«Оптика атмосферы и океана», **23**, № 12 (2010)

УДК 551.501

Пространственная интерполяция метеорологических полей с помощью малопараметрической динамико-стохастической модели с вертикальной компонентой

В.С. Комаров[1], К.Ю. Дубовик[2], Ю.Б. Попов[2], А.В. Лавриненко[1]*

[1]Институт оптики атмосферы им. В.Е. Зуева СО РАН
634021, г. Томск, пл. Академика Зуева, 1
[2]Сургутский государственный университет
628403, г. Сургут, ул. Ленина, 1

Поступила в редакцию 25.02.2010 г.

Рассматривается новый метод пространственной интерполяции полей температуры воздуха и компонентов скорости ветра в нижней тропосфере. Данный метод базируется на алгоритме фильтра Калмана и малопараметрической динамико-стохастической модели с вертикальной компонентой. Рассмотрены результаты качества статистической оценки предложенного алгоритма пространственной интерполяции. Проведено сравнение результатов с алгоритмом, использующим четырехмерную динамико-стохастическую модель.

Ключевые слова: фильтр Калмана, пространственная интерполяция, усвоение данных, численное моделирование, малопараметрическая динамико-стохастическая модель; Kalman filter, spatial interpolation, data assimilation, numerical modeling, low parametrical dynamic-stochastic model.

Введение

В последние годы в практике обработки данных метеорологических наблюдений стала широко применяться процедура усвоения данных, объединяющая в себе решение задач, связанных с прогнозом и объективным анализом метеорологических полей, основанная на одновременном учете как самих измерений, так и результатов прогнозирования по выбранной математической модели. При этом одним из возможных подходов к решению задачи усвоения данных является динамико-стохастический подход, базирующийся на теории фильтрации Калмана [1–3].

В основе динамико-стохастического подхода лежит предположение о том, что состояние атмосферы описывается случайными полями, которые связаны между собой некоторой системой соотношений. В связи с этим процесс пространственной интерполяции при реализации этого подхода состоит из двух этапов:

1) усвоение поступивших метеорологических наблюдений и коррекция параметров математической модели,

2) собственно пространственная интерполяция на основе скорректированной модели.

Следует отметить, что применение фильтра Калмана в процедуре усвоения метеорологических данных, получаемых с помощью современных прогностических моделей, сталкивается с определенными трудностями, вызванными сложностью реализации его алгоритма из-за высокого порядка матриц ковариаций ошибок прогноза [2].

С учетом вышесказанного, а также исходя из необходимости решения задачи пространственной интерполяции для мезометеорологического масштаба нами предлагается упрощенная модель поведения метеорологических параметров в пространстве и во времени с использованием стохастических дифференциальных уравнений. Отличительной особенностью данного подхода является исключение использования процедуры, при которой требуется решение сложной системы гидродинамических уравнений. Кроме того, каждый из интересующих параметров состояния атмосферы рассматривается в отдельности, а его изменения в пространстве и во времени представляют собой стохастический процесс с известными корреляционными свойствами. Согласно такому подходу интерполяция осуществляется в выбранную точку пространства по данным наблюдений только нескольких метеорологических станций. При этом ограничивается размерность вектора состояний, упрощается реализация алгоритма фильтрации и повышается его устойчивость.

Решение проблемы пространственно-временной интерполяции в рамках мезомасштаба диктуется необходимостью метеорологической поддержки различных прикладных задач:

– оценки пространственного распространения техногенных загрязнений на малые расстояния (до 100–200 км) от источников промышленных загрязнений;

* Валерий Сергеевич Комаров; Ксения Юрьевна Дубовик; Юрий Борисович Попов; Андрей Викторович Лавриненко (gfm@iao.ru).

– диагноза и прогноза состояния атмосферы (и в первую очередь температуры и ветра) над неосвещенной в метеорологическом отношении территорией.

