

УСВОЕНИЕ ДАННЫХ: КОНКУРЕНЦИЯ МЕТОДОВ И ПРОБЛЕМА УСВОЕНИЯ СПУТНИКОВЫХ НАБЛЮДЕНИЙ

Т.В. Красюк

*Гидрометеорологический научно-исследовательский центр Российской Федерации
krasjuk@mail.ru*

Введение

Постоянное увеличение массивов наблюдаемой геофизической информации, развитие систем спутниковых (в том числе радиометрических) наблюдений и измерений, высокоскоростных средств связи требует более интенсивного развития методов усвоения данных наблюдений.

В целом современные методы усвоения данных можно разделить на два конкурирующих класса. Первый класс основан на вариационных принципах поиска минимума некоторого функционала (постановка задачи предложена в [7]), описывающего отклонение модельного решения и данных измерений в некоторой заданной метрике как функции начальных (и/или граничных) условий, и получил название «вариационный метод» решения задач усвоения. Четырехмерный вариационный метод базируется на теории возмущений сопряженных уравнений, разработанной академиком Г.И. Марчуком для дискретных моделей динамики атмосферы и океана [5, 6], и успешно применяется на практике, хотя требует для своей реализации значительных вычислительных ресурсов, а также создания и поддержки сопряженной модели.

Алгоритм трехмерного вариационного усвоения 3D-Var впервые был реализован в Национальном центре прогнозов окружающей среды (NCEP) в 1992 году [42]. Затем в 1996 г. в оперативном режиме 3D-Var был реализован в Европейском центре среднесрочных прогнозов погоды (ЕЦСПП). Схема 4D-Var одной из первых была доведена до практического применения в прогностической системе ЕЦСПП в 1997 году [44].

В последнее время в практику стал широко внедряться другой – динамико-стохастический подход, основанный на использовании алгоритма фильтра Калмана [32]. Физический смысл этого подхода следующий: определение связи между известным (наблюдаемым) и неизвестным состоянием атмосферы во времени и/или пространстве, которые в линейном приближении задаются через математическое ожидание и ковариационные функции. Этот подход значительно проще реализуем, чем вариационный. Корректное применение данного метода возможно только для линейных систем [55], в

случае же сильной нелинейности модели и/или негауссовского шума возможно получить оценки неизвестной ковариационной функции методом Монте-Карло и затем решить соответствующее уравнение линейного оптимального фильтра. Один из методов, адаптированных для нелинейных систем, получил название расширенный фильтр Калмана (extended Kalman filter) [16]: нелинейные уравнения, описывающие динамику системы, линеаризуются относительно предыдущего состояния системы, итеративный цикл позволяет найти необходимую оценку состояния динамической системы (эта оценка не обязательно будет оптимальной).

Одно из главных преимуществ Калмановского фильтра для нахождения полей анализа по сравнению с алгоритмом вариационного усвоения данных состоит в том, что этот метод позволяет явно учесть ковариации ошибок прогноза и не накладывает ограничений (как и 4D-Var) на интервал времени, в котором происходит усвоение пришедших ранее данных наблюдений (окно усвоения). К достоинствам калмановского алгоритма следует также отнести его рекуррентную природу, эффективно проявляющуюся при работе в реальном времени, а также возможность априорной оценки точности получаемых результатов средствами самого алгоритма.

1. Задача усвоения и фильтры Калмана

Рассмотрим постановку задачи усвоения данных в общепринятых обозначениях.

x_a – вектор переменных искомого анализа; x_b – вектор переменных модельного состояния атмосферы первого приближения (краткосрочный прогноз, стартовавший с предыдущего шага усвоения): $x_j^b = M(x_{j-1}^a) + \varepsilon_m$, где M – оператор модели; j – шаг усвоения; ε_m – ошибка модели.

Обозначим истинное состояние атмосферы в точках модельной сетки x^t , тогда

$$x^b = x^t + \varepsilon_b, \quad (1)$$

$$y^o = H(x^t) + \varepsilon_r. \quad (2)$$

Здесь y^o – вектор наблюдений; H – оператор наблюдений (в общем случае нелинейный); ε_b – вектор ошибок первого приближения; ε_r – вектор ошибок наблюдений.

Вектор анализа x^a также содержит некоторую ошибку:

$$x^a = x^t + \varepsilon_a. \quad (3)$$

Задача усвоения данных состоит в том, чтобы найти анализ, минимизирующий ошибку

ε_a .

$R = \langle \varepsilon_r \varepsilon_r^T \rangle$ – положительно определенная ковариационная матрица ошибок наблюдений; $B = \langle \varepsilon_b \varepsilon_b^T \rangle$ – ковариационная матрица ошибок первого приближения (квадратная, положительно определенная); $Q = \langle \varepsilon_m \varepsilon_m^T \rangle$ – ковариационная матрица ошибок модели.

Во многих случаях можно предположить, что ошибки первого приближения, наблюдений, а также модели [52] независимы. Кроме того, в исследованиях по усвоению данных обычно считается, что:

- ошибки наблюдений и первого приближения имеют гауссово распределение;
- средние ошибки наблюдений и первого приближения равны нулю.

