

УДК 551.509.313

## Локальный ансамблевый алгоритм усвоения данных для нелинейных геофизических моделей\*

Е.Г. Климова

Федеральный исследовательский центр информационных и вычислительных технологий, просп. Акад. Лаврентьева, 6, Новосибирск, 630090

E-mail: klimova@ict.nsc.ru

Английская версия этой статьи печатается в журнале “Numerical Analysis and Applications” № 1, Vol. 16, 2023.

**Климова Е.Г.** Локальный ансамблевый алгоритм усвоения данных для нелинейных геофизических моделей // Сиб. журн. вычисл. математики / РАН. Сиб. отд-ние. — Новосибирск, 2023. — Т. 26, № 1. — С. 27–42.

Для решения задачи оптимальной оценки интересующих нас величин по данным наблюдений и модели (задачи оптимальной фильтрации) в нелинейном случае может применяться метод частиц, основанный на байесовском подходе. Недостатком классического фильтра частиц является то, что данные наблюдений участвуют только при нахождении весовых коэффициентов, с которыми вычисляется сумма частиц при определении оценки. В настоящей статье рассматривается подход к решению задачи нелинейной фильтрации, использующий представление апостериорной плотности распределения оцениваемой величины в виде суммы с весами гауссовских плотностей распределения. Из теории фильтрации известно, что в случае, если плотность распределения представляет собой сумму с весами функций Гаусса, оптимальной оценкой будет сумма с весами оценок, вычисленных по формулам фильтра Калмана. В настоящей статье предлагается метод решения задачи нелинейной фильтрации, основанный на данном подходе. Для реализации алгоритма используется ансамблевый  $\pi$ -алгоритм, предложенный автором ранее. Ансамблевый  $\pi$ -алгоритм в предлагаемом новом методе применяется при получении ансамбля, соответствующего плотности распределения на шаге анализа. Он представляет собой стохастический ансамблевый фильтр Калмана, при этом является локальным, что позволяет использовать его в геофизических моделях высокой размерности.

DOI: 10.15372/SJNM20230103

**Ключевые слова:** усвоение данных, ансамблевый фильтр Калмана, фильтр частиц, гауссовский смешанный фильтр.

**Klimova E.G.** A local ensemble data assimilation algorithm for nonlinear geophysical models // Siberian J. Num. Math. / Sib. Branch of Russ. Acad. of Sci. — Novosibirsk, 2023. — Vol. 26, № 1. — P. 27–42.

For optimal estimation of quantities of interest from observational data and a model (optimal filtering problem) in the nonlinear case, a particle method based on a Bayesian approach can be used. A disadvantage of the classical particle filter is that the observations are used only to find the weight coefficients with which the sum of the particles is calculated when determining an estimate. The present article considers an approach to solving the problem of nonlinear filtering which uses a representation of the posterior distribution density of the quantity being estimated as a sum with weights of Gaussian distribution densities. It is well-known from filtration theory that if a distribution density is a sum with weights of Gaussian functions, the optimal estimate will be a sum with weights of estimates calculated by the Kalman filter formulas. The present article proposes a method for solving the problem of nonlinear filtering based on this approach. An ensemble  $\pi$ -algorithm proposed earlier by the author is used to implement the method. The ensemble  $\pi$ -algorithm in this new method is used to obtain an ensemble corresponding to the distribution density at the analysis step. This is a stochastic ensemble Kalman filter which is local as well. Therefore, it can be used in high-dimensional geophysical models.

---

\*Работа выполнена в рамках государственного задания Минобрнауки России для Федерального исследовательского центра информационных и вычислительных технологий.

**Keywords:** *data assimilation, ensemble Kalman filter, particle filter, Gaussian mixture filter.*

---

## 1. Введение

Проблема получения оптимальной оценки геофизических полей по данным наблюдений и математической модели (усвоение данных) является одной из современных актуальных задач. К настоящему времени разработано большое количество методов усвоения данных, применяемых при моделировании различных геофизических процессов. В общем случае, если рассматривать геофизические поля и данные наблюдений как случайные переменные, для решения задачи усвоения данных используется байесовский подход [6, 17, 24]. В случае, когда рассматриваются линейная модель и линейный оператор наблюдений, задача оптимальной оценки по данным наблюдений на заданном временном интервале и прогнозам по модели в конце этого интервала называется задачей фильтрации. Решением этой задачи является оценка, формируемая фильтром Калмана [13]. Если оптимальная оценка ищется в заданном временном интервале, такая задача называется задачей сглаживания [13]. При условии линейности модели прогноза и оператора наблюдений эта задача эквивалентна задаче усвоения данных в вариационной постановке (4DVAR) [13, 17].

В работе [7] было предложено применение метода Монте-Карло для решения задач фильтрации и сглаживания (так называемые алгоритмы ансамблевой фильтрации и ансамблевого сглаживания). В этих алгоритмах используется выборка (ансамбль) прогнозов и наблюдений для аппроксимации ковариационных матриц [7].

Существуют проблемы применения ансамблевых алгоритмов в нелинейном случае, а также для систем высокой размерности. Кроме того, ансамблевые алгоритмы имеют проблемы, связанные со сходимостью. Для решения задачи сглаживания в нелинейном случае в статье [8] предлагаются итерационные алгоритмы, использующие вариационную постановку задачи, а также ансамбли прогнозов для вычисления ковариаций и линеаризации операторов прогноза и наблюдений.

В нелинейном негауссовском случае для решения задач оптимальной фильтрации может использоваться метод частиц, основанный на байесовском подходе [17, 23]. Термин “фильтр частиц” используется, когда плотность распределения общего вида (негауссовская) оценивается с помощью ансамбля состояний, называемых частицами [17, 23]. В настоящее время метод частиц становится очень популярным. Однако, применение этого метода в современных моделях атмосферы и океана сложно реализуемо из-за большой размерности моделей. Следует отметить, что размерность вектора прогнозируемых величин в современных моделях достигает порядка  $10^9$ . Также, как отмечено в работе [21], проблемой является возможная расходимость фильтра частиц. Под расходимостью фильтра частиц понимается ситуация, когда все коэффициенты в сумме стремятся к нулю, кроме одного, стремящегося к единице. Для предотвращения расходимости необходимо использование большого количества частиц. Кроме того, в работе [21] показано, что при реализации метода частиц в современных моделях большой размерности требуется применение локализации так же, как это осуществляется в ансамблевых фильтрах. Локализация состоит в понижении порядка используемых в вычислениях векторов и осуществляется различными способами. Это достигается декомпозицией расчетной области, а также ограничением количества данных наблюдений, используемых в конкретных подобластях [11]. В статье [18] предлагается подход к локализации в фильтре частиц по аналогии с локализацией, применяемой в ансамблевом фильтре Калмана. В этой работе

показано, что алгоритм с локализацией сходится даже при небольшом числе частиц.