Здесь необходимо отметить, что исследования в области оценивания и прогноза параметров состояния атмосферы для мезометеорологического масштаба, осуществляемые на основе алгоритма фильтра Калмана, продолжают ранние работы [3–6], выполненные в Институте оптики атмосферы, когда в основу соответствующих алгоритмов были положены малопараметрические динамико-стохастические модели различных видов, на основе алгебраических полиномов, регрессионные и пр.

Несмотря на повышение точности пространственной интерполяции метеорологических полей в области мезомасштаба, нельзя считать эту задачу полностью решенной.

Предложенная нами модель восходит своими корнями к малопараметрической динамико-стохастической модели на основе двумерного уравнения мезомасштабной диффузии, рассмотренной в работах [3, 4]. Отличием от разработанной ранее модели является существование высотной зависимости между оцениваемым параметром и измерениями. Следует отметить, что в ранних моделях присутствовали пространственная и временная зависимости интерполируемой метеорологической величины. Таким образом, мы пытаемся оценить вклад в результирующую оценку информации, поступающей с соседних высотных уровней. Правильность данной гипотезы была практически подтверждена нашими предыдущими исследованиями, проведенными при разработке четырехмерной динамико-стохастической модели, использованной для пространственной интерполяции метеорологических величин [5, 6].

1. Постановка задачи и особенности предложенной методики

Пусть в заданной мезомасштабной области имеются $s + 1$ точек с координатами x_i, y_i, где $i = 0, ..., s$. В нулевой точке с координатами x_0, y_0 метеорологические наблюдения отсутствуют, а в остальных осуществляются аэрологические измерения. Физическая постановка задачи восстановления метеорологического поля ξ в области мезомасштаба заключается в оценке его значения в нулевой точке по измерениям в i-х точках. Аналогично тому, как было сделано в работах [4, 5], процедура оценки значения случайного поля ξ по данным радиозондовых наблюдений выполняется по двухканальной схеме. Согласно ей результирующая оценка значения поля ξ_0 в точке восстановления с координатами x_0, y_0 складывается из суммы оценок – оценки регулярной составляющей поля $\bar{\xi}_0$ и оценки флуктуационной составляющей ξ_0':

$$\xi_0 = \xi_0' + \bar{\xi}_0. \qquad (1)$$

В качестве оценки регулярной составляющей поля в точке восстановления используется средневзвешенное значение для трех ближайших станций, рассчитываемое на фиксированной высоте h:

$$\bar{\xi}_0 = \sum_{i=1}^{3} q_i \xi_i \Bigg/ \sum_{i=1}^{3} q_i, \qquad (2)$$

где ξ_i – измеренное значение поля в i-й точке; $q_i = 1 - \left(\rho_{i0} \Big/ \sum_{i=1}^{3} \rho_{i0} \right)$ – весовой коэффициент; $\rho_{i0} = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}$ – расстояние между i-й станцией и точкой восстановления, а x и y – прямоугольные координаты.

Для оценки флуктуационной составляющей в качестве базовой математической модели предложена динамико-стохастическая модель, состоящая из двух уравнений:

$$\xi_i'(k+1) = \xi_0'(k)(1 - \alpha\Delta t)(1 - \beta\Delta\rho_{i0})(1 - \gamma\Delta h) + \varepsilon_i(k), \qquad (3)$$

$$\xi_0'(k+1) = \xi_0'(k)(1 - \alpha\Delta t) + \varepsilon_0(k), \qquad (4)$$

где α, β – коэффициенты, определяющие временную и пространственную зависимости между данными аэрологических станций в пределах мезомасштабного полигона, а γ – коэффициент, определяющий межуровневую связь метеорологических полей; k – текущее дискретное время; ε – невязка модели, определяющая стохастичность рассматриваемых атмосферных процессов; Δt – интервал времени между отсчетами.

