

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ, УЛУЧШАЮЩИХ СХОДИМОСТЬ АНСАМБЛЕВОГО ФИЛЬТРА КАЛМАНА

Е. Г. Климова

Институт вычислительных технологий СО РАН, 630090, Новосибирск

УДК 551.509.313

Ансамблевый фильтр Калмана является одним из популярных методов усвоения данных при моделировании процессов в атмосфере, водоемах, распространения загрязняющих веществ и т.д. Одной из проблем практической реализации ансамблевых фильтров Калмана является расходимость процедуры усвоения данных со временем из-за убывания дисперсий ошибок оценивания. Существует несколько подходов к решению этой проблемы: применение корректирующего параметра (inflation factor) как мультипликативного, так и аддитивного, задание случайной ошибки модели (stochastic forcing), а также адаптивные процедуры коррекции дисперсий ошибок прогноза. В статье рассматривается общий подход к анализу эффективности таких процедур. Подход основан на формуле для отклонения элемента ансамбля от среднего по ансамблю значения. Свойства процедур, регулирующих сходимость, исследуются для стохастического ансамблевого фильтра Калмана как аналитически, так и с помощью численных экспериментов с 1-мерной моделью Лоренца.

Ключевые слова: усвоение данных, ансамблевый фильтр Калмана, расходимость фильтра Калмана, коррекция ковариационных матриц.

Введение

Лидирующим направлением в работах, посвященных применению фильтра Калмана при усвоении данных, является ансамблевый подход, впервые предложенный в работе [4]. Ансамблевый подход позволяет вычислять матрицы ковариаций ошибок оценивания для нелинейных прогностических моделей. В этом случае рассматривается вариант обобщенного фильтра Калмана, в котором ковариации ошибок прогноза оцениваются с помощью ансамбля прогнозов.

В ансамблевом фильтре Калмана можно выделить два подхода: «стохастический фильтр» и «детерминированный фильтр». В настоящее время проводится большое количество исследований по сравнению стохастических и детерминированных ансамблевых фильтров. Важной проблемой, возникающей при реализации ансамблевого фильтра Калмана, является проблема расходимости фильтра. При этом под расходимостью понимается рост среднеквадратической ошибки при убывании нормы ковариационной матрицы. Такое поведение может быть свойственно и классическому фильтру Калмана, у которого теоретически, при отсутствии шумов модели, норма ковариационной матрицы стремится к нулю. Различные приемы, позволяющие бороться с этим эффектом, разработаны в классических монографиях Jazwinsky [6], [2]. Современные методы, улучшающие сходимость ансамблевого фильтра Калмана, изложены в [7].

В данной работе рассматривается подход к анализу поведения по времени отклонения среднего по ансамблю значения (ансамбль возмущений) и соответствия его поведению теоретической ошибки оценки, основанный на общей асимптотической формуле для отклонения элемента ансамбля от среднего по ансамблю значения. Исследуется влияние на асимптотическое поведение по времени ансамбля возмущений основных приемов, улучшающих сходимость ансамблевого фильтра Калмана: применение аддитивного корректирующего параметра, мультипликативного параметра и адаптивной оценки этих параметров в процедуре усвоения данных.

1 Ансамблевый фильтр Калмана

Запишем нелинейную динамическую систему в виде уравнения процесса

$$x_k^t = f(x_{k-1}^t) + \eta_{k-1}^t \quad (1)$$

и уравнения наблюдений

$$y_k = h(x_k^t) + \varepsilon_k^t.$$

где η_{k-1}^t — вектор “шумов модели”, ε_k^t — вектор ошибок наблюдений, x_k^t — вектор оцениваемых переменных в момент времени t_k , ε_k^t и η_{k-1}^t — Гауссовские случайные переменные: $E[\varepsilon_k^t (\varepsilon_k^t)^T] = R_k^t$, $E[\eta_{k-1}^t (\eta_{k-1}^t)^T] = Q_{k-1}^t$. Будем считать x_k^t истинным значением.