Недостатком классического фильтра частиц является то, что данные наблюдений используются только при оценке весовых коэффициентов для элементов ансамбля (частиц) при вычислении суммы. В настоящее время в ряде работ рассматривается приближенное решение проблемы нелинейной фильтрации, основанное на представлении плотности распределения в виде суммы гауссовских функций. Из теории фильтрации известно, что в случае, если плотность распределения представляет собой сумму с весами функций Гаусса, при выполнении ряда условий оптимальной оценкой будет сумма с весами оценок фильтра Калмана, соответствующих гауссовским функциям распределения [3].

В статье [10] рассматривается алгоритм, основанный на этой теории, с аппроксимацией плотности распределения в виде суммы гауссовских функций. В работе [4] описан вариант этого алгоритма, использующий ансамблевый подход. Предлагаемый алгоритм имеет серьезные проблемы при реализации в случае систем высокой размерности, а также проблемы, связанные с возможной расходимостью фильтра. Авторы [4] рассматривают различные подходы к локализации при применении предлагаемого алгоритма для современных моделей большой размерности.

В статьях [9, 22] предлагаются методы решения задачи нелинейной фильтрации, основанные на объединении ансамблевого фильтра Калмана и фильтра частиц (“bridging the ensemble Kalman and particle filter”). Варианты так называемого гауссовского смешанного фильтра (“Gaussian mixture filter”) обсуждаются в работе [22]. В частности, для предотвращения расходимости фильтра предлагается корректировать весовые коэффициенты при соответствующих частицах с использованием взвешенного среднего исходных коэффициентов и константы, зависящей от количества ансамблей.

Как уже говорилось выше, серьезной проблемой применения фильтра частиц в современных геофизических моделях является высокая размерность системы. В статьях [19, 20] предлагается алгоритм LETKPF (Local Ensemble Transform Kalman Particle Filter). Он представляет собой комбинацию фильтра Калмана и фильтра частиц, при этом при практической реализации используются аналогичные модули алгоритма LETKF [12]. Важным свойством алгоритма LETKF является возможность его использования локально, в связи с этим в настоящее время он очень популярен. Однако алгоритм LETKF представляет собой вариант детерминированного фильтра Калмана [11], поэтому в алгоритме LETKPF требуется дополнительно использовать случайные величины, соответствующие ошибкам наблюдений [19, 20]. Это связано с тем, что элемент ансамбля смешанного гауссовского фильтра имеет вид элемента ансамбля стохастического фильтра Калмана [4, 12].

В настоящей работе рассматривается подход к решению задачи нелинейной фильтрации, основанный на представлении плотности распределения в виде суммы гауссовских функций. При практической реализации алгоритма предлагается использовать ансамблевый  $\pi$ -алгоритм, предложенный в работах [1, 14]. Ансамблевый  $\pi$ -алгоритм представляет собой вариант стохастического ансамблевого фильтра Калмана [11, 15]. Этот алгоритм используется в предлагаемом новом методе для получения ансамбля, соответствующего плотности распределения шага анализа. Ансамблевый  $\pi$ -алгоритм является локальным, что позволяет его эффективно реализовать в случае геофизических моделей высокой размерности. В отличие от подходов к локализации в фильтре частиц, предложенных в [18], предлагаемый в настоящей статье алгоритм использует подходы к локализации, применяемые в ансамблевом фильтре Калмана.

В пункте 1 статьи рассматривается постановка задачи усвоения данных, основанная на байесовском подходе. В п. 2 приводится описание алгоритма оптимальной фильтрации, использующего представление плотности распределения в виде суммы гауссовских

функций. В п. 3 предлагается реализация ансамблевого алгоритма оптимальной нелинейной фильтрации. В п. 4 приводятся результаты тестовых численных экспериментов с нелинейной моделью Lorenz-96. Пункт 5 содержит заключительные выводы.

## 2. Байесовский подход к задаче усвоения данных

Предположим, что математическая модель интересующего нас процесса может быть представлена в виде динамико-стохастической системы:

$$x^{k+1} = f_{k+1,k}(x^k) + \eta^k,$$

где  $x^k$  — вектор размерности  $n$ ,  $k$  — номер шага по времени. Предположим также, что известны данные наблюдений  $y^k$ , представляющие собой вектор размерности  $d$ :

$$y^k = h_k(x^k) + \varepsilon^k.$$

В этих формулах  $\eta^k$  и  $\varepsilon^k$  — случайные ошибки прогноза и наблюдений соответственно,  $f_{k+1,k}$  — оператор модели,  $h_k$  — оператор, переводящий переменные, вычисленные по модели, в наблюдаемые.  $\varepsilon^k$  и  $\eta^k$  гауссовские случайные величины с нулевым математическим ожиданием:  $E[\varepsilon^k (\varepsilon^k)^\top] = R_k$ ,  $E[\eta^k (\eta^k)^\top] = Q_k$ . Символом  $E$  обозначена операция теоретико-вероятностного осреднения. Переменные  $\varepsilon^k$  и  $\eta^k$  не коррелируют между собой а также в различные моменты времени.

Так как предполагается, что прогноз по модели и данные наблюдений являются случайными переменными, они могут быть описаны в терминах плотности распределения. Рассмотрим байесовский подход, состоящий в применении теоремы Байеса для получения оптимальной оценки по данным наблюдений и прогнозу:

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)},$$

где  $p(x|y)$  — условная плотность распределения искомой оценки (анализа),  $p(y|x)$  — условная плотность распределения данных наблюдений,  $p(x)$  есть априорная плотность распределения,  $p(y)$  может рассматриваться как нормализующая константа. Следует отметить, что здесь и далее для всех плотностей распределения используется обозначение “ $p$ ”. Какая конкретно плотность имеется в виду, следует из используемых аргументов.

Существуют следующие методы оценки состояния по данным наблюдений и прогнозу:  $p(x_l|y_{k,1})$ ,  $k > l$ , — прогноз,  $p(x_k|y_{k,1})$  — фильтрация и  $p(x_{k,0}|y_{k,1})$  или оценка  $p(x_l|y_{k,1})$ ,  $0 \leq l \leq k$ , — сглаживание, при этом  $x_{k,0} = \{x_k, x_{k-1}, \dots, x_0\}$ ,  $y_{k,1} = \{y_k, \dots, y_1\}$ . Формулировки и обозначения взяты из обзорной работы [6].

В линейном гауссовском случае решением задачи фильтрации является фильтр Калмана, решением задачи сглаживания — сглаживание Калмана [13]. Для решения задач фильтрации и сглаживания в настоящее время применяется ансамблевый подход, это так называемые ансамблевые алгоритмы фильтрации и сглаживания [7, 8]. В ансамблевом фильтре Калмана в нелинейном случае ковариации прогнозируются с помощью нелинейной модели, при этом условие гауссовости нарушается. Кроме того, в этом случае оценка на шаге анализа будет приближенной и представляет собой линейную оценку наименьших квадратов (linear variance minimizing) [13]. В нелинейном негауссовском случае может быть использован фильтр частиц, который также основан на байесовском подходе [17, 23, 24]. В следующем пункте рассматривается алгоритм усвоения данных,

называемый гауссовский смешанный фильтр (“Gaussian mixture filter”) [22], который является результатом объединения положительных сторон ансамблевого фильтра Калмана и фильтра частиц.