Уравнение (3) описывает пространственно-временную связь между станциями наблюдения и точкой интерполяции, а (4) – эволюцию поля во времени в точке интерполяции. В силу случайности значений метеорологических величин их статистические свойства могут быть описаны соответствующими корреляционными функциями: $\mu(\tau)$ – во времени, $\mu(\rho)$ – в пространстве, $\mu(h)$ – по высоте. Корреляционные функции, описывающие временную, пространственную и высотную зависимость случайных составляющих метеорологических величин, в том числе температуры и скорости ветра, могут быть описаны экспоненциальными выражениями вида

$$\mu(\tau) = \exp(-\alpha\tau), \quad \mu(\rho) = \exp(-\beta\rho), \quad \mu(h) = \exp(-\gamma h),$$

где $\alpha = 1/\tau_0$, $\beta = 1/\rho_0$, $\gamma = 1/h_0$ – коэффициенты, обратно пропорциональные интервалам временной, пространственной и межуровневой корреляции.

По сравнению с малопараметрической динамико-стохастической моделью на основе двумерного уравнения мезомасштабной диффузии [3, 4] в предлагаемой модели появляется новая составляющая, отвечающая за вертикальные корреляционные связи в пределах рассматриваемого полигона.

Неизвестными параметрами рассматриваемой модели будут измерения в нулевой точке с координатами x_0, y_0. Параметры α, β, γ будем задавать, исходя из предварительного анализа радиозондовых данных, характерных для выбранного мезомасштабного полигона. Это справедливо по причине того, что на рассматриваемых интервалах времени между наблюдениями и в пределах выбранного мезомасштаба (в нашем случае с максимальными размерами 500×500 км) используемые метеорологические поля можно считать однородными и изотропными.

Общим требованием при синтезе алгоритмов оценивания неизвестных параметров динамической системы является возможность их описания с помощью системы дифференциальных или разностных уравнений первого порядка [4], матричная форма записи которых имеет следующий вид:

$$\mathbf{x}_{k+1}^{t} = \boldsymbol{\psi}_{k} \cdot \mathbf{x}_{k}^{t} + \boldsymbol{\omega}_{k}^{t}, \quad (5)$$

где $\mathbf{x}_{k+1}^{t}$ — вектор состояния, включающий в себя неизвестные и подлежащие оценке параметры для момента времени k; $\boldsymbol{\psi}_k$ — матрица перехода для дискретной системы; $\boldsymbol{\omega}_k^t$ — вектор шумов состояния. Стоит отметить, что в нашем случае, когда коэффициенты α, β и γ задаются априорно или оцениваются отдельно от алгоритма фильтра Калмана, вектор состояния включает всего один элемент — искомую метеорологическую величину в точке восстановления с координатами x_0, y_0 на фиксированной высоте h в момент времени k.

Для того чтобы матрица перехода содержала только временную зависимость, а матрица наблюдений — пространственную и межуровневые связи, произведем ряд математических преобразований. При сопоставлении выражений (4) и (5) матрица перехода будет состоять всего из одного элемента и ее вид

$$\psi_k = \|1 - \alpha\Delta t\|. \quad (6)$$

Математическая модель измерений, по данным которых в алгоритме фильтра Калмана проводится оценка состояния системы, в общем случае описывается аддитивной смесью полезного сообщения и ошибки измерения:

$$\mathbf{y}_k^o = \xi_k^o = \mathbf{H}_k \mathbf{x}_k^t + \boldsymbol{\varepsilon}_k^o, \quad (7)$$

где $\mathbf{y}_k^o$ – вектор фактических центрированных измерений размерностью $(m \times 1)$ (здесь и далее для простоты записи штрих опущен); $\mathbf{H}_k$ – матрица наблюдений размерностью $(m \times 1)$, определяющая функциональную связь между истинными значениями переменных состояния и фактическими измерениями; $\boldsymbol{\varepsilon}_k^o$ – вектор ошибок измерений в момент времени k (шум измерений) размерностью $m \times 1$.