Стохастический ансамблевый фильтр Калмана состоит из ансамбля прогнозов $\{x_k^{f,n}, n = 1, \dots, N\}$

$$x_k^{f,n} = f(x_{k-1}^{a,n}) + \eta_{k-1}^n \quad (2)$$

и ансамбля анализов $\{x_k^{a,n}, n = 1, \dots, N\}$

$$x_k^{a,n} = x_k^{f,n} + K_k(y_k^n + \varepsilon_k^n - h(x_k^{f,n})). \quad (3)$$

Для осуществления стохастического фильтра Калмана требуется задание ансамбля ошибок наблюдений $\{\varepsilon_k^n, n = 1, \dots, N\}$ а также ансамбля ошибок прогноза $\{dx_k^{f,n} = x_k^{f,n} - \overline{x_k^{f,n}}, n = 1, \dots, N\}$, где $\overline{x_k^{f,n}} \cong \frac{1}{N} \sum_{n=1}^N x_k^{f,n}$ и ансамбля шумов модели $\{\eta_{k-1}^n, n = 1, \dots, N\}$: $E[\eta_{k-1}^n (\eta_{k-1}^n)^T] = Q_k$. Матрица K_k имеет вид

$$K_k = P_k^f H_k^T (H_k P_k^f H_k^T + R_k)^{-1},$$

где P_k^f и R_k — матрицы, оцениваемые по ансамблю

$$P_k^f \triangleq \frac{1}{N-1} \sum_{n=1}^N dx_k^{f,n} (dx_k^{f,n})^T, \quad R_k \triangleq \frac{1}{N-1} \sum_{n=1}^N \varepsilon_k^n (\varepsilon_k^n)^T,$$

H_k — линеаризованный оператор $h(x_k^{f,n})$ относительно $\overline{x_k^{f,n}}$:

$$h(x_k) \cong h(\overline{x_k^{f,n}}) + H_k \varepsilon_k^f.$$

Зададим ошибку анализа $\{dx_k^{a,n} = x_k^{a,n} - \overline{x_k^{a,n}}, n = 1, \dots, N\}$, где $\overline{x_k^{a,n}} \cong \frac{1}{N} \sum_{n=1}^N x_k^{a,n}$. Тогда матрица ковариаций ошибок анализа может быть вычислена по формуле

$$P_k^a \triangleq \frac{1}{N-1} \sum_{n=1}^N dx_k^{a,n} (dx_k^{a,n})^T.$$

Формулы (2)–(3) представляют собой стохастический ансамблевый фильтр Калмана. Эти формулы можно записать в следующем виде:

$$x_k^n = (I - K_k H_k) [f(x_{k-1}^n) + \eta_{k-1}^n] + K_k(y_k^n + \varepsilon_k^n),$$

где x_k^n — ансамбль оцениваемых переменных. Оптимальной оценкой в ансамблевом фильтре Калмана является среднее по ансамблю значение $\overline{x_k^n}$. Отклонение от среднего (spread) моделирует ошибку оценки $dx_k^n \triangleq x_k^n - \overline{x_k^n}$. Уравнение для отклонения от среднего имеет вид

$$dx_k^n = (I - K_k H_k)(f(x_{k-1}^n) - \overline{f(x_{k-1}^n)}) + \eta_{k-1}^n + K_k \varepsilon_k^n. \quad (4)$$

В то же время “теоретическая” ошибка оценки $dx_k^t \triangleq x_k^t - \overline{x_k^n}$ удовлетворяет уравнению

$$dx_k^t = (I - K_k H_k)(f(x_{k-1}^t) - \overline{f(x_{k-1}^n)}) + \eta_{k-1}^n + K_k \varepsilon_k^t. \quad (5)$$

Ансамблей анализов, соответствующих ковариационной матрице P_k^a и среднему значению $\overline{x_k^{a,n}}$ может быть множество. В частности, детерминированные фильтры вычисляют ансамбль полей с соответствующим средним значением и матрицей ковариаций. Однако, ансамбль анализов, задаваемый по формуле (3), имеет важное преимущество, состоящее в том, что уравнение для отклонения от среднего по ансамблю значения (4) будет иметь такой же вид, как и уравнение для теоретической ошибки оценки (5).