### 3. Алгоритм оптимальной фильтрации, основанный на представлении плотности распределения в виде суммы гауссовских функций

Предположим, что плотность распределения интересующего нас процесса представляет собой сумму гауссовских функций [3]. Пусть плотность распределения на шаге прогноза имеет вид

$$p(x) = \sum_{i=1}^L \alpha_i g(x - m_i, P_i),$$

где  $g$  есть функция Гаусса,  $m_i$  — векторы размерности  $n$  и  $P_i$  — положительно определенные матрицы,  $L$  и коэффициенты  $\{\alpha_i, i = 1, \dots, L\}$   $\sum_{i=1}^L \alpha_i = 1$  заданы. Из теории, изложенной в монографии [3] известно, что в этом случае теоретико-вероятностное среднее значение и ковариационная матрица случайного вектора  $x$  есть, соответственно,

$$\hat{x} = \sum_{i=1}^L \alpha_i m_i, \quad (1)$$

$$\hat{P} = \sum_{i=1}^L \alpha_i \left[ P_i + (\hat{x} - m_i)(\hat{x} - m_i)^\top \right]. \quad (2)$$

Алгоритм усвоения данных обычно состоит из шага прогноза по модели и шага анализа (оценки искомого значения по данным наблюдений и прогнозу).

В монографии [3] показано, что если прогноз  $x_k$  и данные наблюдений  $y_k$  известны на  $k$ -м шаге по времени, то условную плотность распределения можно представить в виде

$$p(x_k | y_k) = \sum_{i=1}^L \tilde{\alpha}_{i,k} g(x_k - m_a^{i,k}, P_a^{i,k}), \quad (3)$$

где

$$m_a^{i,k} = m_f^{i,k} + K_{i,k} [y_k - h(m_f^{i,k})], \quad (4)$$

$$K_{i,k} = P_f^{i,k} H_{i,k}^\top (H_{i,k} P_f^{i,k} H_{i,k}^\top + R_{i,k})^{-1}, \quad H_{i,k}^\top = \left. \frac{\partial h(x)}{\partial x} \right|_{x=m_f^{i,k}} \quad (5)$$

$$P_a^{i,k} = (I - K_{i,k} H_{i,k}) P_f^{i,k}, \quad (6)$$

$$\tilde{\alpha}_{i,k} = \alpha_{i,k} \frac{g[y_k - h(x_f^{i,k}), H_k P_f^{i,k} H_k^\top + P_k]}{\sum_{l=1}^L g[y_k - h(x_f^{l,k}), H_k P_f^{l,k} H_k^\top + P_k]}. \quad (7)$$

$\left( H_{i,k}^\top = \left. \frac{\partial h(x)}{\partial x} \right|_{x=m_f^{i,k}} \right)$ : это означает, что компонент  $(i, j)$  матрицы  $H_{i,k}^\top$  есть частная производная по  $x_j$   $i$ -го компонента  $h(\cdot)$ .

Таким образом, оценка (1) представляет собой сумму с весами значений, полученных в  $L$  алгоритмах обобщенного фильтра Калмана (4)–(6).

На шаге прогноза имеем

$$p(x_{k+1}|y_k) = \sum_{i=1}^L \tilde{\alpha}_{i,k} g[x_{k+1} - f(m_a^{i,k}), F_{i,k} P_a^{i,k} F_{i,k}^\top + Q_k], \quad (8)$$

где  $f$  есть нелинейный оператор модели,  $F_{i,k} = \left. \frac{\partial f(x)}{\partial x} \right|_{x=m_a^{i,k}}$ .

В [3] показано, что это соотношение имеет место при условии, что норма матрицы  $P_f^{i,k}$  уменьшается со временем.

Применение алгоритма усвоения данных, основанного на общем представлении плотности распределения связано с дополнительной проблемой оценки весовых коэффициентов в сумме гауссовских функций. Будем рассматривать подход, позволяющий использовать упрощенную версию аппроксимации плотности распределения. В статье [10] предложен алгоритм, основанный на приведенных выше формулах, при этом рассматривается упрощенная версия аппроксимации плотности распределения: для заданной выборки из  $N$  независимых случайных величин с плотностью распределения  $p$ , оценка плотности распределения  $\hat{p}$  может быть представлена в виде суммы  $N$  гауссовских функций (индекс  $k$  опущен):

$$\hat{p}(x) = \frac{1}{N} \sum_{i=1}^N g(x - x_i, P), \quad (9)$$

где  $P$  — это положительно определенная матрица. Матрица  $P$  может быть приближенно оценена по выборке  $\{x_i, i=1, \dots, N\}$ :  $\hat{P} = \gamma^2 \tilde{P}$ , где  $\tilde{P} = \frac{1}{N-1} X X^\top$ ,  $X = \{x_1 - \bar{x}_1, \dots, x_N - \bar{x}_N\}$ ,  $\bar{x}_i = \frac{1}{N} \sum_{i=1}^N x_i$ , и  $\gamma$  есть параметр, подбираемый эмпирически [10, 22]. Если плотность распределения на шаге прогноза имеет вид (9), на шаге анализа, с использованием данных наблюдений  $y$ , плотность распределения будет описываться формулами (4)–(7), где  $P_f = \hat{P}$ ,  $m_f^i = x_i$ .

Далее, на шаге прогноза, алгоритм, предложенный в [10], реализуется на основе соотношения (8).

Серьезной проблемой этой версии фильтра частиц является его расходимость со временем, а именно, один из весовых коэффициентов  $\alpha_i$  становится близким к 1 и остальные коэффициенты приближаются к нулю. В этом случае прибегают к процедуре “перезапуска” (resampling), при которой выборка генерируется в соответствии с плотностью распределения, и после шага прогноза алгоритм соответствует начальному шагу, а ковариационная матрица приближенно вычисляется по выборке. В работе [17] рассматривается версия алгоритма, в котором такая процедура производится после каждого шага анализа.

#### 4. Практическая реализация алгоритма нелинейной фильтрации, основанного на ансамблевом приближении

Рассмотрим ансамблевый алгоритм нелинейной фильтрации, основанный на разложении (9). При этом ковариации ошибок прогноза вычисляются по ансамблю прогнозов, а при реализации шага анализа применяется ансамблевый  $\pi$ -алгоритм [1, 14].

#### 4.1. Ансамблевый $\pi$ -алгоритм

Для осуществления практической реализации рассматриваемого алгоритма используется вариант стохастического ансамблевого фильтра Калмана в форме квадратного корня (ансамблевый  $\pi$ -алгоритм), предложенный в работах [1, 14]. Этот алгоритм обладает свойством локальности: ансамбль ошибок анализа вычисляется для каждого узла сетки независимо. Все арифметические операции производятся с матрицами, порядок которых равен размерности ансамбля.