Фильтр Калмана представляет собой последовательный алгоритм, в котором ищется решение с учетом данных наблюдений (в момент анализа), прогноза на интервале между предыдущим и текущим сроком анализа и ковариациями полей первого приближения. Уравнения прогностической модели используются для экстраполяции во времени ковариаций ошибок анализа, которые включают ковариации ошибок самой прогностической модели. Поскольку классический фильтр Калмана, как указано выше, сформулирован для линейных систем, то решением задачи является¹:
$$x^a = x^b + K[y^o - H(x^b)].$$

При этом оптимальная весовая матрица $K = BH^T(R + HBH^T)^{-1}$.

В данном случае H – либо изначально линейный оператор, либо линеаризация исходного оператора наблюдений.

2. Ансамблевые фильтры Калмана

В схемах усвоения данных, основанных на ансамблевом подходе, ансамбль полей первого приближения генерируется из ансамбля начальных условий, распределенных согласно результатам предыдущего анализа. Преимуществом данного подхода перед четырехмерным вариационным усвоением и расширенным фильтром Калмана является то, что нет необходимости в трудоемкой разработке и поддержании линеаризованной прогностической модели. Кроме того, использование данного подхода позволяет с минимальными затратами провести распараллеливание вычислений на многопроцессорной вычислительной технике.

Различают два основных класса ансамблевых фильтров: стохастические и

¹ в постановке BLUE (Best Linear Unbiased Estimate)

детерминированные. Основное различие между существующими схемами, основанными на ансамблевом подходе, состоит в генерации ансамбля анализов. Одно семейство схем (стохастические ансамблевые фильтры Калмана) основано на возмущении наблюдений [22, 25–28]. Различные наборы данных наблюдений получаются при помощи добавления случайного шума к реальным данным наблюдений, где компоненты случайного шума генерируются согласно ковариационной матрице ошибок наблюдений. Одна из таких стохастических схем [27] является оперативной схемой усвоения для ансамблевого прогноза в метеослужбе Канады.

Другое семейство схем – детерминированные ансамблевые фильтры Калмана, в том числе, Калмановские фильтры квадратного корня [49] (Kalman square-root Filter). При данном подходе анализ проводится только один раз для получения среднего анализа и ковариационной матрицы ошибок анализа, а затем генерируются ансамблевые возмущения анализа (для среднего анализа) с использованием ковариационной матрицы ошибок анализа. Поскольку существует бесконечное множество методов задания возмущений по матрице ковариаций ошибок анализа, то согласно данному подходу может быть получено много различных схем [49]: ансамблевый трансформированный фильтр Калмана ETKF [12], ансамблевый корректирующий фильтр EAKF [10], ансамблевый фильтр квадратного корня EnSRF [53] (Ensemble Square Root Filter), локальный ансамблевый трансформированный фильтр Калмана LETKF [30] и др.

J.S. Whitaker и T.H. Namill [53] показали, что фильтры квадратного корня более точны, чем фильтры, использующие возмущения наблюдений. Они избегают дополнительной ошибки выборки, которая вводится через возмущения наблюдения случайной ошибкой. Однако W.G. Lawson и J.A. Hansen [35] сделали предположение, что фильтры, использующие возмущения наблюдений, могут лучше учитывать нелинейность, чем фильтры квадратного корня.

Обзор ансамблевых фильтров Калмана приведен в [21] и [40]. Теоретически оба варианта ансамблевых фильтров должны давать близкие результаты, при этом на первый план выходит вычислительная эффективность, ключевым аспектом повышения которой стала локализация.

3. Локализация в ансамблевых фильтрах

При локализации уменьшается размерность пространства, в котором решается задача. Локализация ковариаций ограничивает использование ковариационной информации ансамбля для небольшого подмножества переменных, определенных в точках модельной сетки в пределах некоторого локального региона, намного меньшего, чем все физическое

пространство модели. Локальный подход делает возможным оценку глобальных статистик ковариации ошибок поля первого приближения, основанных на малых ансамблях. Локализация осуществляется за счет вычисления анализа в точках сетки только с учетом наблюдений, находящихся в пределах корреляционного радиуса.

При локализации подавляются мнимые ковариации между далеко расположенными точками. Еще одним достоинством локализации [37, 41] является возможность эффективного распараллеливания вычислений, т.к. анализ можно вычислять параллельно в разных точках сетки. Основным недостатком локализации является то, что она может создавать дисбаланс между различными полями, а также внутри поля (при переходе от точки к точке сетки некоторые наблюдения начинают использоваться, другие, наоборот, перестают).

Среди всех разновидностей ансамблевый фильтр Калмана (АнКФ) наибольшее распространение получил локальный АнКФ с преобразованием ансамбля (Local Ensemble Transform Kalman Filter, LETKF) [31]. Алгоритм LETKF был разработан Оттом [30, 31] и опробован как на имитируемых наблюдениях в идеальной модели [46], так и наблюдениях реальной атмосферы [38,45,54]. Алгоритм LETKF рассмотрен более подробно в [9].

LETKF тестировался в квазиоперативном режиме в Японии, с 2011 г. применяется в оперативном прогнозе в Италии, а в сочетании с 4D-Var (генерация начальных данных) используется для оперативной вероятностной системы прогноза в Англии.

