

## Параллельная реализация локального ансамблевого фильтра Калмана для усвоения атмосферных данных

© В.Г. Мизяк, А.В. Шляева, М.А. Толстых

ФГБУ «Гидрометцентр России», Москва, 123242, Россия

*Одним из основных путей уменьшения ошибок численного прогноза погоды является улучшение оценки состояния атмосферы, используемой в качестве начальных данных для прогностической модели. Вычислительная сложность алгоритмов усвоения данных диктует необходимость их эффективной параллельной реализации. В работе описана параллельная реализация алгоритма усвоения данных с помощью локального ансамблевого фильтра Калмана с преобразованием ансамбля. Приведены результаты ускорения и эффективности параллельной реализации. Показано, что неравномерное распределение вычислительной нагрузки между процессами ограничивает эффективность параллельного алгоритма.*

*Работа выполнена при поддержке грантов в форме субсидий Минобрнауки России по соглашениям 8350 и 8344 от 17.08.2012 г.*

**Ключевые слова:** параллельный алгоритм, ансамблевые фильтры Калмана, усвоение данных.

**Введение.** На протяжении многих лет численный прогноз погоды остается одной из тех задач, решение которой требует огромных вычислительных ресурсов.

Система численного прогноза погоды включает в себя два основных компонента:

1) модель атмосферы (состоит из динамического блока, решающего систему дифференциальных уравнений в частных производных, описывающую процессы гидротермодинамики, явно разрешаемые при заданном уровне детализации, и параметризаций процессов подсеточного масштаба);

2) система усвоения данных наблюдений (подготовка начальных условий для старта модели атмосферы).

Одним из основных путей уменьшения ошибок численного прогноза погоды является улучшение оценки состояния атмосферы, используемой в качестве начальных данных для прогностической модели. В качестве исходных для системы усвоения данных выступают данные наблюдений, например, наземных станций, радиозондов, спутниковые наблюдения ( $y^0 \in R^m$ ,  $m \sim 10^5$  без использования спутниковых наблюдений,  $m \sim 10^7$  с использованием спутниковых наблюдений); а также первое приближение, обычно краткосрочный прогноз модели на заданный срок ( $x^b \in R^n$ , в современных глобальных

моделях атмосферы используется порядка  $10^5 \dots 10^6$  узлов по горизонтали при разрешении  $0,2 \dots 1^\circ$  и  $10^2$  вертикальных уровней, т. е.  $n \sim 10^7 \dots 10^8$ ). Оба источника данных содержат априори неизвестную ошибку. Задача усвоения данных заключается в получении поля анализа — наилучшей оценки состояния атмосферы в заданный момент времени.

Для глобальных моделей атмосферы, а также используемых для них наблюдений можно ограничиться стандартным на сегодняшний день в области усвоения данных предположением о гауссовости, несмещенности и взаимной независимости ошибок наблюдений и первого приближения. В этом случае ошибки первого приближения и наблюдений можно описать их матрицами ковариаций:  $P^b$  — матрица ковариаций ошибок первого приближения,  $R$  — матрица ковариаций ошибок наблюдений.

Решением задачи о нахождении оптимальной оценки состояния атмосферы является вектор анализа  $x^a \in R^n$ :

$$x^a = x^b + K \left[ y^o - H(x^b) \right], \quad (1)$$

где  $K$  — весовая матрица;  $H$  — оператор наблюдений, который проектирует вектор из пространства модели в пространство наблюдений.

Предполагается, что существует линеаризация последнего оператора, т. е. матрица  $H$  размерностью  $m \times n$ , такая, что  $H(x + \delta x) - H(x) \approx H(\delta x)$ . Далее везде в качестве оператора наблюдений рассматривается линейный оператор.

При сделанных выше предположениях оптимальная в смысле метода наименьших квадратов матрица  $K$  может быть вычислена по формуле

$$K = P^b H^T (R + H P^b H^T)^{-1}. \quad (2)$$

Как правило, полагают, что ошибки различных наблюдений не коррелированы и матрица  $R$  диагональна. Наибольшую сложность представляет построение матрицы ковариаций ошибок первого приближения.

В настоящее время широко распространены две группы методов усвоения данных: вариационное усвоение и усвоение с помощью ансамблевых фильтров Калмана. Кроме того, активно развивается и гибридное усвоение, учитывающее достоинства обеих групп методов.