Ансамблевый  $\pi$ -алгоритм представляет собой стохастический фильтр, в котором шаг анализа производится только для среднего по ансамблю значения:

$$\bar{x}^a = \bar{x}^f + \frac{1}{N-1} X^a (HX^a)^\top R^{-1} [y - \bar{h}(x^f)]. \quad (10)$$

Здесь  $x^a$  — оценка ансамблевого фильтра Калмана на шаге анализа. Матрица (ансамбль) возмущений на шаге анализа вычисляется через матрицу возмущений ошибок прогноза с использованием  $N \times N$  матрицы трансформаций

$$(X^a)^\top = (I + \Pi^\top)^{-1} (X^f)^\top, \quad (11)$$

где

$$\Pi^\top = (C + 0, 25I)^{\frac{1}{2}} - 0, 5I,$$

$$C = \frac{1}{N-1} (X^f)^\top H^\top R^{-1} (HX^f + S).$$

Матрицы  $X^{a,f}$  содержат отклонения  $x^{a,f}$  от среднего по ансамблю значения,  $S$  есть матрица, столбцами которой являются векторы  $\varepsilon_k^n$  — ансамбль ошибок наблюдений,  $I$  — единичная матрица. Более детальное описание содержится в работах [1, 14].

Элементы матрицы  $\Pi$  вычисляются для заданных матриц  $H$  и  $R$  (оператор наблюдений и ковариационная матрица ошибок наблюдений) для каждого узла сетки независимо. Следовательно, алгоритм может применяться локально в заданных подобластях. Следует отметить, что таким же свойством локальности обладает популярный алгоритм LETKF [12], являющийся детерминированным фильтром.

#### 4.2. Реализация шага анализа нелинейного алгоритма

Согласно теории оптимальной фильтрации, если апостериорная плотность распределения представлена в виде суммы гауссовских функций, то на шаге анализа она имеет вид [3]:

$$p(x|y) = \sum_{i=1}^N \tilde{\alpha}_i g(x - x_a^i, P_a), \quad (12)$$

где средние значения в функциях Гаусса вычисляются по формулам фильтра Калмана:  $x_a^i = x_f^i + K(y - Hx_f^i)$ ,  $K = P_a H^\top R^{-1}$ ,  $P_a = (I - KH)P_f$ . Коэффициенты  $\tilde{\alpha}_i$  определяются по формуле (7) ( $\alpha_i = 1$ ).

Ансамбль оценок на шаге анализа  $x_a^i$  вычисляется по формуле шага анализа фильтра Калмана для среднего по ансамблю (10). При этом вместо средних значений  $\overline{x^f}$  и  $\overline{x^a}$  используются средние значения гауссовских функций в сумме (12). Так как рассматривается вариант алгоритма, в котором матрица  $P_f$  в сумме (9) не зависит от  $i$ , ковариационная матрица  $P_a$  тоже не зависит от  $i$  и может быть вычислена по выборке  $X^a = \{dx_1^a, \dots, dx_N^a\}$ :  $P_a = \frac{1}{N-1} X^a (X^a)^\top$ .

Для осуществления шага прогноза требуется генерировать выборку (ансамбль анализов), соответствующую плотности распределения (12). В статье [5] предлагается генерировать количество элементов выборки, соответствующее слагаемому “ $i$ ”, пропорциональное весовому коэффициенту  $\tilde{\alpha}_i$ ;  $j$ -й элемент выборки, соответствующий слагаемому “ $i$ ” в сумме (12), может быть представлен в виде [4]:

$$x_{i,j}^a = x_{i,j}^f + K(y + \varepsilon_j - Hx_{i,j}^f),$$

при этом отклонение от  $i$ -го среднего имеет вид

$$dx_j^a = dx_j^f + K(\varepsilon_j - Hdx_j^f).$$

Как видно из этих формул, возмущения  $dx_j^a$  удовлетворяют уравнению для возмущений ансамбля в стохастическом фильтре Калмана [11, 15]. Уравнение для  $X^a = \{dx_1^a, \dots, dx_N^a\}$  то же самое, что и для ансамблевого  $\pi$ -алгоритма [1, 14], так как, как уже отмечалось выше, матрица  $P_f$  в сумме (9) не зависит от  $i$ . Следовательно, матрица  $X^a$  может быть вычислена по формуле (11).

Таким образом, при генерировании ансамбля анализов количество элементов ансамбля, соответствующее  $i$ -й функции Гаусса, пропорционально весовому коэффициенту  $\tilde{\alpha}_i$ . При этом элементы ансамбля вычисляются добавлением к  $x_a^i$  нужного количества возмущений  $dx_j^a$ .

### 4.3. Алгоритм приближенного решения задачи нелинейной фильтрации

Перечислим основные этапы, из которых состоит процедура реализации предлагаемого алгоритма. Пусть в начальный момент времени  $t = 0$  задана выборка (ансамбль), соответствующая ковариационной матрице  $\frac{1}{1+\gamma^2} P_0$ , где  $P_0$  — известная ковариационная матрица ошибок первого приближения (прогноза),  $\gamma$  — это эмпирически подбираемый параметр [10]. Нормализация вводится, поскольку, если считать, что выборка соответствует гауссовскому разложению, то ковариация будет вычисляться по формуле (2).

**Шаг прогноза.** Предположим, что после шага прогноза плотность распределения имеет вид  $p(x|y) = \frac{1}{N} \sum_{i=1}^N g(x - x_f^i, P_f)$ , где  $P_f = \gamma^2 \hat{P}$ ,  $\hat{P}$  — это оценка ковариации по ансамблю прогнозов.

**Шаг анализа.** Шаг анализа состоит из следующих процедур:

- 1) вычисление ансамбля возмущений  $dx_j^a$  по ансамблю возмущений  $dx_j^f$ ;
- 2) вычисление матрицы  $P_a$  и  $N$  анализов “средних” для соответствующих гауссовских функций  $\{x_a^i, i = 1, \dots, N\}$  используя ансамбль возмущений  $dx_j^a$ ;
- 3) вычисление выборки  $x_{i,j}^a$ , соответствующей плотности распределения из формулы (12).

По полученному ансамблю анализов производится ансамбль прогнозов по модели и

повторятся шаг 1. В данной работе рассматривается алгоритм с “перезапуском” (resampling) после каждого шага анализа. То есть после шага прогноза алгоритм возвращается к шагу 1, при этом ковариационная матрица вычисляется по выборке.

Предлагаемый алгоритм представляет собой приближенное решение задачи нелинейной фильтрации, обладает свойством локальности, а также является эффективным. Локальность алгоритма обусловлена свойствами  $\pi$ -алгоритма: шаг анализа может производиться для каждого шага сетки независимо. Алгоритм является эффективным, так как арифметические операции на шаге анализа производятся с матрицами порядка размера ансамбля.

Легко видеть, что алгоритм становится обычным ансамблевым фильтром Калмана при  $\gamma = 1$  и если в формуле (12) для всех  $i$   $x_i^a = \bar{x}^a$ ,  $\tilde{\alpha}_i = \frac{1}{N}$ .

#### 4.4. Вычисление весовых коэффициентов

Для предотвращения расходимости алгоритма рассмотрен метод локализации (декомпозиции). Он состоит в разбиении всей области на подобласти и проведение анализа в каждой подобласти независимо. Отметим, что при таком подходе возникает проблема сглаживания на границах подобластей. Такой подход предлагается, например, для фильтра частиц в работе [18]. Важным отличием применяемого в настоящей работе алгоритма является использование гауссовского смешанного фильтра, основанного на оценке плотности распределения случайной ошибки прогноза с помощью суммы гауссовских функций. В этом случае может применяться подход к локализации, принятый при практической реализации ансамблевого фильтра Калмана. Так как свойства ансамблевого алгоритма, используемого на шаге анализа, близки к свойствам алгоритма LETKF, рассматривается используемая в этом алгоритме так называемая  $R$ -локализации, состоящая в поэлементном умножении матрицы  $R$  на функцию расстояния от точки наблюдения до узла сетки [11].