Авторы [30] расширили ансамблевый фильтр Калмана до 4-мерного случая (4-D EnKF), который учитывает дополнительно еще и время. Данный метод позволил усваивать асинхронные данные (время поступления данных наблюдений и время выхода данных первого приближения не совпадает). В методе 4-D EnKF данные наблюдений представляются как линейная комбинация ансамблевых возмущений в момент времени произведенного наблюдения. Подобная линейная комбинация членов ансамбля может быть использована для перемещения наблюдения вперед (или назад) во времени к моменту проведения анализа. Этот простой метод дает ансамблевому фильтру Калмана возможности метода 4D-Var для усвоения наблюдений в момент времени, когда они были сделаны, и позволяет использовать наблюдения из будущего, когда они становятся доступны (например, в задачах реанализа).

Многие проблемы АнКФ так или иначе решены практическим способом, но до сих пор не разработано строгое обоснование этого метода усвоения данных. Нерешенные проблемы (негауссовость ошибок, учет и оценка ошибки модели) во многом те же, что и у 4D-Var. Все варианты АнКФ требуют больших вычислительных ресурсов, как и 4D-Var. Однако реализовать АнКФ проще, чем 4D-Var, при этом АнКФ не зависит от модели (если

рассматривается одна и та же среда и одни и те же наблюдения).

4. Решение некоторых проблем усвоения спутниковых данных

Как правило, современные измерительные спутниковые системы развиваются в направлении увеличения спектрального разрешения. При этом возникает несколько проблем: 1) существенно увеличивается количество доступных данных наблюдений, обработка которых становится весьма затруднительным процессом; 2) с увеличением спектрального разрешения уменьшается энергия полезного сигнала в каждом канале. Сужение измерительных каналов уменьшает отношение сигнал/шум в каждом отдельном канале, что естественно ухудшает возможность восстановления искомым параметров атмосферы.

Методы и проблемы вариационного усвоения данных спутниковых измерений для определения физических параметров атмосферы и океана подробно рассмотрены в исследовании Н.В. Уварова [8].

Общая проблема методов усвоения данных (и не только для спутниковых наблюдений) – негауссовость ошибок, – решается на практике пока только в отдельных частных случаях с помощью замены исследуемых переменных на переменные, распределение ошибок которых близко к гауссовскому. Часто оказывается эффективной нормализация переменной, которая, однако, приводит к дополнительным корреляциям. Частный пример решения проблемы предложили авторы [36] для влажности – новая переменная представляла собой отношение наблюдаемой удельной влажности к насыщенной удельной влажности, значение которой берется из первого приближения прогноза, поэтому корреляции с наблюдаемой температурой (характерные для нормированной обычным образом влажности) отсутствуют.

Поиск решения проблемы негауссовости ошибок в приложении к методам калмановской фильтрации в совокупности с модификацией фильтра для нелинейного оператора наблюдений продолжается (см., например [59]).

Ненулевые средние ошибки наблюдений также создают дополнительные сложности усвоения спутниковых данных. Эта проблема преодолевается с помощью замены переменных или расширения вектора состояния, например за счет вектора смещений (одна из соответствующих схем будет рассмотрена ниже).

К характерным трудностям усвоения спутниковых данных с помощью методов калмановской фильтрации отнесем следующие.

1. Ансамблевые фильтры Калмана требуют длительного времени раскрутки (spin-up), чтобы достичь асимптотического уровня точности по сравнению с вариационными методами (3D-Var или 4D-Var). На практике могут возникать ситуации, когда EnKF

разгоняется слишком медленно, чтобы быть полезным.

Предложенный в [33] эвристический алгоритм позволяет ускорить «разгон» EnKF с помощью скользящего среднего от ансамбля, а также многократного использования наблюдений в каждом окне усвоения, в том числе наблюдений в последующие моменты времени. Такая схема получила название «gunning in place» (RIP) [33] – «бег на месте». Данная схема включена в LETKF и реализована в квазигеострофической модели [33], требующей очень много времени «на раскрутку», особенно когда она инициализируется из случайных равномерно распределенных начальных возмущений. Результаты показывают, что применение схемы RIP позволяет LETKF раскручиваться и сходиться к оптимальному уровню ошибки быстрее, чем 3D-Var или 4D-Var. При этом дополнительные вычисления «на месте» требуются только на начальном этапе раскрутки LETKF алгоритма.

2. Дисбаланс переменных вследствие локализации. Возможное решение – инициализация модели цифровым фильтром [59].

3. Численные эксперименты [17] подтвердили наличие двух проблем, связанных с неустранимой вертикальной локализацией спутниковых данных в пространстве наблюдений:

1) для интегральных (глобальных) наблюдений расстояние и форма локализации недостаточно определены;

2) ширина весовой функции наблюдений приводит к тому, что функция локализации либо слишком широка (локализация неэффективна, появляются ложные ковариации), либо слишком узка (подавляются или исключаются реальные ковариации ошибок между спутниковыми каналами).

