // *Proceedings of IGARSS 2003.* Toulouse. 2003. P. 21–25.
116. Sakawa K. Optimal filtering in linear distributed parameter system // *Int. J. Contr.* 1972. V. 16. No 1. P. 115–127.
117. Sakov P., Counillon F., Bertino L. et al. TOPAZ4: An Ocean-Sea Ice Data Assimilation System For The North Atlantic And Arctic // *Ocean Sci.* 2012. V. 8. P. 633–656.
118. Sakov P., Sandery P.A. Comparison of EnOI and EnKF regional ocean reanalysis systems // *Ocean Modelling.* 2015. V. 89. P. 45–60.
119. Sarkisyan A.S., Ibrayev R.A., Iakovlev N.G. High resolution and four-dimensional analysis as a prospect for ocean modeling // *Russ. J. Numer. Anal. Math. Modelling.* 2010. V. 25. P. 477–496.
120. Sasaki YK. A fundamental study of the numerical prediction based on the variational principle. // *J. Meteorol. Soc. Jpn.* 1955. V. 33. P. 262–275.
121. Schiller A., Brassington G.B (Eds.). *Operational oceanography in the 21st century.* Springer Dordrecht: The Netherlands. 2011. 745 p.
122. Schrum C., Backhaus J.O. Sensitivity of atmosphere-ocean heat exchange and heat content in the North Sea and the Baltic Sea // *Tellus - Ser. A Dyn. Meteorol. Oceanogr.* 1999. V. 51(4). P. 526–549.

123. Shlyayeva A.V., Tolstykh M.A., Mizyak V.G., Rogutov V.S. Local ensemble transform Kalman filter data assimilation system for the global semi-Lagrangian atmospheric model // *Russ. J. Num. An. & Math. Mod.* 2013. V. 28. No. 4. P. 419 – 441.
124. Sluka T.C., Penny S.G., Kalnay E., Miyoshi T. Assimilating atmospheric observations into the ocean using strongly coupled ensemble data assimilation // *Geophys. Res. Lett.* 2016. V. 43, P. 752–759.
125. Smith R.D., Maltrud M.E., Bryan F.O., Hecht M.W. Numerical Simulation of the North Atlantic Ocean at $(1/10)^\circ$ // *J. Phys. Oceanography.* 2000. V. 30. P. 1532–1561.
126. Smith S., Cummings J.A., Rowley C. Validation Test Report for the Navy Coupled Ocean Data Assimilation 3D Variational Analysis (NCODA-VAR) System, Version 3.43. 2012.
127. Tanajura C.A.S., Belyaev K. A sequential data assimilation method based on the properties of a diffusion-type process // *Appl. Math. Model.* 2009. V. 33. No. 5. P. 2165–2174.
128. Tanajura C.A.S., Santana A.N., Mignac D., Lima L.N., Belyaev K., Ji-Ping X.. The REMO Ocean Data Assimilation System into HYCOM (RODAS_H): General Description and Preliminary Results // *Atmospheric and Oceanic Science Letters.* 2014. V.7, No. 5. P. 464–470.
129. The standardization forum for the Message Passing Interface (MPI). URL: <http://mpi-forum.org/>
130. Van Leeuwen P.J. Efficient nonlinear data-assimilation in geophysical fluid dynamics // *Computers & Fluids.* 2011. V. 46. P. 52–58.
131. Van Leeuwen P.J. Particle filtering in geophysical systems // *Monthly Weather Rev.* 2009. V. 137. P. 4089–4114.
132. Veronis G. Large scale ocean circulation // *Advances in Applied Mathematics.* 1973. V. 13. P. 2–92.

133. Whitaker J.S., Hamill T.M., Wei X., Song Y., Toth Z. Ensemble Data Assimilation with the NCEP Global Forecast System // *Mon. Wea. Rev.* 2008. V. 136. P. 463–482.
134. Xie, J., Zhu. Ensemble optimal interpolation schemes for Assimilating Argo profiles into a hybrid coordinate ocean model // *Ocean Modell.* 2010. V. 3. P. 283–298.
135. Yang S., Corazza M., Carrassi A., Kalnay E., Miyoshi T. Comparison of Local Ensemble Transform Kalman Filter, 3DVAR, and 4DVAR in a Quasigeostrophic Model // *Mon. Wea. Rev.* 2009.V. 137. P. 693–709.