Как отмечено выше, на шаге анализа используется ансамблевый  $\pi$ -алгоритм [1, 14]. Это локальный алгоритм, т. е. анализ для заданного узла сетки может осуществляться независимо, если используются одни и те же данные наблюдений. Пусть вычислительная область разделена на подобласти и предположим, что для анализа в подобласти используются одни и те же данные наблюдений. Множества данных наблюдений, используемые в разных подобластях, не пересекаются. Пусть  $H = [H_1, \dots, H_M]$ , где  $H_m$ ,  $m = 1, \dots, M$ , есть оператор наблюдений для анализа в  $m$ -й подобласти. После шага анализа вычисляются весовые коэффициенты по формуле (индекс  $k$  опущен):

$$\tilde{\alpha}_i = \alpha_i \frac{g[y - h(x_f^i), HP_f^i H^\top + R]}{\sum_{n=1}^N g[y - h(x_f^n), HP_f^n H^\top + R]}.$$

Рассмотрим подход к вычислению весовых коэффициентов, основанный на декомпозиции вычислительной области. Пусть

$$\tilde{\alpha}_i = \text{const} \times \alpha_i \exp \left[ -0.5(y - h(x_f^i))^\top (HP_f^i H^\top + R)^{-1} (y - h(x_f^i)) \right].$$

Предположим, что матрица ошибок наблюдений блочно-диагональная:

$$R = \begin{bmatrix} R_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & R_M \end{bmatrix}.$$

Тогда весовые коэффициенты для шага анализа в  $m$ -й подобласти можно представить в виде

$$\tilde{\alpha}_i^{(m)} = \text{const} \times \alpha_i^{(m)} \exp \left[ -0.5 (y_m - H_m x_f^i)^\top (H_m P_f^i H_m^\top + R_m)^{-1} (y_m - H_m x_f^i) \right],$$

где  $y_m$  — это подвектор вектора данных  $y$ , состоящий из данных, используемых в  $m$ -й подобласти. В этом случае коэффициенты могут быть вычислены локально на шаге анализа для подобласти.

## 5. Численные эксперименты с моделью Lorenz-96

### 5.1. Модель Lorenz-96

Были проведены численные эксперименты по апробации предложенного алгоритма с нелинейной моделью Lorenz-96 [16]. 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 систем, обладающая рядом свойств атмосферных моделей. Для решения системы уравнений используется конечно-разностная схема Рунге-Кутты четвертого порядка. Шаг по времени задавался равным  $\delta t = 0.05$ , что соответствует 6 часам ( $t = 1$  для 5 суток),  $F_0 = 8$ . В алгоритмах оценивания в качестве модели используется дискретный аналог этого уравнения для расчета изменений по времени вектора размерности  $J$ .

С моделью Lorenz-96 проводились численные эксперименты с модельными данными. Для этого моделировалось состояние, считавшееся “истиной”. Для моделирования “истины” в экспериментах по усвоению данных были проведены вычисления по заданному начальному значению  $x_0^t \cong N(F_0/4; F_0/2)$  на  $N_t$  шагов по времени [16]. Через  $N(a, b)$  обозначена случайная величина, распределенная по нормальному закону с математическим ожиданием  $a$  и дисперсией  $b$ . В численных экспериментах результат прогноза на  $N_t = 1000$  шагов по времени брался в качестве начального значения для прогноза “истины”.

### 5.2. Численные эксперименты

Начальное значение прогноза и данные наблюдений моделировались добавлением случайной ошибки к “истинному” значению. В качестве критерия точности прогноза с усвоением рассматривалось среднеквадратическое отклонение от “истины”. Начальные данные для шага прогноза задавались добавлением возмущения к “истинным” начальным данным:  $x_d(0) = x_t(0) + \delta$ ,  $\delta \sim N(0, s_0)$ .

Для проведения численных экспериментов были заданы следующие величины: ансамбль начальных полей  $x^n(0) = x_d(0) + \delta^n$ ,  $\delta^n \sim N(0, s_0)$ ,  $n = 1, \dots, N_{\text{ens}}$ ; данные наблюдений  $y_0^n = x_t(0) + \delta_0^n$ ,  $\delta_0^n \sim N(0, \varepsilon_0)$ ; а также ансамбль возмущений данных наблюдений  $\delta_0^n \sim N(0, \varepsilon_0)$ ,  $n = 1, \dots, N_{\text{ens}}$ . В численных экспериментах предполагалось, что “шум модели”  $\eta^n = 0$ , при моделировании “истины”  $\eta^t \sim N(0, s_1)$ ,  $s_1 = 0.01$ . Через  $N_{\text{ens}}$  обозначено количество элементов ансамбля. Предполагается, что данные наблюдений известны во

всей вычислительной области. Прогноз производился на  $N_t = 3000$  шагов по времени. В численных экспериментах задавались следующие значения параметров:  $s_0 = \varepsilon_0 = 1$ .

Во всех численных экспериментах  $R = \varepsilon_0^2 I$ . Вычислительная область разделялась на подобласти размером в 5 узлов сетки. Для анализа внутри подобласти брались все данные из этой подобласти. При реализации алгоритма использовалась процедура  $R$  — локализации [11]. При применении такой локализации вместо матрицы  $R$  используется матрица  $\tilde{R} = R \circ e^{-0.5(\rho_{ik}/bc)^2}$ , где  $\rho_{ik}$  — это расстояние между узлом сетки и наблюдением, через “ $\circ$ ” обозначена операция поэлементного умножения. Во всех экспериментах брались значения  $id = 5$ ,  $bc = 5\Delta x$  ( $\Delta x$  — шаг сетки).

Во всех экспериментах параметр  $\gamma$  в формуле для оценки  $P_f$  задавался равным 0.8. Это значения брались на основе рекомендаций, приведенных в статье [22].

В работе [22] отмечается, что ковариационная матрица, соответствующая плотности распределения (12) имеет вид  $(1 + \gamma^2)\hat{P}$ , где  $\hat{P}$  — это матрица, вычисленная по ансамблю. Для того, чтобы дисперсии соответствовали дисперсиям, вычисленным по выборке, может быть использовано следующее преобразование ансамбля анализов:

$$x_a^{i'} = \beta x_a^i + (1 - \beta)N^{-1} \sum_{i=1}^N x_a^i \quad (13)$$

вместо  $x_a^i$ , где  $\beta = \sqrt{1 - \gamma^2}$  (предполагается, что  $\gamma < 1$ ). В численных экспериментах элементы ансамбля анализов корректировались по формуле (13).