5. Стратегия LETKF усвоения с интерполяцией оператора наблюдений

Исследования [11] показали, что LETKF может эффективно усваивать данные спутниковых наблюдений, если применить определенную стратегию усвоения, включающую механизм для выбора наблюдений, которые усваиваются в заданной точке сетки, и технологию коррекции смещения наблюдений на ансамблевой основе [11, 23].

Данная стратегия имеет некоторые характерные отличия и при усвоении «обычных» наблюдений. При этом в пространстве состояний модели рассматриваются траектории $\gamma_n = x(t)$, где $t \in \tau_n = [t_n - \Delta t/2; t_n + \Delta t/2]$, где t_n – время анализа; $x(t)$ – вектор конечномерного образа состояния атмосферы на сетке модели. Ансамбль анализов используется как начальные условия для получения ансамбля траекторий первого

приближения $\gamma_n^{b(k)}$. Разработка шага обновления состояния LETKF похожа на другие современные схемы.

Для локального вектора состояний, формируемого из переменных модели в точке модельной сетки l , LETKF генерирует K членов ансамбля локальных анализов при помощи вычисления ансамбля весовых векторов:

$$x_l^{a(k)} = \bar{x}_l^b + X_l^b w_l^{a(k)}, k=1, \dots, K. \quad (4)$$

Наилучшей оценкой локального состояния является среднее по ансамблю анализов

$$\bar{x}_l^a = \bar{x}_l^b + X_l^b \bar{w}_l^a. \quad (5)$$

Для обычных наблюдений усвоение происходит по следующему алгоритму.

На предварительном (нулевом) шаге определяется оператор наблюдений $h(\gamma_n)$. При этом в двух горизонтальных пространственных измерениях $h(\gamma_n)$ является простой билинейной интерполяцией.

После интерполяции по горизонтали, выполняется вертикальная линейная интерполяция в сигма-системе координат, полагая приземное давление равным его значению в поле первого приближения.

Шаг 1. Оператор наблюдений применяется к каждому члену ансамбля траекторий первого приближения для получения ансамбля $y^{b(k)}$ прогнозных значений в точках наблюдения. Вычисляется среднее \bar{y}^b по ансамблю $y^{b(k)}$ и строится матрица Y^b , столбцами которой являются векторы, полученные путем вычитания \bar{y}^b из каждого члена ансамбля $y^{b(k)}$.

Шаг 2 – локализация. Вектор \bar{y}_l^b и матрицы Y_l^b и R_l формируются путем выбора тех компонент, которые связаны с выбранным набором локальных наблюдений.

Шаг 3. Вычисляется весовой вектор:

$$\bar{w}_l^a = \tilde{P}_l^a (Y_l^b)^T R_l^{-1} (y_l^0 - y_l^b), \quad (6)$$

$$\text{где } \tilde{P}_l^a = [(k-1)I / \rho + (Y_l^b)^T R_l^{-1} Y_l^b]^{-1}, \quad (7)$$

$\rho \geq 1$ – фактор мультипликативной инфляции и I – единичная матрица.

$$\text{Шаг 4. Вычисляется матрица } W_l^a = [(k-1)\tilde{P}_l^a]^{1/2}.$$

Шаг 5. Весовой вектор \bar{w}_l^a добавляется к каждой строке W_l^a . Столбцами

результатирующей матрицы являются члены ансамбля весовых векторов.

6. Стратегия LETKF усвоения спутниковых данных

Для процедуры усвоения спутниковых наблюдений [38] предполагается поправка для оператора наблюдений:

$$\hat{h}(\gamma, \beta) = h(\gamma) + b, \quad (8)$$

удовлетворяющая соотношению

$$\hat{y}^0 = \hat{h}(\gamma^t) + \varepsilon. \quad (9)$$

Размерность J векторов \hat{h} , \hat{y}^0 и b равна общему числу спутниковых каналов, по которым усваиваются наблюдения. Вектор \hat{y}^0 состоит из компонентов y^0 , а вектор коррекции параметров β определяется из следующих условий: каждый компонент вектора коррекции оценивается линейной комбинацией

$$b_j = \beta_j^0 + \sum_{i=1}^I \beta_j^i p_i, j = 1, \dots, J \quad (10)$$

с набором предикторов $p_i(t)$. Типичными примерами предикторов модели являются температура и толщина различных слоев атмосферы, тогда как примером для предиктора наблюдения является угол сканирования, при котором прибор принимает спутниковое наблюдение. В рассматриваемой формулировке набор предикторов является одинаковым для всех наблюдений, которые формируют \hat{y}^0 , но коэффициенты $\beta_j^i(t)$ различны для разных каналов и инструментов.

Пока $b_j = \beta_j^0$, все предикторы равны нулю. Назовем β_j^0 отсечкой для канала, связанного с j -ой компонентой вектора коррекции b .

Оценки параметров смещения получим с помощью расширения вектора состояния (например, [20]). Для определения расширенного вектора состояния дополняем вектор состояния x вектором β , имеющим $(I + 1) \times J$ компонент:

$$z = \begin{bmatrix} x \\ \beta \end{bmatrix}. \quad (11)$$

Оценку расширенного вектора состояния получаем путем применения алгоритма LETKF к вектору z вместо вектора x .