**Усвоение данных с помощью ансамблевых фильтров Калмана.** Ансамблевые фильтры Калмана были предложены Эвенсеном [1]. Они представляют собой синтез идей метода Монте-Карло и фильтра Калмана — циклического алгоритма поиска оптимальной оценки состояния линейной системы.

На первом этапе (шаг прогноза) моделируется состояние системы ( $x^b$  в наших обозначениях) и матрица ковариаций  $P^b$ . На втором этапе (шаг анализа) спрогнозированное состояние корректируется с учетом имеющихся на момент анализа наблюдений и матрицы ковариаций  $R$ . Такой подход хорошо зарекомендовал себя в реальных системах, имеющих не очень высокие порядки. Однако применение классического фильтра Калмана при подготовке начальных данных для современных численных моделей атмосферы практически невозможно по двум причинам:

1) порядок вектора состояния может достигать значения  $10^8$ , вычисление матрицы ковариаций  $P^b$  в этом случае требует огромных вычислительных затрат;

2) система уравнений и оператор наблюдений существенно нелинейные, а значит, полученное решение может оказаться неоптимальным.

Идея ансамблевых фильтров Калмана заключается в статистической оценке матрицы ковариаций  $P^b$  по ансамблю первых приближений, размерность которого значительно меньше размерности модели (десятки или сотни). При таком подходе матрица ковариаций  $P^b$  имеет структуру, которая обусловлена реальными свойствами процессов, происходящих в атмосфере. Кроме того, ансамблевая фильтрация достаточно проста (в сравнении с вариационными методами), хотя предполагает дополнительные вычислительные затраты для расчета большего числа прогнозов.

При усвоении с помощью ансамблевых фильтров Калмана в цикле усвоения (анализ-прогноз) участвует не один вектор первого приближения, а ансамбль первых приближений  $x^b = \{x^{b(1)}, x^{b(2)}, \dots, x^{b(k)}\}$ , где  $k$  — размерность ансамбля (обычно 40...200). С помощью такого ансамбля производится оценка матрицы ковариаций ошибок первого приближения:

$$P^b = (k - 1)^{-1} X^b (X^b)^T, \quad (3)$$

где  $X^b = \{x^{b(1)} - \bar{x}^b, x^{b(2)} - \bar{x}^b, \dots, x^{b(k)} - \bar{x}^b\}$ ,  $\bar{x}^b = \frac{1}{k} \sum_{i=1}^k x^{b(i)}$  — среднее по ансамблю.

Матрица ковариаций ошибок анализа может быть вычислена следующим образом:

$$P^a = (I - KH) P^b = P^b - P^b H^T (R + HP^b H^T)^{-1} HP^b, \quad (4)$$

где  $I$  — единичная матрица размерностью  $n \times n$ .

**Локальный ансамблевый фильтр Калмана с преобразованием ансамбля.** В данной работе обсуждается разработанная параллельная реализация схемы усвоения на базе локального ансамблевого фильтра Калмана с преобразованием ансамбля (Local Ensemble Transform Kalman Filter, LETKF) [3].

Локальные ансамблевые фильтры используют локализацию ковариаций ошибок наблюдений, при которой учитываются только те наблюдения, которые расположены в определенной области вокруг точки сетки, при этом функция ковариации зависит от расстояния между точкой наблюдения и обрабатываемой точкой. В данной работе используется гладкая функция ковариации пятого порядка Гаспари — Кона [4].

Одним из достоинств локализации является возможность эффективного распараллеливания алгоритма усвоения по данным, так как вычисление анализа можно проводить параллельно в разных точках сетки.

Основная идея LETKF (и ETKF, [5]) — проводить вычисления не в физическом пространстве модели, а в пространстве ансамбля (меньшей размерности  $k$ ).

Пусть

$$x = \bar{x}^b + X^b w, \quad (5)$$

где  $w$  — некоторый вектор из пространства ансамбля,  $X^b$  — оператор проецирования из пространства ансамбля в модельное пространство.

Тогда, если  $w$  — случайный вектор с нормальным законом распределения, нулевым математическим ожиданием и матрицей ковариаций  $(k-1)^{-1}I$ , то  $x$  — случайный вектор с математическим ожиданием  $\bar{x}^b$  и матрицей ковариаций  $P^b = (k-1)^{-1}X^b(X^b)^T$ .