Предложенный выше алгоритм является вариантом фильтра частиц и, следовательно, может расходиться со временем. Для предотвращения расходимости применялась модификация коэффициентов, предложенная в работе [22]:

$$\hat{\alpha}_i = \tilde{\alpha}_i \times w + (1 - w) \times N^{-1}, \quad (14)$$

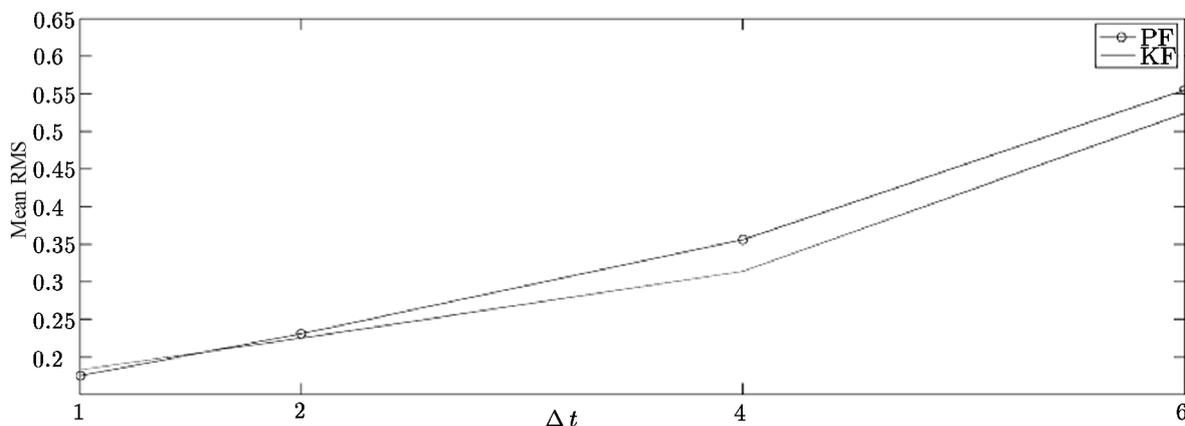
где  $w$  — это параметр, подбираемый эмпирически.

Для выбора оптимального значения  $w$  проводились численные эксперименты с предложенным алгоритмом и различными значениями параметра  $w$  на  $N_t = 3000$  шагов по времени, при этом сравнивались среднеквадратические ошибки прогноза с усвоением. В этих экспериментах бралось количество элементов ансамбля  $N_{\text{ens}} = 50$ , усвоение данных осуществлялось на каждом шаге по времени. Рассматривались следующие значения коэффициента  $w$ : 0, 0.3, 0.5 и 0.7. Лучшие результаты получились при выборе  $w = 0.3$ . Это значение параметра  $w$  было использовано в последующих экспериментах для модификации весовых коэффициентов алгоритма с помощью формулы (14).

Предлагаемый алгоритм сравнивался с ансамблевым фильтром Калмана. При этом под фильтром Калмана подразумевается алгоритм, описанный в п. 4.3. В ансамблевых алгоритмах при практической реализации используется прием для предотвращения быстрого убывания следа ковариационной матрицы, состоящий в умножении элементов ансамбля на константу, называемую “inflation factor” [11]. При реализации алгоритма фильтра Калмана использовалось значение “inflation factor” равное 1.2. В фильтре частиц “inflation factor” не использовался.

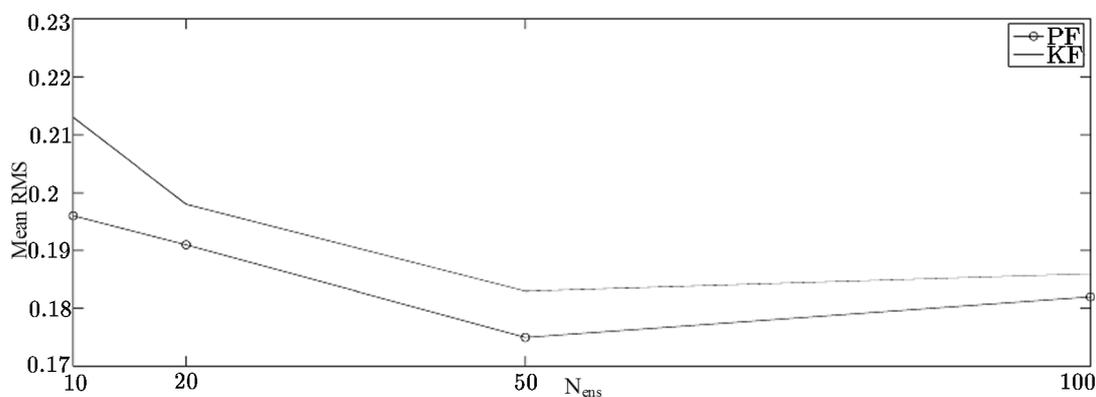
Для оценки свойств предлагаемого алгоритма проводились численные эксперименты с усвоением с частотой 1, 2, 4, 6 шагов по времени (т.е. на каждом первом шаге по времени, втором и т.д.). Эксперименты проводились для  $N_{\text{ens}} = 50$ .

На рис. 1 приводится среднее по времени значение среднеквадратической ошибки для нового алгоритма (фильтр частиц) (PF) и фильтра Калмана (KF). Как видно из рисунка, фильтр частиц имеет преимущество только при усвоении на каждом шаге по времени, что согласуется с результатами, приведенными в [18].



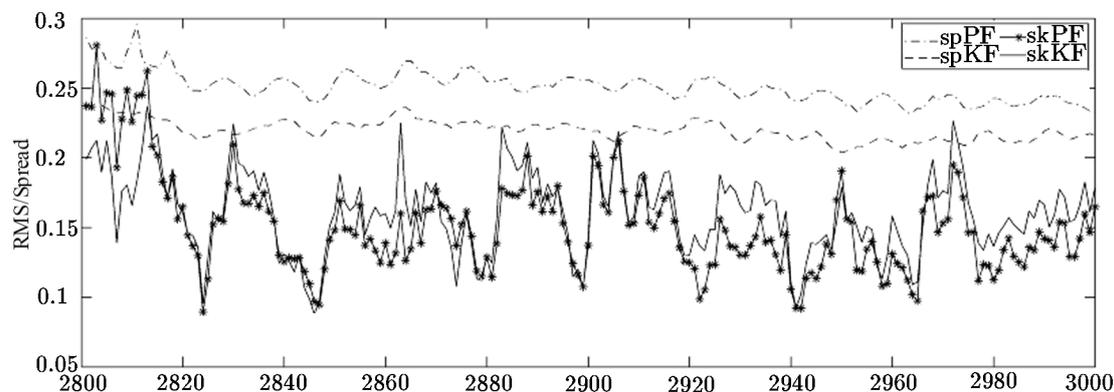
**Рис. 1.** Средняя по времени среднеквадратическая ошибка прогноза с усвоением для ансамблевого фильтра Калмана и фильтра частиц. Количество элементов ансамбля 50

На рис. 2 приводится среднее по времени значение среднеквадратической ошибки при усвоении на каждом шаге по времени для количества элементов ансамбля  $N_{\text{ens}}$  равного 10, 20, 50 и 100. Как видно из рисунка, фильтр частиц имеет преимущество по сравнению с фильтром Калмана (от 2 до 7%) во всех случаях.