2 Асимптотическое поведение отклонения от среднего значения в ансамблевом фильтре Калмана

Запишем уравнение для отклонения от среднего в стохастическом фильтре Калмана в виде

$$dx_k^n = (I - K_k H_k)(F_k dx_{k-1}^n + \eta_{k-1}^n) + K_k \varepsilon_k^n. \quad (6)$$

Здесь вместо нелинейного оператора модели f берется линеаризованный оператор F_k . Это уравнение можно переписать как

$$dx_k^n = \Psi(k, 0) dx_0^n + \sum_{i=1}^k \Psi(k, i) [K_i \varepsilon_i^n + (I - K_i H_i) \eta_{i-1}^n], \quad (7)$$

где $\Psi(k, i) \triangleq \prod_{j=i+1}^k (I - K_j H_j) F_j$, если $k > i$, $\Psi(k, k) \equiv I$. Отметим, что аналогичное уравнение для ошибки классического фильтра Калмана получено в работе [3]. Как видно из формул, для того, чтобы $\|\Psi(k, 0)\| \rightarrow 0$ при $k \rightarrow \infty$ достаточно выполнения условия

$$\|(I - K_k H_k) F_k\| < 1 \quad (8)$$

на каждом шаге по времени. Это условие сходимости фильтра Калмана приводится в монографии Anderson et al. [2]. В этом случае со временем начальное значение возмущения перестает оказывать влияние на текущее значение ансамбля возмущений, поскольку $\|\Psi(k, 0)\| \rightarrow 0$ при $k \rightarrow \infty$.

Ошибка оценки, равная отклонению среднего значения от “истины” удовлетворяет аналогичному (6) уравнению

$$dx_k^t = (I - K_k H_k)(F_k dx_{k-1}^t + \eta_{k-1}^t) + K_k \varepsilon_k^t. \quad (9)$$

Как видно из уравнений для dx_k^n (6) и dx_k^t (9), моделируемая ошибка оценки близка к “теоретической” ошибке, если случайные векторы моделируемых ошибок наблюдений и шумов модели имеют те же ковариационные матрицы, что и истинные.

3 Способы улучшения сходимости ансамблевого фильтра Калмана

В современных методах усвоения данных, основанных на ансамблевом фильтре Калмана, рассматривается множество способов улучшения сходимости алгоритма [7]. Выделим среди них наиболее часто используемые:

1. мультипликативный корректирующий параметр (multiplicative inflation);
2. аддитивный корректирующий параметр (additive inflation);
3. адаптивная оценка корректирующих параметров (adaptive inflation).

Асимптотическая формула для ансамбля возмущений (7) позволяет проанализировать влияние каждого из этих способов. Рассмотрим модификацию ансамбля возмущений после шага прогноза в общем виде:

$$\widetilde{dx}_k^n = (I - K_k H_k)(\alpha_k F_k dx_k^n + \alpha_k \eta_{k-1}^n + \beta_k^n) + K_k \varepsilon_k^n,$$

где α_k — заданный параметр, β_k^n — случайный вектор с заданной ковариационной матрицей. Тогда асимптотическая формула для стохастического фильтра будет иметь вид

$$\widetilde{dx}_k^n = \tilde{\Psi}(k, 0) dx_0^n + \sum_{i=1}^k \tilde{\Psi}(k, i) [\tilde{K}_i \varepsilon_i^n + (I - \tilde{K}_i H_i) \alpha_i \eta_{i-1}^n + (I - \tilde{K}_i H_i) \beta_{i-1}^n],$$