Все основные расчеты проводятся в пространстве ансамбля, в котором вычисляется анализ  $\bar{w}^a$ , после чего в пространстве модели — вектор анализа

$$\bar{x}^a = \bar{x}^b + X^b \bar{w}^a. \quad (6)$$

Ниже приведен алгоритм LETKF:

1) вычислить ансамбль (глобальных) первых приближений в точках наблюдений  $y_g^b = H(x_g^b)$ , среднее по ансамблю  $\bar{y}_g^b$  и матрицу отклонений от среднего  $Y_g^b$ . Здесь нижний индекс  $g$  обозначает, что соответствующие векторы определены во всем пространстве;

2) вычислить среднее по ансамблю первых приближений  $\bar{x}_g^{-b}$  и матрицу отклонений от среднего  $X_g^b$ ;

3) выбрать из  $y_g^o$  локальные наблюдения  $y^o$ , находящиеся в круге заданного радиуса, центр которого — в рассматриваемой точке сетки;

4) вычислить матрицу  $C = (Y^b)^T R^{-1}$ ;

5) вычислить матрицу  $\tilde{P}^a = ((k-1)I / \rho + CY^b)^{-1}$ ,  $\rho$  — множитель инфляции ансамбля;

6) вычислить матрицу  $W^a = [(k-1)\tilde{P}^a]^{1/2}$ ;

7) вычислить вектор среднего анализа в пространстве ансамбля  $\bar{w}^a = \tilde{P}^a C (y^o - \bar{y}^b)$  и добавить к каждому столбцу матрицы  $W^a$  (получим матрицу  $w^a$ );

8) вычислить векторы анализов  $x^a = \bar{x}^b + X^b w^a$ .

Один из опытов реализации схемы усвоения с помощью LETKF описан в статье [6]. В этой работе приводится описание реализации, в которой применялось гибридное распараллеливание с использованием библиотек MPI и OpenMP. Вся расчетная область разбивалась на горизонтальные широтные полосы для распараллеливания с помощью MPI. Внутри каждой локальной области циклы по вертикальной координате были распараллелены с помощью OpenMP. Для различных сочетаний количества членов ансамбля и количества процессоров была достигнута параллельная эффективность, близкая к 100 % [6].

Как показали исследования [7], для эффективного выполнения счета задачи на всех доступных узлах вычислительной системы необходим алгоритм сбалансированной нагрузки узлов. Простое деление расчетной области на равномерные широтно-долготные квадраты при увеличении числа доступных процессоров не дает ожидаемого выигрыша. Происходит это ввиду возрастающего количества пересылок и различного количества обрабатываемых данных: процессоры, которые обрабатывают хорошо наблюдаемые регионы, тратят на расчет гораздо больше времени, чем те, что обрабатывают регионы с малым количеством наблюдений.

Для оптимального с этой точки зрения разбиения расчетной области на заданное количество регионов авторами [7] была применена бисекция. Первоначально авторы изучили распределение наблюдений по локальным областям. Далее, учитывая это распределение, в Северном и Южном полушариях были найдены по два таких разбиения, при которых наблюдалась наименьшая разница между временами, затраченными на обработку каждым из двух процессоров (отдельно в Северном и отдельно в Южном полушариях). После этого регион, в котором время счета было больше, делился на две меньшие

области. Процесс продолжался до тех пор, пока не становилось задействованным требуемое число процессоров. Этот и другие аспекты данной реализации более подробно можно посмотреть в [7].

**Параллельная версия алгоритма LETKF.** Поскольку в схеме усвоения с помощью локального ансамблевого фильтра Калмана вычисления анализа в каждой точке сетки не зависят от вычислений в других точках сетки, алгоритм хорошо распараллеливается по данным. В нашей реализации данные (узлы регулярной широтно-долготной сетки) распределяются между вычислительными процессами равномерно с точки зрения количества обрабатываемых узлов по широтным полосам.

Перед вычислением анализа происходит пересылка данных между процессами.

Далее для каждого узла сетки выполняются шаги 3–8 алгоритма: находятся локальные наблюдения, вычисляется матрица ковариаций ошибок наблюдений, вычисляется весовая матрица  $W^a$ , или матрица анализа в пространстве ансамбля, и находится вектор анализа в данной точке в пространстве модели.

После вычисления значений анализа в каждой точке сетки выполняется сбор данных от всех вычислительных процессов для получения полных полей анализа.