**Рис. 2.** Средняя по времени среднеквадратическая ошибка прогноза с усвоением на каждом шаге по времени для ансамблевого фильтра Калмана и фильтра частиц

На рис. 3 приводится среднеквадратическая ошибка (rms) и след ковариационной матрицы (sp) при усвоении на каждом шаге по времени с использованием фильтра частиц (sk PF, sp PF) и фильтра Калмана (sk KF, sp KF). След ковариационной матрицы представляет собой теоретическую ошибку оценивания, в то время как среднеквадратическая ошибка характеризует реальное значение ошибки модельного эксперимента. Эксперименты проводились для количества элементов ансамбля  $N_{\text{ens}} = 100$ . Оценки приводятся для временного интервала с шагами по времени и от 2800 до 3000. Как видно из рис. 3, новый алгоритм имеет преимущество по сравнению с фильтром Калмана на ряде временных интервалов.



**Рис. 3.** Среднеквадратическая ошибка и след ковариационной матрицы для ансамблевого фильтра Калмана и фильтра частиц. Количество элементов ансамбля 100

Следует отметить, что рассматриваемая в данной работе задача зависит от многих параметров, таких как  $\gamma$ ,  $w$ ,  $N_{\text{ens}}$ , частота усвоения по времени, распределение данных наблюдений, параметры локализации и т. д. Проведенные в работе численные эксперименты показали, что при выборе определенных значений параметров алгоритм имеет преимущество по сравнению с фильтром Калмана. Понятно, что выбираемые параметры задачи зависят от конкретной практической проблемы.

## 6. Заключение

В статье предлагается новый алгоритм усвоения данных, который применим в случае нелинейной модели. Алгоритм основан на представлении плотности распределения в виде суммы гауссовских функций и применении локального стохастического фильтра Калмана с трансформацией ансамбля возмущений (ансамблевого  $\pi$ -алгоритма). Приводятся результаты численных экспериментов по оценке свойств алгоритма с одномерной нелинейной моделью. Показано, что в ряде случаев предлагаемый алгоритм является более точным, чем ансамблевый фильтр Калмана. Однако точность алгоритма зависит от таких параметров, как количество элементов ансамбля, частота усвоения и ряда других.

Целью данной работы было показать, что для решения задачи нелинейной фильтрации может быть использован разработанный ранее ансамблевый  $\pi$ -алгоритм. При этом предлагаемый алгоритм имеет важное свойство локальности и может применяться в современных геофизических моделях. В настоящей работе основное внимание было уделено реализации шага анализа гауссовского смешанного фильтра. Отметим, что алгоритм может быть обобщен на случай плотности распределения более общего вида.

Таким образом, можно выделить следующие преимущества предложенного алгоритма:

- 1) шаг анализа алгоритма может проводиться с использованием разработанных для ансамблевых фильтров технологий, включая методы локализации;
- 2) существует версия алгоритма, являющаяся ансамблевым фильтром Калмана;
- 3) алгоритм может быть обобщен на случай плотности распределения более общего вида.

## Литература

1. **Климова Е.Г.** Стохастический ансамблевый фильтр Калмана с трансформацией ансамбля возмущений // Сиб. журн. вычисл. математики / РАН. Сиб. отд-ние. — Новосибирск, 2019. — Т. 22, № 1. — С. 27–40. Перевод: Klimova E.G. The Kalman stochastic ensemble filter with transformation of perturbation ensemble // Numerical Analysis and Applications. — 2019. — Vol. 12, № 1. — P. 26–36.
2. **Anderson D.L., Anderson S.L.** A Monte-Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts // Monthly Weather Review. — 2019. — Vol. 127. — P. 2741–2758.
3. **Anderson B.D., Moore J.B.** Optimal filtering. — New Jersey: Prentice-Hall, 1979.
4. **Bengtsson T., Snyder C., Nychka D.** Toward a nonlinear ensemble filter for high-dimensional systems // J. Geophysical Research. — 2003. — <https://doi.org/10.1029/2002JD002900>.
5. **Carpenter J., Clifford P., Fearnhead P.** Improved particle filter for nonlinear problems // IEE Proc. Radar, Sonar Navig. — 1999. — <https://doi.org/10.1049/ip-rsn:19990255>.
6. **Carrassi A., Bocquet M., Bertino L., Evensen G.** Data assimilation in the geosciences: An overview of methods, issues and perspectives // Wiley interdisciplinary reviews: Climate Change. — 2018. — <https://doi.org/10.1002/wcc535>.
7. **Evensen G.** Data assimilation. The ensemble Kalman filter. — Berlin, Heidelberg: Springer-Verlag, 2009.
8. **Evensen G.** Analysis of iterative ensemble smoother for solving inverse problems // Computational Geosciences. — 2019. — Vol. 22. — P. 885–908. — <https://doi.org/10.1007/s10596-018-9731-y>.
9. **Frei M., Kunsch H.R.** Bridging the ensemble Kalman and particle filter // Biometrika. — 2013. — Vol. 100, iss. 4. — P. 781–800. — <https://doi.org/10.1093/biomet/ast020>.
10. **Hoteit I., Pham D.-T., Triantafyllou G., Korres G.** A new approximate solution of the optimal nonlinear filter for data assimilation in meteorology and oceanography // Monthly Weather Review. — 2008. — Vol. 138. — P. 317–334.
11. **Houtekamer, H.L. Zhang, F.** Review of the ensemble Kalman filter for atmospheric data assimilation // Monthly Weather Review. — 2016. — Vol. 144. — P. 4489–4532.
12. **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.
13. **Jazwinski, A.H.** Stochastic processes and filtering theory. — New York: Academic Press, 1970.
14. **Klimova E.** A suboptimal data assimilation algorithm based on the ensemble Kalman filter // Quarterly Journal of the Royal Meteorological Society. — 2012. — Vol. 138. — P. 2079–2085.
15. **Klimova E.G.** Bayesian approach to data assimilation based on ensembles of forecasts and observations // IOP Conf. Series: Earth and Environmental Science. — 2019. — <https://doi.org/10.1088/1755-1315/386/1/012038>.
16. **Lorenz E.N., Emanuel K.A.** Optimal sites for supplementary weather observations: simulation with a small model // Journal of the Atmospheric Sciences. — 1998. — Vol. 55. P. 319–414.
17. **Nakamura G., Potthast R.** Inverse Modeling. — 2015. — <https://doi.org/10.1088/978-0-7503-1218-9>.
18. **Poterjoy J.** A localized particle filter for high-dimensional nonlinear systems // Monthly Weather Review. — 2016. — Vol. 144. — P. 59–76. — <https://doi.org/10.1175/MWR-D-15-0163.1>.
19. **Potthast R., Walter A., Rhodin A.** A localized adaptive particle filter within an operational NWP Framework // Monthly Weather Review. — 2019. — Vol. 147. — P. 345–362. — <https://doi.org/10.1175/MWR-D-18-0028.1>.