Компоненты смещения $\beta^{b(k)}$ ансамбля первого приближения в момент t_n

предполагаются равными значениям анализов $\beta^{a(k)}$ параметров смещения в предыдущих анализах t_{n-1} .

Второй важной проблемой является нелокальный характер оператора наблюдений для излучения: в отличие от обычных наблюдений, оператор для данного наблюдения зависит полностью от данных атмосферной модели в горизонтальной плоскости наблюдения. Это свидетельствует о том, что для наблюдений за излучением необходимо изменить второй шаг стратегии вертикальной локализации.

Измененная стратегия отбора данных основана на функции вертикального взвешивания w_l для каждого наблюдения на всех уровнях модели. Для выбора уровней модели, где данное наблюдение учитывается при усвоении, применим стратегию отсечки, предложенную в [23]: 1) выбираем параметр отсечки $0 < \eta \leq 1$; 2) для каждого члена ансамбля находим уровень модели l_{\max}^k , где w_l^k принимает свое максимальное значение; 3) ищем слои наименьшей l_{bottom} и наибольшей l_{top} глубины вокруг уровня l_{\max}^k , где весовая функция удовлетворяет условию $w_l^k \leq \eta w_{\max}^k$; 4) для слоев, указанных в предыдущем шаге, вычисляем среднее по ансамблю l_{top} верхнего индекса и среднее по ансамблю l_{bottom} нижнего индекса; 5) усваиваются наблюдения из слоев, ограниченных l_{bottom} и l_{top} . Значение параметра отсечки η , которое обеспечивает анализ с приемлемой точностью при минимальных вычислительных затратах, находится в численных экспериментах.

Получаем следующие изменения в основных шагах LETKF алгоритма:

Нулевой шаг: оператор наблюдения определяется по формуле (8).

Шаг 1: ансамбль $y^{b(k)}$ прогнозных значений в точках наблюдения получается в результате применения оператора наблюдений \hat{h} к ансамблю траекторий первого приближения $\mathcal{Y}_n^{b(k)}$.

Шаг 2: наблюдения, которые формируют \hat{y}_l^0 компоненту y_l^0 на различных точках вертикальной сетки l , выбираются для усвоения с помощью пошаговой стратегии отсечки.

Шаги 3–5 алгоритма, которые определяют весовые коэффициенты $w_l^{a(k)}$ для расчета анализа локального расширенного вектора состояний, не отличаются от алгоритма для обычных наблюдений.

Компоненты $x^{a(k)}$ анализа расширенного вектора $z^{a(k)}$, $k = 1, \dots, K$, как и прежде,

получаются объединением компонент вектора $x_l^{a(k)}$. Для получения глобального анализа параметров смещения $\beta^{a(k)}$ из параметров коррекции для всех спутниковых каналов, усваиваемых в данной точке, вычисляются поправки к компонентам расширенного вектора состояния. Для получения отдельной оценки каждого из параметров смещения усредняются локальные оценки параметров по всем уровням. Взвешивание с обратной дисперсией гарантирует, что в местах, где неопределенность в оценке заданного параметра смещения имеет больший вклад, весовая поправка к глобальной оценке этого параметра смещения меньше.

Применение описанной стратегии для усвоения данных AMSU-A для яркостной температуры, в дополнение к другим оперативным наблюдениям (как температуры, так и ветра), привело к статистически значимому улучшению анализа ветра и температуры в Южном полушарии [11].

Заключение: конкуренция методов и гибридные схемы усвоения

Основной проблемой использования алгоритма фильтра Калмана для усвоения данных наблюдений в реальных прогностических моделях атмосферы и океана является большая размерность матриц ковариаций ошибок прогноза. Однако в настоящее время продолжают разрабатываться усовершенствованные методы и алгоритмы.

Один из подходов заключается в использовании упрощенных прогностических моделей при вычислении матрицы ковариаций ошибок прогноза (предложен в [19]) и называется субоптимальным фильтром Калмана [24, 50]. В [3, 34] был предложен алгоритм усвоения данных на основе адаптивного субоптимального фильтра Калмана. В [4, 18] предлагается алгоритм оценки шумов модели по данным наблюдений с использованием вектора «невязок» (разности между наблюдаемыми и спрогнозированными значениями) в процедуре фильтра Калмана. Алгоритм основан на совместном использовании упрощенной модели для вычисления матрицы ковариаций ошибок прогноза (субоптимальный фильтр Калмана) и адаптивной оценки шумов модели по данным наблюдений.

Несколько исследовательских групп разработали эффективные ансамблевые Кальмановские фильтры, которые были успешно протестированы в реальных наблюдениях за атмосферой в глобальных (например [29, 38, 39, 48, 54]) и локальных областях (например, [27, 54]). Исследования показывают, что стохастические ансамблевые фильтры дают лучшие результаты, чем 3D-Var схемы, но хуже, чем 4D-Var [54].

В сравнении, проведенном в [54], показано, что применение ансамблевого фильтра квадратного корня (EnSRF) и LETKF для глобальной модели среднесрочных прогнозов

NCER с разрешением T62L38, при использовании всех доступных оперативных наблюдений для атмосферы (исключая спутниковые данные), позволяет получить лучшие прогнозы, чем использование оперативной системы 3D-Var на том же наборе данных наблюдений.