**Результаты экспериментов.** Разработанная программная реализация ансамблевого фильтра Калмана была протестирована на задаче усвоения атмосферных данных для модели численного прогноза погоды ПЛ-АВ [8]. ПЛ-АВ — вычислительно эффективная полулагранжева глобальная конечно-разностная модель общей циркуляции атмосферы. В модели применяется набор параметризаций процессов подсеточного масштаба из модели ALADIN/ALARO. Программный комплекс модели был успешно реализован на параллельных вычислительных системах с общей и распределенной памятью с применением технологий распараллеливания OpenMP и MPI. Модель является основной глобальной оперативной моделью в Гидрометцентре России.

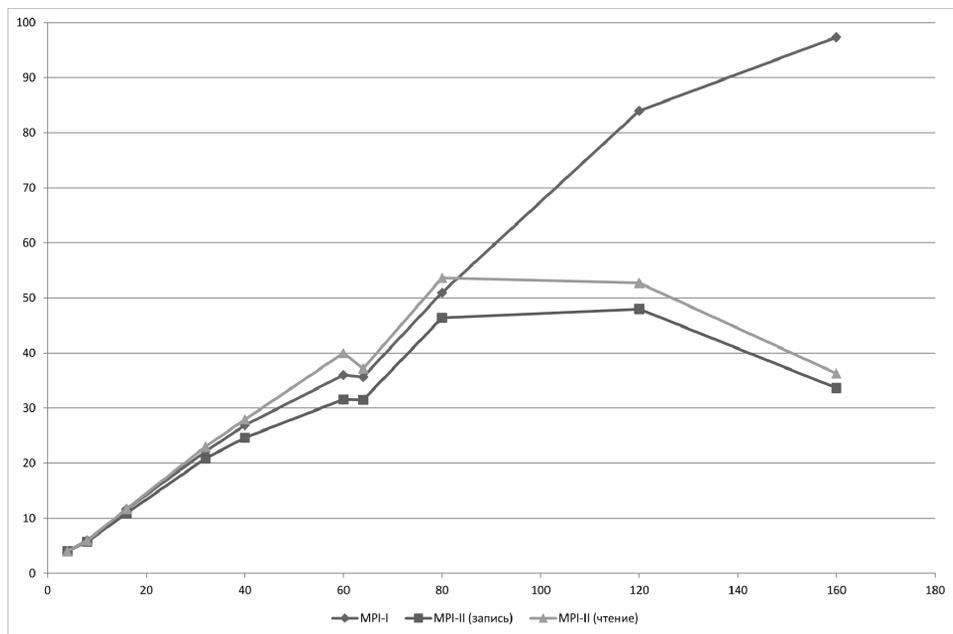
Для тестирования схемы усвоения данных использовалась версия модели ПЛАВ с горизонтальным разрешением  $1,4 \times 1,125^\circ$  ( $256 \times 161$  узел сетки), 28 уровней по вертикали. Радиус горизонтальной локализации — 1500 км, использовались наблюдения наземных станций, радиозондов и спутниковых наблюдений SATOB (общее число наблюдений — около 25 000).

Были выполнены тестовые запуски прототипа LETKF с различным числом участников ансамбля (40 и 60) на вычислительной системе SGI Altix 4700, которая установлена в ГВЦ Росгидромета. SGI Altix 4700 — система, состоящая из 13 узлов, каждый из которых

имеет 128 ядер (1664 процессорных ядра Intel Itanium 2 9140M, оперативная память 6,6 ТВ) пиковой производительностью 11 Тфлопс.

Расчеты производились для следующих значений числа процессов: 4, 8, 16, 32 и 64 (степени двойки), 40, 80, 160 (распределение примерно равного числа широтных полос между процессами).

График параллельного ускорения расчетов представлен на рис. 1.