где $\tilde{\Psi}(k, i) \triangleq \prod_{j=i+1}^k (I - \tilde{K}_j H_j) F_j \alpha_j$, \tilde{K}_j вычислена по модифицированным ковариационным матрицам. Как легко видеть из приведенных формул, ансамбль возмущений, полученный после модификации, не соответствует ансамблю ошибок. В случае $\beta_k^n = 0$ получаем вариант мультипликативного корректирующего параметра (multiplicative inflation). Рассмотрим вариант $\alpha_k = 1$, $\beta_k^n = \delta Q_k \xi_k^n$, где

$$\delta Q_k = (\alpha_k^2 - 1) P_k^f, \quad (10)$$

ξ^n — Гауссовский случайный вектор $N(0, I)$, I — единичная матрица. В этом случае получим вариант аддитивного корректирующего параметра (additive inflation), при котором ковариационные матрицы будут совпадать, а вид ансамбля возмущений отличаться. Также из асимптотических формул, приведенных в этом пункте, следует, что в случае аддитивного корректирующего параметра поведение ансамбля возмущений стохастического фильтра ближе к поведению ошибки, чем в случае мультипликативного корректирующего параметра (матрицы $\tilde{\Psi}$ и $\tilde{\Psi}_t$ совпадают).

4 Адаптивные алгоритмы

Популярным подходом к улучшению сходимости фильтра Калмана, предлагаемым в классической литературе, является адаптивная оценка матрицы ковариаций «шумов модели» по данным о невязках [6]. Классический адаптивный алгоритм оценки матрицы ковариаций Q , предлагаемый в этой монографии, является аналогом подхода, называемого «additive inflation». При этом дисперсии ковариационной матрицы уточняются с учетом данных о невязках.

Существует большое количество адаптивных алгоритмов оценки корректирующего параметра в ансамблевом фильтре Калмана [7]. Рассмотрим один из алгоритмов адаптивного уточнения дисперсий прогнозируемых ансамблей по данным о невязках (residuals) в ансамблевом фильтре Калмана. Пусть в точках наблюдений “ i ” в момент времени t_k известны невязки $r_{k,i} = y_{0,i} - H\bar{x}_k^f$. Предположим, по аналогии с алгоритмом, предложенным в [6], что $r_{k,i}^2 = (\sigma_{k,i}^f)^2 + \varepsilon_{0,i}^2$. То есть, данные о невязках в точках наблюдений содержат информацию о дисперсиях ошибок прогноза (с учетом «шумов модели»). Будем оценивать среднее по области значение мультипликативного корректирующего параметра как

$$\alpha^2 = \frac{\langle r_{k,i}^2 - \overline{\varepsilon_{0,i}^2} \rangle}{\langle (\sigma_{k,i}^f)^2 \rangle}, \quad (11)$$

где $\langle \cdot \rangle$ — среднее по области значение. В этом случае этот параметр не является константой, а оценивается в каждый момент усвоения, при этом возможны случаи, когда он становится меньше 1, либо намного больше 1. Если полученное значение отрицательно, будем считать его равным 1. Так как невязки вычисляются по данным о прогнозе, адаптивная оценка мультипликативного параметра осуществляется на шаге прогноза. Этот алгоритм является аналогом алгоритма, предложенного в [10] для детерминированного фильтра Калмана. Предлагаемый способ уточнения дисперсий учитывает также дисперсию «шумов модели», которые, как правило, неизвестны.

Таким образом, могут быть рассмотрены варианты адаптивного уточнения дисперсий в виде аддитивного корректирующего параметра или мультипликативного корректирующего параметра.