До настоящего времени окончательно не ясно, сможет ли EnKF превзойти 4D-Var. В связи с этим продолжаются активные исследования, в основном – численные эксперименты, направленные на комплексное сравнение таких характеристик двух методов усвоения, как точность, вычислительная сложность и устойчивость (см. например [56–58, 60]).

Главная цель исследования [48] состоит в том, чтобы сделать обоснованное сравнение между LETKF и операционным 4D-Var. Предложен и осуществлен метод коррекции дополнительного смещения в LETKF для спутниковых наблюдений. Развитие и LETKF и 4D-Var в JMA обладает существенными преимуществами в связи с синергетическим эффектом конкуренции и является в настоящее время рекомендуемой многими исследователями стратегией.

Несмотря на то, что два подхода к усвоению данных – вариационный и стохастический (на основе фильтров Калмана) – оперируют принципиально разными алгоритмами, они имеют общую суть – метод наименьших квадратов. В последнее время появились гибридные схемы, объединяющие два подхода к усвоению данных. В этих схемах в основном с помощью ансамблей генерируется матрица ковариаций ошибок прогноза, которая используется в вариационном алгоритме (например [56–58]). Данные методы обычно предлагают вполне работоспособный алгоритм без доказательства его оптимальности и/или сходимости.

В [1, 2] предлагается принципиально иной метод усвоения, сочетающий вариационный и калмановский подходы. Предлагаемый метод основан на доказательстве сходимости линейного фильтра к диффузионному процессу, из которой следует система уравнений для оптимального фильтра, отличная от уравнений Калмана.

Таким образом, конкурентное взаимодействие методов усвоения данных ведет не только к совершенствованию каждого из них в отдельности, но и к развитию «смешанных» алгоритмов. Из самых свежих исследований, проводящих сравнение и синтез алгоритмов, стоит выделить [14, 15, 43, 47, 48].

Список использованных источников

1. *Беляев К.П.* Методы усвоения данных в гидродинамических моделях циркуляции и их применения для анализа состояния и изменчивости Мирового океана // Автореферат диссертации на соискание ученой степени доктора физико-математических наук. – М., 2011.

2. *Беляев К.П., Тучкова Н.П., Кирхнер И.* Метод коррекции модельных расчетов по данным измерений, основанный на диффузионном приближении, и его применения для анализа гидрофизических характеристик // Математическое моделирование. – 2009. – Т. 21, № 3. – С. 53–68.
3. *Климова Е.Г.* Методика усвоения данных наблюдений на основе адаптивного субоптимального фильтра Калмана // Труды Международной конференции «Математические методы в геофизике». Новосибирск, 2003. – Т. 2. – С. 400–404.
4. *Климова Е.Г.* Алгоритм усвоения данных наблюдений на основе адаптивного субоптимального фильтра Калмана // Метеорология и гидрология. – 2005. – № 3. – С. 24–35.
5. *Марчук Г.И.* Методы вычислительной математики. – М.: Наука, 1977. – 340 с.
6. *Марчук Г.И.* Основные и сопряженные уравнения динамики атмосферы и океана // Метеорология и гидрология. – 1974. – № 2. – С.17–34.
7. *Пененко В.В.* Методы численного моделирования атмосферных процессов. – Л.: Гидрометеиздат, 1981. – С. 352.
8. *Уваров Н.В.* Определение физических параметров атмосферы и океана методом вариационного усвоения данных спутниковых измерений // Диссертация на соискание ученой степени кандидата физико-математических наук. – Москва, 2007.
9. *Шляева А.В., Мизяк В.Г., Толстых М.А.* Экспериментальная система усвоения данных на основе локального ансамблевого фильтра Калмана для глобальной модели атмосферы // Труды Гидрометцентра России. – Вып. 348. – 2012. – См. наст. сборник.
10. *Anderson J.L.* An ensemble adjustment filter for data assimilation // Mon. Wea. Rev. – 2001. – Vol. 129. – P. 2884–2903.
11. *Aravequia J.A., Szunyogh I., Fertig E.J., Kalnay E., Kuhl D., Kostelich E.J.* Evaluation of a strategy for the assimilation of satellite radiance observations with the Local Ensemble Kalman Filter // Mon. Wea. Rev. – 2009. – Vol. 139, Issue 6 (June 2011). – P. 1932–1951.
12. *Bishop C.H., Etherton B.J., Majumdar S.* Adaptive sampling with the ensemble transform Kalman filter. Part I: theoretical aspects // Mon. Wea. Rev. – 2001. – Vol. 129. – P. 420–436.
13. *Bonavita M., Torrisi L., Marcucci F.* The ensemble Kalman filter in an operational regional NWP system: Preliminary results with real observations // Quart. J. R. Meteorol. Soc. – 2008. – Vol. 134. – P. 1733–1744.
14. *Buehner M., Houtekamer P.L., Cecilien C., Herschel L.M., Bin H.* Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part I: Description and Single-Observation Experiments // Mon. Wea. Rev. – 2010. – Vol. 138, Issue 5 (May 2010). – P. 1550–1566.
15. *Buehner M., Houtekamer P.L., Cecilien C., Herschel L.M., Bin H.* Intercomparison of Variational Data Assimilation and the Ensemble Kalman Filter for Global Deterministic NWP. Part II: One-Month Experiments with Real Observations // Mon. Wea. Rev. – 2010. – Vol. 138, Issue 5 (May 2010). – P. 1567–1586.