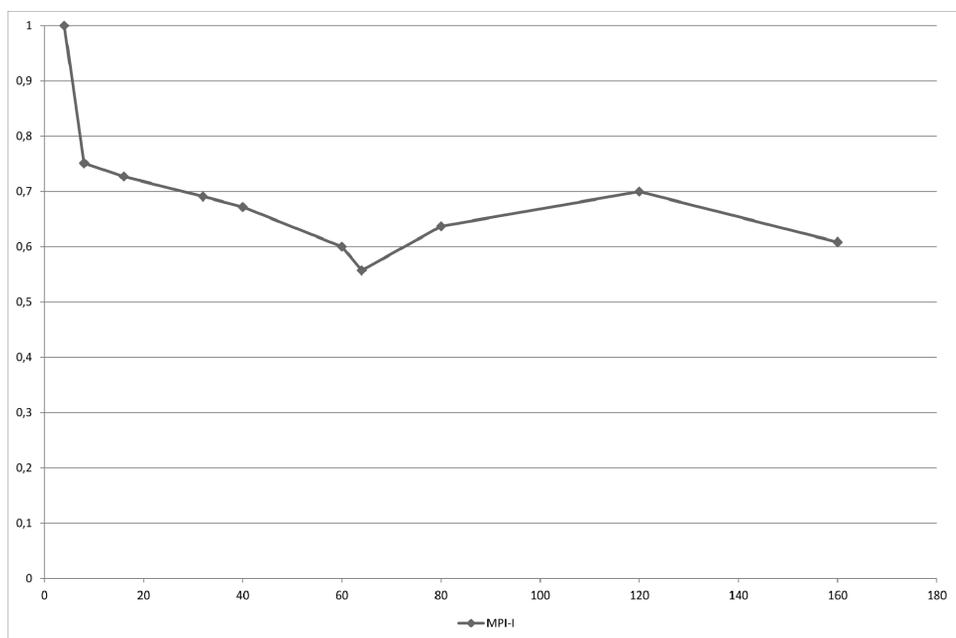
**Рис. 1.** Ускорение при использовании 60 членов ансамбля

Из графиков видно, что рост ускорения близок к линейному для 60 участников ансамбля. На рис. 2 приведена эффективность параллельной реализации при использовании MPI-I для 60 участников ансамбля.

Видно, что эффективность уменьшается с увеличением числа процессов. Это объясняется тем, что неравномерностью распределения вычислений между процессами. С точки зрения количества обрабатываемых узлов сетки данные распределяются равномерно, при этом вычислительная сложность задачи зависит от числа обрабатываемых локальных (расположенных на расстоянии не больше заданного) наблюдений. Наблюдения распределены неравномерно, кроме того, широтные полосы в областях, расположенных у полюсов, покрывают меньшую площадь, чем широтные полосы в областях, ближних к экватору.

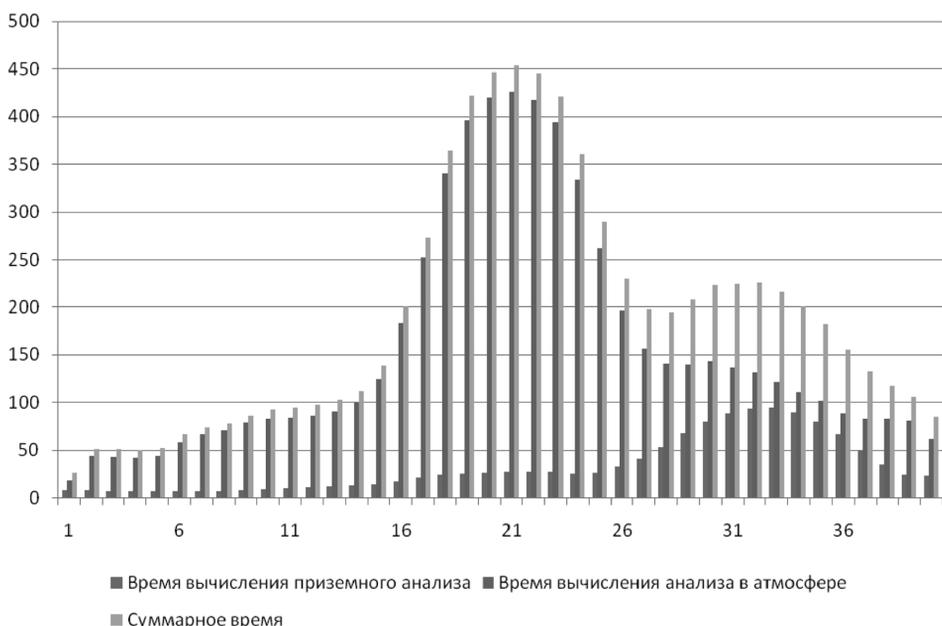
Таким образом, большая часть вычислительных процессов ожидает вычисления анализа процессами, на которые за счет большей площади и числа наблюдений приходится большая нагрузка. Распреде-

ние вычислений между процессами (для 40 из них) показано на рис. 3. Видно, что для вычисления приземного анализа существенную роль играет расположение наблюдений (процессы с 21 по 40 обрабатывают широты Северного полушария, в котором находится больше приземных наблюдений). На время вычисления анализа в свободной атмосфере оказывает заметное влияние различная площадь широтных полос в приполюсных и приэкваториальных областях. Было опробовано распределение широтных полос между процессами, равномерное по косинусу широты. Распределение вычислений показано на рис. 4. При его использовании время счета сократилось на 20 %.