5 Численные эксперименты

В численных экспериментах использовался вариант стохастического ансамблевого фильтра Калмана, предложенный в [8], который реализован в виде квадратного корня (ансамблевый π -алгоритм). Этот алгоритм обладает свойством локальности, то есть, вычисление ансамбля ошибок анализа производится в каждом узле сетки независимо. Алгоритм является экономичным: все арифметические операции осуществляются с матрицами размерность которых равна размерности ансамбля. Описание метода реализации этого алгоритма, использованного в численных экспериментах, приводится в [1].

Численные эксперименты проводились с моделью Lorenz-96 [9]. Уравнения модели имеют вид:

$$\begin{aligned} \frac{dx_j}{dt} &= (x_{j+1} - x_{j-2})x_{j-1} - x_j + F_0, \quad j = 1, \dots, J \\ x_{-1} &= x_{J-1}, \quad x_{J+1} = x_1, \end{aligned}$$

где x_1, \dots, x_J ($J=40$) — прогнозируемые переменные. Это одна из простейших возможных систем, которая воспроизводит свойства многих атмосферных моделей. Для решения системы уравнений используется конечно-разностная схема Рунге-Кутты 4-го порядка точности. Шаг по времени $\Delta t = 0.05$, что соответствует 6 часам ($t=1$ принимается за 5 суток), $F_0=8$. Для моделирования “истины” в численных экспериментах

по усвоению данных, так же как в [9] рассматривалось начальное значение $x_0^t \cong N(F_0/4; F_0/2)$ и проводился расчет на N_t шагов по времени. Через $N(a, b)$ обозначена случайная величина, распределенная по нормальному закону с математическим ожиданием, равным a и дисперсией, равной b . В численных экспериментах в качестве начальных данных для “истины” брался прогноз на $N_t=1000$ шагов по времени. Для получения начальных данных для прогноза по модели \hat{x}_0 к «истинным» начальным данным добавлялось возмущение $x_d(0) = x_t(0) + \delta$, $\delta \sim N(0, s_0)$.

Для организации численных экспериментов задавались: ансамбль начальных полей $x^n(0) = x_d(0) + \delta^n$, $\delta^n \sim N(0, s_0)$, $n = 1, \dots, N_{\text{ens}}$; наблюдения $y_0 = x_t(0) + \delta_0$, $\delta_0 \sim N(0, \varepsilon_0)$; ансамбль наблюдений с возмущениями $y_0^n = y_0 + \delta_0^n$, $\delta_0^n \sim N(0, \varepsilon_0)$, $n = 1, \dots, N_{\text{ens}}$. Во всех экспериментах “шумы модели” $\eta^n = 0$, при моделировании “истины” $\eta^t = 0.01$. Через N_{ens} обозначено число элементов ансамбля. Данные наблюдений считались известными во всей области интегрирования. Прогноз осуществлялся в течение $N_t=3000$ шагов по времени, усвоение проводилось на каждом шаге по времени и через каждые четыре шага по времени. Численные эксперименты проводились для значений $s_0 = \varepsilon_0 = 1$, $N_{\text{ens}}=20$.

Во всех численных экспериментах рассматривался вариант $R = \varepsilon_0^2 I$. При анализе в узле сетки l брались данные наблюдений из интервала $(l - id, l + id)$. При этом при анализе в узле сетки l вместо матрицы R бралась матрица $\hat{R} = R \circ e^{-0.5(\rho_{il}/bc)^2}$, где ρ_{il} — расстояние между узлом сетки и наблюдением, “ \circ ” — знак поэлементного умножения. В проводимых экспериментах брались значения $id = 5$, $bc = 5\Delta x$ (Δx — шаг сетки).