В связи с тем что вертикальные корреляционные связи полей метеорологических величин достаточно слабые, при формировании вектора измерений в его состав мы будем включать измерения только на соседних высотных уровнях, лежащих выше и ниже от восстанавливаемого уровня, причем для всех s измерительных станций. Таким образом, он может быть записан в виде

$$\mathbf{y}_k^o = \\ = \left\| y_{11}^o(k), y_{12}^o(k), y_{13}^o(k), y_{21}^o(k), y_{22}^o(k), y_{23}^o(k), \ldots, y_{s3}^o(k) \right\|^{\mathrm{T}}, \quad (8)$$

где первый элемент индекса показывает номер точки наблюдения (станции), второй — номер высотного уровня, участвующего в рассмотрении, а значение в скобках — рассматриваемый момент времени. Следовательно, в векторе измерений последовательно друг за другом записаны укороченные вертикальные профили измеренных метеорологических величин, получаемых со всех точек наблюдения, участвующих в рассмотрении в заданный момент времени, поэтому $m = s \times 3$.

Зададим теперь матрицу наблюдений $\mathbf{H}_k$. Из сопоставления выражений (3) и (8) видно, что элементами матрицы являются значения коэффициентов β и γ, характеризующие пространственные и вертикальные связи полей полигона:

$$\mathbf{H}_k = \left\| \begin{matrix} \gamma_1(1-\beta\rho_{01}) \\ (1-\beta\rho_{01}) \\ \gamma_2(1-\beta\rho_{01}) \\ \gamma_1 \cdot 1-\beta\rho_{02}) \\ \ldots \\ \gamma_2(1-\beta\rho_{0s}) \end{matrix} \right\|, \quad (9)$$

где γ_1 и γ_2 — коэффициенты межуровневой корреляции между нижним (верхним) и фиксированным h уровнем; ρ_{0s} — расстояние между точкой восстановления и станцией с номером s.

В модели была учтена ситуация, когда уровень, на котором происходит восстановление метеорологической величины, является крайним нижним или крайним верхним. Для этого в векторе измерений использовали два верхних или два нижних уровня соответственно. Выражения для подобной ситуации в статье не будут представлены.

После определения всех элементов, входящих в выражения (6) и (8), задачу оценивания сводим к использованию линейного фильтра Калмана, обеспечивающего оценку элементов вектора состояния с минимальными среднеквадратическими ошибками.

В этом случае уравнение оптимального оценивания вектора состояния имеет вид [7]:

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k(\mathbf{y}_k^0 - \mathbf{H} \cdot \mathbf{x}_k^f), \quad (10)$$

где $\mathbf{x}_k^a$ – вектор проанализированных значений (оценка вектора состояния) в момент времени k; $\mathbf{x}_k^f$ – вектор предсказанных значений в момент времени k; $\mathbf{K}_k$ – матрица весовых коэффициентов. Расчет весовых коэффициентов в линейном фильтре Калмана осуществляется по рекуррентным матричным уравнениям:

$$\mathbf{K}_k = \mathbf{P}_k^f \mathbf{H}^{\mathrm{T}} (\mathbf{H}\mathbf{P}_k^f \mathbf{H}^{\mathrm{T}} + \mathbf{R}_k)^{-1},$$

$$\mathbf{P}_k^f = \mathbf{F}_{k-1}\mathbf{P}_{k-1}^a \mathbf{F}_{k-1}^{\mathrm{T}} + \mathbf{Q}_{k-1}, \quad \mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k\mathbf{H})\mathbf{P}_k^f,$$

где $\mathbf{P}_k^a = <(\mathbf{x}_k^a - \mathbf{x}_k^t)(\mathbf{x}_k^a - \mathbf{x}_k^t)^{\mathrm{T}}>$ – ковариационная матрица ошибок оценивания; $\mathbf{P}_k^f = <(\mathbf{x}_k^f - \mathbf{x}_k^t)(\mathbf{x}_k^f - \mathbf{x}_k^t)^{\mathrm{T}}>$ –

ковариационная матрица ошибок предсказания; $\mathbf{F}_{k-1} = \dfrac{\partial(\boldsymbol{\psi}_{k-1}(\mathbf{x}^a_{k-1}))}{\partial \mathbf{x}^a_{k-1}}$ — матрица Якоби от вектор-функции $\boldsymbol{\psi}_{k-1}(\mathbf{x}^a_{k-1})$; $\mathbf{I}$ — единичная матрица.