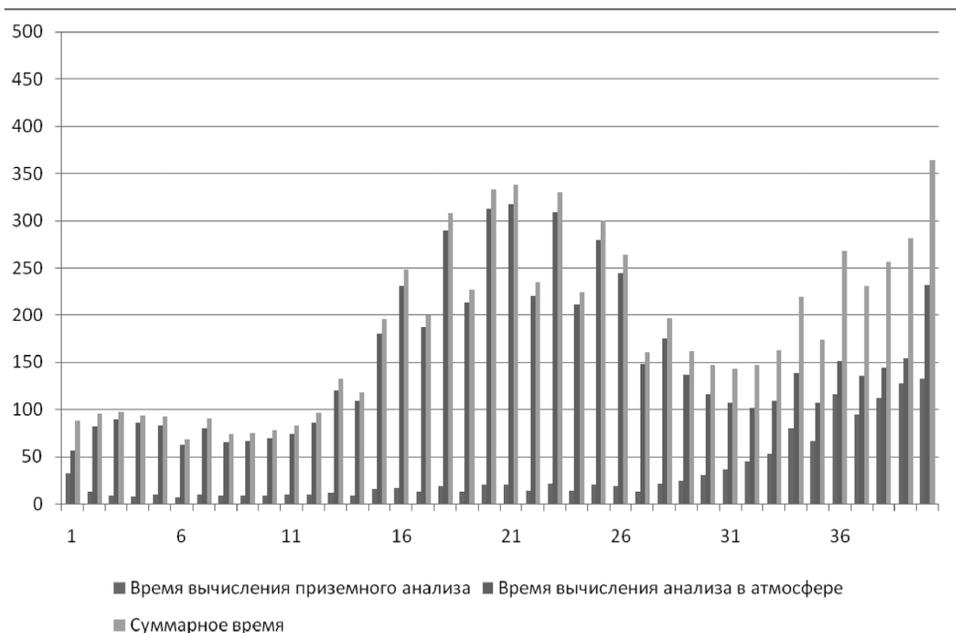


**Рис. 2.** Эффективность при использовании 60 членов ансамбля

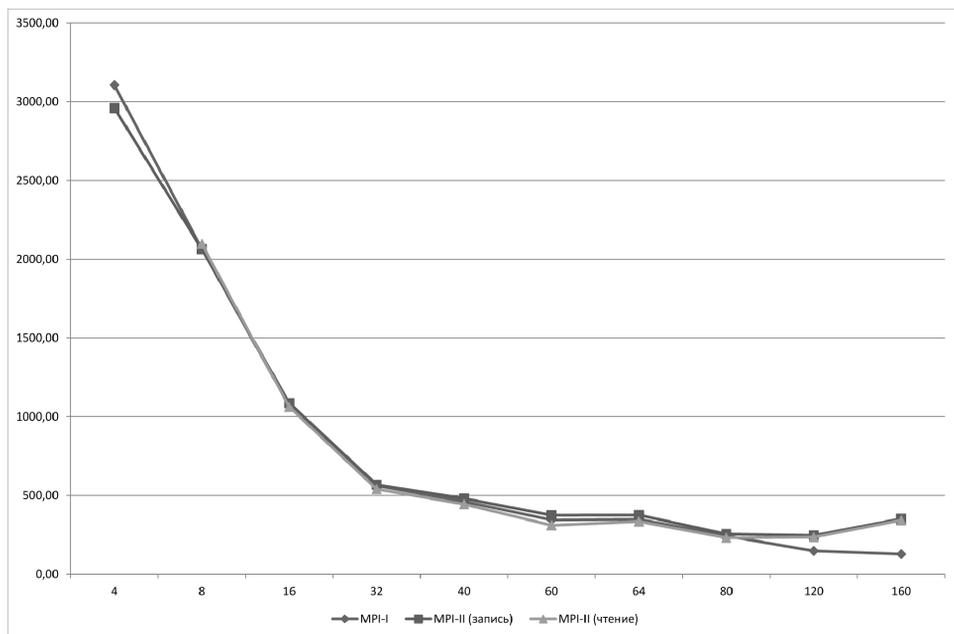
Известно, что использование односторонних коммуникаций, например, из стандарта SHMEM, может ускорить работу параллельной программы (например, [8]), поэтому в данной работе сравнивается также время работы программы при использовании двухсторонних коммуникаций MPI-I и односторонних коммуникаций из стандарта MPI-II (в двух вариантах — использование односторонней записи (MPI\_PUT) и одностороннего чтения (MPI\_GET)). Вариант передачи сообщений с помощью функционала библиотеки SHMEM признан нами неперспективным в связи с ожидаемым переходом на другую архитектуру. На рис. 5 приведены графики времени выполнения программы при использовании 40 членов ансамбля.