Численные эксперименты проводились для 10 вариантов «истины», все оценки вычислялись как среднее по этим 10 вариантам. Рассматривались следующие оценки:

$$rms = \frac{1}{K} \sum_{k=1}^K \left\{ \frac{1}{L} \sum_{i=1}^L (\bar{x}_{k,i} - x_{k,i}^t)^2 \right\}^{1/2},$$

среднеквадратическая ошибка, средняя по $K=10$ вариантам вычислений (k — номер варианта, i — номер узла сетки), в каждом варианте задаются случайные значения ошибок начальных данных, ошибок данных наблюдений и ансамбля возмущений в начальный момент времени. Если рассматривать “истину” как случайную величину, то значение rms соответствует следу ковариационной матрицы ошибки (ковариация оценивается по выборке из 10 значений).

$$sp = \frac{1}{K} \sum_{k=1}^K \left\{ \frac{1}{N} \sum_{n=1}^N \left[\frac{1}{L} \sum_{i=1}^L (\bar{x}_{k,i}^n - x_{k,i}^n)^2 \right]^{1/2} \right\},$$

среднее значение следа ковариационной матрицы, вычисленной по ансамблю (n — номер элемента ансамбля).

Оценки rms и sp вычислялись на каждом шаге по времени. Также вычислялись средние значения этих оценок по временному интервалу $\{1001, \dots, 3000\}$:

$$\overline{rms} = \frac{1}{N_t - 1001} \sum_{m=1001}^{N_t} rms_m,$$

$$\overline{sp} = \frac{1}{N_t - 1001} \sum_{m=1001}^{N_t} sp_m.$$

Кроме того, вычислялось значение корреляции между rms и sp на этом временном интервале.

С описанным выше ансамблевым π -алгоритмом были проведены численные эксперименты в одинаковых условиях. Модификация ансамбля возмущений проводится после шага прогноза. В первой серии экспериментов данные наблюдений задавались на каждом шаге по времени, во второй серии — на каждом четвертом. В каждой серии проводилось 4 эксперимента:

1. Эксперимент 1 — при усвоении использовался мультипликативный параметр 1.1 после шага прогноза.
2. Эксперимент 2 — при усвоении использовался аддитивный параметр, соответствующий мультипликативному параметру 1.1 (формула (10)).
3. Эксперимент 3 — при усвоении использовался мультипликативный параметр, оцениваемый адаптивно так, как это описано в п.5 (формула (11)).

4. Эксперимент 4 — при усвоении использовался аддитивный параметр, соответствующий адаптивному мультипликативному параметру (формула (10)).

На рис. 1 приводятся значения средних по времени rms (rm) и sp (sp), а также коэффициента корреляции (cr), полученных в экспериментах 1–4 в первой и второй сериях экспериментов. Как видно из этого рисунка, использование аддитивного параметра увеличивает корреляцию между rms и sp по сравнению с мультипликативным параметром. Во второй серии экспериментов адаптивный алгоритм уменьшает среднеквадратическую ошибку. Кроме того, адаптивный алгоритм повышает коэффициент корреляции во всех экспериментах, при этом сами значения rms увеличиваются.