16. *Burgers G., van Leeuwen P.J, Evensen G.* Analysis scheme in the ensemble Kalman filter // *Mon. Wea. Rev.* – 1998. – Vol. 126. – P. 1719–1724.
17. *Campbell W.F., Bishop C.H., Hodyss D.* Vertical covariance localization for satellite radiances in ensemble Kalman filters // *Mon. Wea. Rev.* – 2010 . – Vol. 138. – P. 282–290.
18. *Daley R.* *Atmospheric Data Analysis.* – Cambridge, UK: Cambridge University Press, 1991. – 457 p.
19. *Dee D.P.* Simplification of the Kalman filter for meteorological data assimilation // *Quart. J. R. Meteorol. Soc.* – 1991. – Vol. 117. – P. 385–397.
20. *Dee D.P.* Bias and data assimilation // *Quart. J. R. Meteorol. Soc.* – 2005. – Vol. 131. – P. 3323–3343.
21. *Evensen G.* The ensemble Kalman filter: theoretical formulation and practical implementation // *Ocen. Dyn.* – 2003. – Vol. 53. – P. 343–367.
22. *Evensen G., van Leeuwen P.J.* Assimilation of geosat altimeter data for the agulhas current using the ensemble Kalman filter with a quasi-geostrophic model // *Mon. Wea. Rev.* – 1996. – Vol. 124. – P. 85–96.
23. *Fertig E.J. et al.* Observation bias correction with an ensemble Kalman filter // *Tellus.* – 2009. – Vol. 61A. – P. 210–226.
24. *Ghil M., Malanotte-Rizzoli P.* Data assimilation in meteorology and oceanography // *Adv. Geophys.* – 1991. – Vol. 33. – P. 141–266.
25. *Hamill T.M., Snyder C.* A hybrid ensemble Kalman filter-3d variational analysis scheme // *Mon. Wea. Rev.* – 2000. – Vol. 128. – P. 2905–2919.
26. *Hamill T.M., Whitaker J., Snyder C.* Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter // *Mon. Wea. Rev.* – 2001. – Vol. 129. – P. 2776–2790.
27. *Houtekamer P.L., Mitchell H.L.* Data assimilation using an ensemble Kalman filter technique // *Mon. Wea. Rev.* – 1998. – Vol. 126. – P. 796–811.
28. *Houtekamer P.L., Mitchell H.L.* A sequential ensemble Kalman filter for atmospheric data assimilation // *Mon. Wea. Rev.* – 2001. – Vol. 129. – P. 123–137.
29. *Houtekamer P.L., Mitchell H.L., Pellerin G., Buehner M., Charron M., Spacek L., Hansen B.* Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations// *Mon. Wea. Rev.* – 2005. – Vol. 133. – P. 604–620.
30. *Hunt B.R., Kalnay E., Kostelich E.J. et al.* Four-dimensional ensemble Kalman filtering // // *Tellus.* – 2004. – Vol. 56A. – P. 273–277.
31. *Hunt B.R., Kostelich E.J., Szunyogh I.* Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. // *Physica D.* – 2007. – Vol. 230. – P. 112–126.
32. *Kalman R.* A new approach to linear filtering and prediction problem // *J. Basic Engrg.* – 1960. – Vol. 1. – P. 35– 45.
33. *Kalnay E., Yang S-C.* Notes and Correspondence Accelerating the spin-up of Ensemble Kalman Filtering // *Quart. J. R. Meteorol. Soc.* – 2010. – Doi:10.1002/qj.652.