**Рис. 3.** Распределение вычислений (по времени) между процессами при распределении данных равномерно по широте



**Рис. 4.** Распределение вычислений (по времени) между процессами при распределении данных равномерно по косинусу широты

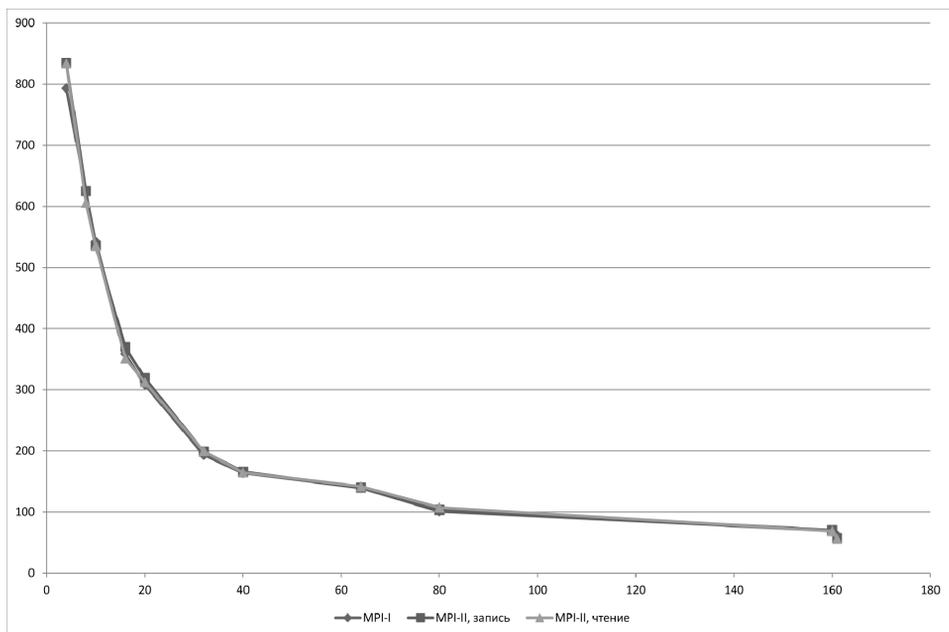


**Рис. 5.** Время выполнения (в мс) на Altix 4700 при использовании 40 членов ансамбля

Однако, как видно из графиков времени выполнения, использование односторонних коммуникаций не дало выигрыша во времени. Возможным объяснением этого факта является ограниченная поддержка стандарта MPI-II в используемой на вычислительной системе SGI Altix 4700 библиотеке MPI.

Также были выполнены расчеты анализа для 40 участников ансамбля на пилотной кластерной системе РСК, установленной в октябре 2011 г. в ГВЦ Росгидромета. «РСК Торнадо» состоит из 96 вычислительных узлов, каждый из которых содержит по два шестиядерных процессора Intel Xeon 5680 с тактовой частотой 3,33 ГГц и 48 Гб оперативной памяти. Коммуникационная сеть построена на базе интерфейса Infiniband QDR. Время выполнения расчетов приведено на рис. 6. Видно, что при малом числе используемых процессов система РСК в 3–4 раза быстрее, чем вычислительная система SGI Altix 4700, однако при большом количестве процессов выигрыш меньше примерно в 2 раза. Это объясняется потерями на ожидание завершения обменов, которые вызваны несовершенством используемого алгоритма распределения данных по процессам.

**Выводы.** Одним из основных путей уменьшения ошибок численного прогноза погоды является улучшение оценки состояния атмосферы, используемой в качестве начальных данных для прогностической модели. Вычислительная сложность алгоритмов усвоения данных диктует необходимость их эффективной параллельной реализации.