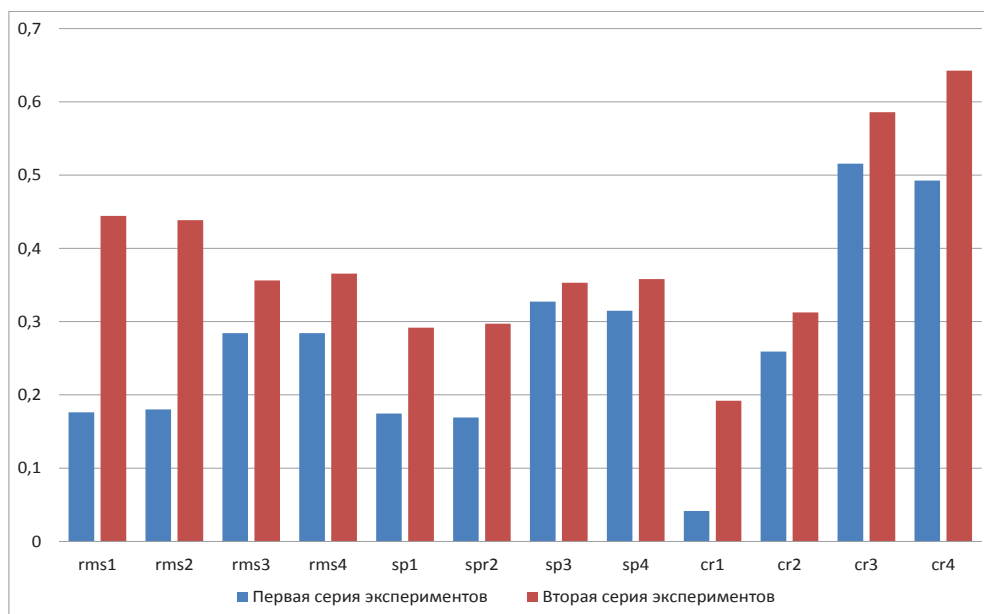


Рис. 1: Среднеквадратическая ошибка (rms), след ковариационной матрицы (sp) и коэффициент корреляции между rms и sp (cr).

Заключение

В работе исследуются основные факторы, влияющие на поведение ансамбля возмущений в ансамблевом фильтре Калмана. Получена общая формула для ансамбля возмущений в стохастическом ансамблевом фильтре Калмана. Рассмотрено влияние применения различных способов улучшения сходимости на асимптотическое поведение ансамблевого фильтра Калмана. Из проведенных исследований можно сделать следующие выводы:

1. применение аддитивного корректирующего параметра и мультипликативного корректирующего параметра близко по смыслу и оказывает похожее влияние на поведение ковариационных матриц, в то же время оно изменяет общую формулу для ансамбля возмущений;
2. асимптотическая формула для ансамбля возмущений при использовании аддитивного корректирующего параметра ближе к асимптотической формуле для ошибки оценки, чем при использовании мультипликативного корректирующего параметра;
3. адаптивный алгоритм увеличивает корреляцию по времени между следом ковариационной матрицы и среднеквадратической ошибкой оценки.

Список литературы

- [1] Климова Е.Г. Экономичный алгоритм усвоения данных, основанный на стохастическом ансамблевом фильтре Калмана // Труды Международной конференции «Актуальные проблемы вычислительной и прикладной математики. 2015», 19–23 октября 2015 г., г. Новосибирск. 2015. С. 339–344.
- [2] Anderson, B., Moore, J. Optimal filtering. Prentice-Hall, 1979.
- [3] Cohn, S.E., Dee, D.P. . Observability of discretizes partial differential equations // SIAM J. Numer. Anal. 1988. V. 25, iss. 3, P. 586–617.
- [4] Evensen, G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics // Journal Geophysical Research. 1994. V. 99, P. 10143–10162.
- [5] Evensen, G. Data assimilation. The ensemble Kalman filter, Spriger-Verlag, Berlin Heideberg, 2009.
- [6] Jazwinski, A.H. Stochastic processes and filtering theory. Academic Press, New York, 1970.
- [7] Houtekamer, H.L. Zhang, F. Review of the ensemble Kalman filter for atmospheric data assimilation // Monthly Weather Review. 2016. V. 144, P. 4489–4532.
- [8] Klimova E. A suboptimal data assimilation algorithm based on the ensemble Kalman filter // Quarterly Journal of the Royal Meteorological Society. 2012. V. 138, P. 2079–2085.
- [9] Lorenz, E.N., Emanuel, K.A. Optimal sites for supplementary weather observations: simulation with a small model // Journal of the Atmospheric Sciences. 1998. V. 55, P. 399–414.
- [10] Wang, X., Bishop, C.H. A comparison of breeding and ensemble transform Kalman filter ensemble forecast schemes // Journal of the Atmospheric Sciences. 2003. V. 60, P. 1140–1158.

*Екатерина Георгиевна Климова — д.ф.-м.н., ст. науч.сотр. Института
вычислительных технологий СО РАН;
e-mail: klimova@ict.nsc.ru.
Дата поступления — 30 мая 2017 г.*