34. *Klimova E.G.* Adaptive algorithm of suboptimal Kalman filter // Research Activities in Atmospheric and Ocean Modeling. – 2004. – Report. № 34. – P. 0117–0118. – <http://collaboration.cmc.ec.gc.ca/science/wgne>.
35. *Lawson W.G., Hansen J.A.* Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth // Mon. Wea. Rev. –2004. – Vol. 132. –P. 1966–1981.
36. *Liu J., Li H., Kalnay E., Szunyogh I., Kostelich E.J.* Univariate and multivariate assimilation of AIRS humidity retrievals with the Local Ensemble Transform Kalman Filter // Mon. Wea. Rev. – 2009 . – Vol. 137. – P. 3918–3932.
37. *Lorenc A.* A global three-dimensional multivariate statistical interpolation scheme // Mon. Wea. Rev. – 1981. – No. 109. – P. 701–721.
38. *Miyoshi T., Sato Y.* Assimilating satellite radiances with a local ensemble transform Kalman filter (LETKF) applied to the JMA global model (GSM) // SOLA. – 2007. – Vol. 135. – P. 37–40.
39. *Miyoshi, T., Yamane S.* Local ensemble transform Kalman filter with an AGCM at a T159/L48 resolution // Mon. Wea. Rev. – 2007. – Vol. 135. – P. 3841–3861.
40. *Nerger L., Hiller W., Scroeter J.* A comparison of error subspace Kalman filter // Tellus. – 2005. – Vol. 57A. – P. 715–735.
41. *Ott E., Hunt B. R., Szunyogh I. et al.* Exploiting local low dimensionality of the atmospheric dynamics for efficient Kalman filtering // ArXiv: archive/paper 020458. – <http://arxiv.org/abs/physics/020358>. 2002.
42. *Parrish D.F., Derber J.C.* The national meteorological center's spectral statistical interpolation analysis system // Mon. Wea. Rev. – 1992. – Vol. 120. – P. 1747–1763.
43. *Piccolo C.* Growth of Forecast Errors from Covariances Modeled by 4DVAR and ETKF Methods // Mon. Wea. Rev. – 2011. – Vol. 139, Issue 5 (May 2011). – P. 1505–1518.
44. *Rabier F., Järvinen H., Klinker E. et al.* The ECMWF operational implementation of four dimensional variational assimilation. Part I: experimental results with simplified physics // Quart. J. R. Meteorol. Soc. – 2000. – Vol. 126. – P. 1143–1170.
45. *Szunyogh I., Kostelich E.J., Gyarmati G., Kalnay E., Hunt B.R., Ott E., Satterfield E., Yorke J.A.* A local ensemble transform Kalman filter data assimilation system for the NCEP global model // Tellus. – 2008. – Vol. 60A. – P. 113–130.
46. *Szunyogh I., Kostelich E.J., Gyarmati G., Patil D.J., Hunt B.R., Kalnay E., Ott E., Yorke J.A.* Assessing a local ensemble Kalman filter: Perfect model experiments with the National Centers for Environmental Prediction global model // Tellus. – 2005. – Vol. 57A. – P. 528–545.
47. *Takemasa M.* The Gaussian Approach to Adaptive Covariance Inflation and Its Implementation with the Local Ensemble Transform Kalman Filter // Mon. Wea. Rev. – 2011. – Vol. 139, Issue 5 (May 2011). – P. 1519–1535.

48. Takemasa M., Yoshiaki S., Takashi K. Ensemble Kalman Filter and 4D-Var Intercomparison with the Japanese Operational Global Analysis and Prediction System // *Mon. Wea. Rev.* – 2010. – Volume 138, Issue 7 (July 2010). – P. 2846–2866.
49. Tippett M.K., Anderson J. L., Bishop C.H. et al. Ensemble square-root filters // *Mon. Wea. Rev.* – 2002. – Vol. 131. – P. 1485–1490.
50. Todling R., Cohn S. Suboptimal schemes for atmospheric data assimilation based on the Kalman filter // *Mon. Wea. Rev.* – 1994. – Vol. 122. – P. 2530–2557.
51. Torn R.D., Hakim G.J. Performance characteristics of a pseudo-operational ensemble Kalman filter // *Mon. Wea. Rev.* – 2008. – Vol. 136. – P. 3497–3963.
52. Tsyrlunikov M.D. Stochastic modelling of model errors: a simulation study // *Quart. J. R. Meteorol. Soc.* – 2005. – Vol. 131, Issue 613 (October 2005. Part C). – P. 3345–3371.
53. Whitaker J.S., Hamill T.H. Ensemble data assimilation without perturbed observations // *Mon. Wea. Rev.* – 2002. – Vol. 130. – P. 1913–1924.
54. 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. – Vol. 136. – P. 463–482.
55. Wiener N., Masany P. Prediction theory of multivariate stochastic processes // *Acta Math.* – 1957. – Vol. 98. – P. 111–150.
56. Wang X., Snyder C., Hamill T.M. On the theoretical equivalence of differently proposed ensemble/3D-Var hybrid analysis schemes // *Mon. Wea. Rev.* – 2007. – Vol. 135. – P. 222–227.
57. Wang X., Hamill T.M., Whitaker J.S., Bishop C.H. A comparison of hybrid ensemble transform Kalman filter-OI and ensemble square-root filter analysis schemes // *Mon. Wea. Rev.* – 2007. – Vol. 135. – P. 1055–1076.
58. Wang X., Barker D.M., Snyder C., Hamill T.M. A hybrid ETKF–3DVAR data assimilation scheme for the WRF Model. Part II: Real observation experiments // *Mon. Wea. Rev.* 2008. – Vol. 136. – P. 5132–5147.
59. Yang S-C, Kalnay E. Handling nonlinearity and non-Gaussianity in Ensemble Kalman Filter. A special collection ‘Intercomparisons of 4D-Variational Assimilation and the Ensemble Kalman Filter // *Mon. Wea. Rev.* – 2010 (submitted).
60. Yang S-C, Corazza M, Carrassi A, Kalnay E, Miyoshi T. Comparison of ensemble-based and variational-based data assimilation schemes in a quasi-geostrophic model // *Mon. Wea. Rev.* – 2009. – Vol. 137. – P. 639–709.

Поступила в редакцию 25.09.2012 г.