**Рис. 6.** Время выполнения (в мс) на «РСК Торнадо» при использовании 40 членов ансамбля

Нами разработан параллельный алгоритм локального ансамблевого фильтра Калмана для усвоения метеорологических данных наблюдений. Тестирование показало, что разработанная реализация позволяет достичь ускорения в 100 раз при распараллеливании на 160 процессорах (размерность ансамбля равна 60). При меньшей размерности задачи (размерность ансамбля — 40 участников) эффективное распараллеливание ограничено 100 вычислительными процессорами. Показано, что неравномерное распределение вычислительной нагрузки между процессорами ограничивает эффективность параллельного алгоритма. В дальнейшем планируется усовершенствование алгоритма распределения данных (в зависимости от плотности наблюдений в областях), а также реализация более эффективного алгоритма выбора локальных наблюдений из множества всех наблюдений. В перспективе предполагается реализовать данную схему усвоения в оперативном режиме.

## ЛИТЕРАТУРА

- [1] Evensen G. Sequential data assimilation with a nonlinear quasigeostrophic model using Monte-Carlo methods to forecast error statistic. *J. Geophys. Res.*, 1994, vol. 99, pp. 10143–10162.
- [2] Kalman R.E. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 1960, vol. 82, pp. 35–45.

- [3] Hunt B.R., Kostelich E.J., Szunyogh I. Efficient data assimilation for spatio-temporal chaos: a Local Ensemble Transform Kalman Filter. *Physica D*, 2007, vol. 230, pp. 112–126.
- [4] Gaspari G., Cohn S.E. Construction of correlation functions in two and three dimensions. *Quarterly Journal of Royal Meteorological Society*, 1999, vol. 125, pp. 723–757.
- [5] Bishop C.H., Etherton B.J., Majumdar S.J. Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Monthly Weather Review*, 2001, vol. 129, pp. 420–436.
- [6] Miyoshi T., Yamane S. Local Ensemble Transform Kalman Filtering with an AGCM at a T159/L48 Resolution. *Monthly Weather Review*, 2007, vol. 135, N 11. pp. 3841–3861.
- [7] Szunyogh I. et al. *A local ensemble transform Kalman filter data assimilation system for the NCEP global model*. Tellus A, 2007.
- [8] Толстых М.А. *Глобальная полулагранжева модель численного прогноза погоды*. Обнинск, ОАО ФОР, 2010, 111 с.
- [9] Корж А.А. SHMEM+ на суперкомпьютерах BLUE GENE/P и «Ломоносов». *Вычислительные методы и программирование*, 2010, с. 123–129.

Статья поступила в редакцию 10.06.2013

Ссылку на эту статью просим оформлять следующим образом:

Мизяк В.Г., Шляева А.В., Толстых М.А. Параллельная реализация локального ансамблевого фильтра Калмана для усвоения атмосферных данных. *Инженерный журнал: наука и инновации*, 2013, вып. 6. URL: <http://engjournal.ru/catalog/it/hidden/1102.html>

**Мизяк Василий Геннадьевич** окончил факультет «Фундаментальные науки» МГТУ им. Н.Э. Баумана в 2007 г. Программист отдела среднесрочных прогнозов погоды ФГБУ «Гидрометцентр России». Автор 3 статей в области параллельных вычислений и усвоения данных. e-mail: [vmizyak@gmail.com](mailto:vmizyak@gmail.com)

**Шляева Анна Викторовна** родилась в 1984 г., окончила кафедру «Программное обеспечение ЭВМ и информационные технологии» МГТУ им. Н.Э. Баумана в 2007 г. Канд. техн. наук, доцент кафедры «Программное обеспечение ЭВМ и информационные технологии», старший научный сотрудник ФГБУ «Гидрометцентр России». Автор 15 научных работ. Области научных интересов: математическое моделирование, статистический анализ данных, параллельные вычисления. e-mail: [shlyaeva@gmail.com](mailto:shlyaeva@gmail.com)

**Толстых Михаил Андреевич** родился в 1963 г., окончил Московский физико-технический институт в 1986 г. Д-р физ.-мат. наук, ведущий научный сотрудник ФГБУН Институт вычислительной математики РАН и заведующий лабораторией ФГБУ «Гидрометцентр России». Автор 68 научных работ. Области научных интересов: вычислительные методы в геофизической гидродинамике, численный прогноз погоды, параллельные вычисления. e-mail: [tolstykh@inn.ras.ru](mailto:tolstykh@inn.ras.ru)