2. Результаты апробации разработанного алгоритма

Описанная выше методика была использована в задаче пространственной интерполяции мезомасштабных полей температуры и ветра. С этой целью были взяты массивы данных наблюдений за 2009 г. на пяти аэрологических станциях: Москва, Рязань, Сухиничи, Смоленск и Курск. При этом в качестве контрольных точек, в которые проводилась пространственная интерполяция, использовались, как и в [5, 6], станции Москва и Смоленск. Кроме того, оценка предложенного алгоритма проводилась применительно к решению задачи численного анализа пространственного распространения облака загрязняющих веществ, когда в качестве данных температуры и ветра берутся не уровневые, а усредненные по высоте значения, т.е. $<\xi>_{h_0,h}$ (здесь h_0 — высота земной поверхности, h — высота верхней границы этого слоя). Поэтому, как в работах [5, 6], в качестве исходных данных нами взяты средние в слое значения температуры $<T>_{h_0,h}$, зональной $<U>_{h_0,h}$ и меридиональной $<V>_{h_0,h}$ составляющих скорости ветра, рассчитанные для слоев от 0 до 8 км по данным на изобарических поверхностях и уровнях особых точек.

Для реализации алгоритма калмановской фильтрации, использующего динамико-стохастическую модель с вертикальной компонентой, были заданы следующие начальные условия: $\alpha = 0{,}9$, $1/\beta = 1200$, $\gamma_1 = 1$, $\gamma_2 = 1$ для температуры воздуха, а для ортогональных компонентов скорости ветра $\alpha = 0{,}3$, $1/\beta = 700$, $\gamma_1 = 0{,}5$, $\gamma_2 = 0{,}5$. В качестве начальных значений параметров $\mathbf{x}^a_0$ и $\mathbf{P}^a_0$ берутся значения $\mathbf{x}^a_0 = 0$ и $\mathbf{P}^a_0 = 10$. Диагональные элементы матрицы ковариаций шумов наблюдений $\mathbf{R}_k = \sigma^2_\xi \mathbf{I}$ задаются, исходя из величин ошибок радиозондовых измерений, заимствованных из [7]: для температуры $\sigma_\xi = 1$ °C, для ортогональных составляющих скорости ветра $\sigma_\xi = 1$ м/с. Матрица ковариаций шумов состояния $\mathbf{Q}_k = \mathbf{I}$ задается единичной.

Рассмотрим теперь собственно результаты статистической оценки качества и эффективности предложенного алгоритма в задаче пространственной интерполяции полей температуры воздуха и ортогональных компонент ветра. Для этого проведем сравнение нашей модели и четырехмерной динамико-стохастической модели, подробно описанной в работах [5, 6]. При этом результаты статистических оценок для сравниваемых моделей получены для идентичных выборок исходных данных.

На рис. 1 и 2 в качестве примера представлены зависимости среднеквадратических погрешностей пространственной интерполяции (экстраполяции) температуры воздуха и скорости зонального и меридионального ветра, полученные для лета и зимы на контрольных станциях Москва и Смоленск на основе предложенного алгоритма.