20. **Robert S., Leuenberger D., Kunsch H.R.** A local ensemble transform Kalman particle filter for convective scale data assimilation // Quarterly Journal of the Royal Meteorological Society. — 2018. — Vol. 144. — P. 1279–1296. — <https://doi.org/10.1002/qj.3116>.
21. **Snyder C., Bengtsson T., Morzfeld M.** Performance bounds for particle filters using the optimal proposal // Monthly Weather Review. — 2015. — Vol. 143. — P. 4750–4761. — <https://doi.org/10.1175/MWR-D-15-0144.1>.
22. **Stordal A.S., Karlsen H.A., Nævdal G., Skaug H.J., Valles B.** Bridging the ensemble Kalman filter and particle filter: the adaptive Gaussian mixture filter // Computational Geosciences. — 2011. — Vol. 15. — P. 293–305. — <https://doi.org/10.1007/s10596-010-9207-1>.
23. **van Leeuwen P.J., Kunsch H.R., Nerger L., Potthast R., Reich S.** Particle filters for high-dimensional geoscience applications: A review // Quarterly Journal of the Royal Meteorological Society. — 2019. — Vol. 145. — P. 2335–2365. — <https://doi.org/10.1002/qj.3551>.
24. **Vetra-Carvalho S., van Leeuwen P.J., Nerger L. et al.** State-of-the-art stochastic data assimilation methods for high-dimensional non-Gaussian problems // Tellus A. — 2018. — Vol. 70. — P. 1–43. — <https://doi.org/10.1080/16000870.2018.1445364>.

*Поступила в редакцию 27 июня 2022 г.*

*После исправления 19 октября 2022 г.*

*Принята к печати 23 ноября 2022 г.*

## Литература в транслитерации

1. **Klimova E.G.** Stokhasticheskii ansamblevyy fil'tr Kalmana s transformaciei ansamblya vozmuschenii // Sib. zhurn. vychisl. matematiki / RAN. Sib. otd-nie. — Novosibirsk, 2019. — Т. 22, № 1. — С. 27–40. Перевод: Klimova E.G. The Kalman stochastic ensemble filter with transformation of perturbation ensemble // Numerical Analysis and Applications. — 2019. — Vol. 12, № 1. — P. 26–36.
2. **Anderson D.L., Anderson S.L.** A Monte-Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts // Monthly Weather Review. — 2019. — Vol. 127. — P. 2741–2758.
3. **Anderson B.D., Moore J.B.** Optimal filtering. — New Jersey: Prentice-Hall, 1979.
4. **Bengtsson T., Snyder C., Nychka D.** Toward a nonlinear ensemble filter for high-dimensional systems // J. Geophysical Research. — 2003. — <https://doi.org/10.1029/2002JD002900>.
5. **Carpenter J., Clifford P., Fearnhead P.** Improved particle filter for nonlinear problems // IEE Proc. Radar, Sonar Navig. — 1999. — <https://doi.org/10.1049/ip-rsn:19990255>.
6. **Carrassi A., Bocquet M., Bertino L., Evensen G.** Data assimilation in the geosciences: An overview of methods, issues and perspectives // Wiley interdisciplinary reviews: Climate Change. — 2018. — <https://doi.org/10.1002/wcc535>.
7. **Evensen G.** Data assimilation. The ensemble Kalman filter. — Berlin, Heidelberg: Springer-Verlag, 2009.
8. **Evensen G.** Analysis of iterative ensemble smoother for solving inverse problems // Computational Geosciences. — 2019. — Vol. 22. — P. 885–908. — <https://doi.org/10.1007/s10596-018-9731-y>.
9. **Frei M., Kunsch H.R.** Bridging the ensemble Kalman and particle filter // Biometrika. — 2013. — Vol. 100, iss. 4. — P. 781–800. — <https://doi.org/10.1093/biomet/ast020>.
10. **Hoteit I., Pham D.-T., Triantafyllou G., Korres G.** A new approximate solution of the optimal nonlinear filter for data assimilation in meteorology and oceanography // Monthly Weather Review. — 2008. — Vol. 138. — P. 317–334.

11. **Houtekamer, H.L., Zhang, F.** Review of the ensemble Kalman filter for atmospheric data assimilation // *Monthly Weather Review*. — 2016. — Vol. 144. — P. 4489–4532.
12. **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.
13. **Jazwinski, A.H.** *Stochastic processes and filtering theory*. — New York: Academic Press, 1970.
14. **Klimova E.** A suboptimal data assimilation algorithm based on the ensemble Kalman filter // *Quarterly Journal of the Royal Meteorological Society*. — 2012. — Vol. 138. — P. 2079–2085.
15. **Klimova E.G.** Bayesian approach to data assimilation based on ensembles of forecasts and observations // *IOP Conf. Series: Earth and Environmental Science*. — 2019. — <https://doi.org/10.1088/1755-1315/386/1/012038>.
16. **Lorenz E.N., Emanuel K.A.** Optimal sites for supplementary weather observations: simulation with a small model // *Journal of the Atmospheric Sciences*. — 1998. — Vol. 55. P. 319–414.
17. **Nakamura G., Potthast R.** Inverse Modeling. — 2015. — <https://doi.org/10.1088/978-0-7503-1218-9>.
18. **Poterjoy J.** A localized particle filter for high-dimensional nonlinear systems // *Monthly Weather Review*. — 2016. — Vol. 144. — P. 59–76. — <https://doi.org/10.1175/MWR-D-15-0163.1>.
19. **Potthast R., Walter A., Rhodin A.** A localized adaptive particle filter within an operational NWP Framework // *Monthly Weather Review*. — 2019. — Vol. 147. — P. 345–362. — <https://doi.org/10.1175/MWR-D-18-0028.1>.
20. **Robert S., Leuenberger D., Kunsch H.R.** A local ensemble transform Kalman particle filter for convective scale data assimilation // *Quarterly Journal of the Royal Meteorological Society*. — 2018. — Vol. 144. — P. 1279–1296. — <https://doi.org/10.1002/qj.3116>.
21. **Snyder C., Bengtsson T., Morzfeld M.** Performance bounds for particle filters using the optimal proposal // *Monthly Weather Review*. — 2015. — Vol. 143. — P. 4750–4761. — <https://doi.org/10.1175/MWR-D-15-0144.1>.
22. **Stordal A.S., Karlsen H.A., Nævdal G., Skaug H.J., Valles B.** Bridging the ensemble Kalman filter and particle filter: the adaptive Gaussian mixture filter // *Computational Geosciences*. — 2011. — Vol. 15. — P. 293–305. — <https://doi.org/10.1007/s10596-010-9207-1>.
23. **van Leeuwen P.J., Kunsch H.R., Nerger L., Potthast R., Reich S.** Particle filters for high-dimensional geoscience applications: A review // *Quarterly Journal of the Royal Meteorological Society*. — 2019. — Vol. 145. — P. 2335–2365. — <https://doi.org/10.1002/qj.3551>.
24. **Vetra-Carvalho S., van Leeuwen P.J., Nerger L. et al.** State-of-the-art stochastic data assimilation methods for high-dimensional non-Gaussian problems // *Tellus A*. — 2018. — Vol. 70. — P. 1–43. — <https://doi.org/10.1080/16000870.2018.1445364>.