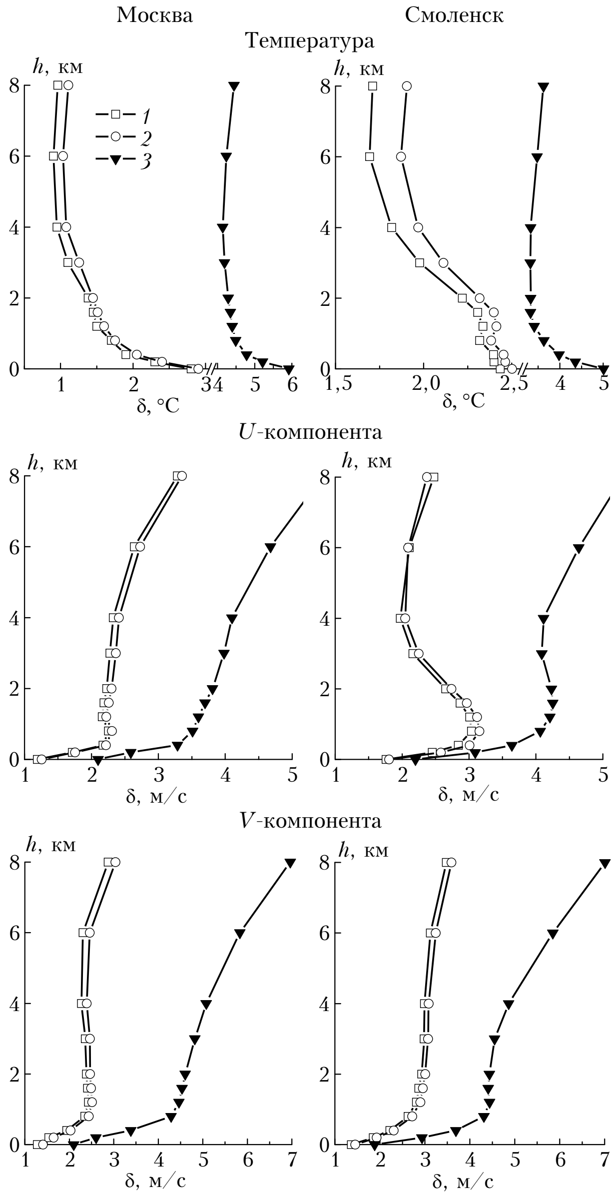

Рис. 1. Зависимости среднеквадратических погрешностей δ пространственной интерполяции средних в слое значений метеорологических полей от высоты для лета на основе малопараметрической динамико-стохастической модели с вертикальной компонентой (*1*), четырехмерной динамико-стохастической модели (*2*) и соответствующие среднеквадратические отклонения (*3*)

Для сравнения приводятся эти же зависимости для случая, когда используется четырехмерная динамико-стохастическая модель регрессионного типа.

Анализ рис. 1 и 2 показывает, что предложенный алгоритм может быть с успехом использован на практике, поскольку среднеквадратические погрешности пространственной интерполяции T, U и V варьируют независимо от высотного уровня атмосферы и сезона в пределах 0,8–2,5 °C (для температуры воздуха) и 1,0–4,5 м/с (для ортогональных составляющих скорости ветра).

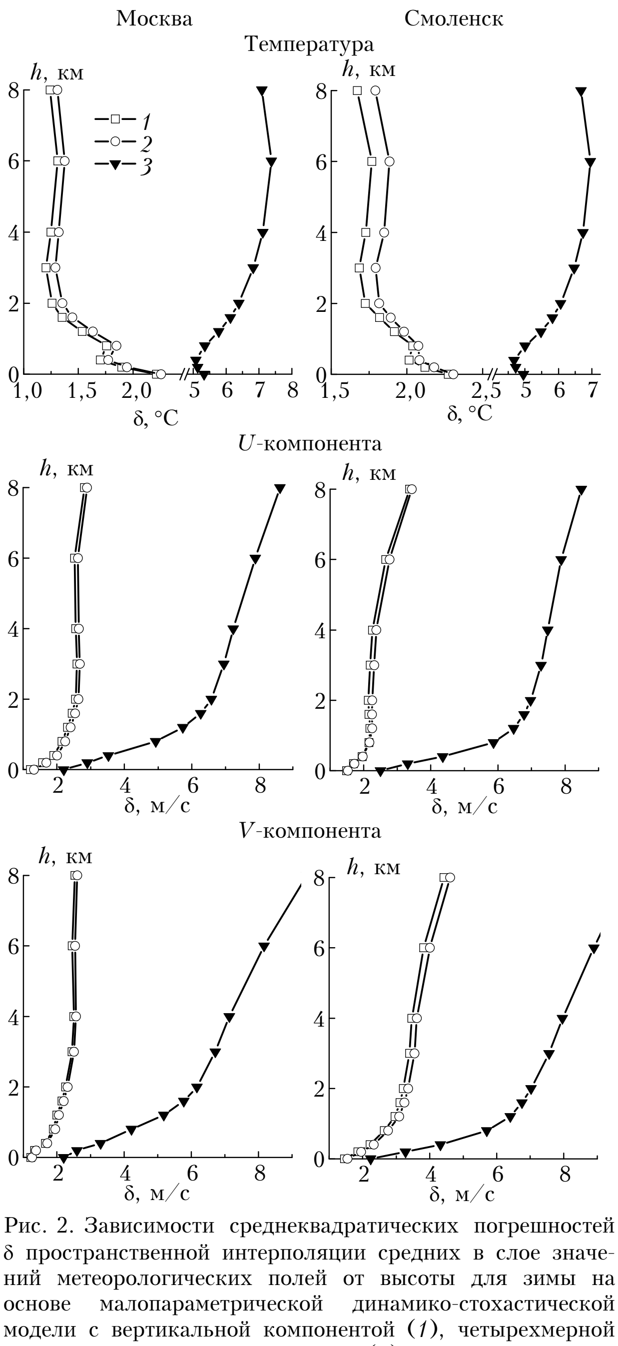

Рис. 2. Зависимости среднеквадратических погрешностей δ пространственной интерполяции средних в слое значений метеорологических полей от высоты для зимы на основе малопараметрической динамико-стохастической модели с вертикальной компонентой (*1*), четырехмерной динамико-стохастической модели (*2*) и соответствующие среднеквадратические отклонения (*3*)

Кроме того, алгоритм дает небольшой, но устойчивый выигрыш по точности перед алгоритмом на основе четырехмерной модели, когда независимо от слоя атмосферы он составляет от 1,1 до 1,3 раза для температуры воздуха и от 1,1 до 1,2 раза для компонентов скорости ветра.

Таким образом, качественная оценка алгоритма фильтра Калмана с малопараметрической динамико-стохастической моделью с вертикальной компонентой показывает, что данный алгоритм является достаточно эффективным и может быть с успехом применен при решении различных прикладных задач.

В дальнейшем предполагается провести оценку параметров α, β и γ на основе оценок радиусов пространственной, межуровневой и временной корреляции для заданного мезомасштабного полигона в темпе поступления исходной информации со станций наблюдения. Эта гипотеза нашла свое подтверждение в результате тестирования модели при варьировании набора коэффициентов α, β и γ для различных метеорологических величин и высот, и были выявлены определенные закономерности.

1. *Важник А.И., Свиренко П.И., Фролов А.В., Цветков В.И., Веселова Г.К., Булдовский Г.С.* Глобальная схема дискретного четырехмерного усвоения данных Всемирной службы погоды и результаты ее оперативных испытаний: Информ. сб. Гидрометеорол. научно-исслед. центра РФ, № 28. СПб.: Гидрометеоиздат, 2001. С. 3–17.
2. *Климова Е.Г.* Методика усвоения данных метеонаблюдений на основе обобщенного субоптимального фильтра Калмана // Метеорол. и гидрол. 1997. № 11. С. 55–65.
3. *Комаров В.С., Попов Ю.Б., Суворов С.С., Кураков В.А.* Динамико-стохастические методы и их применение в прикладной метеорологии. Томск: Изд-во ИОА СО РАН, 2004. 236 с.
4. *Komarov V.S., Il'in S.N., Kreminskii A.V., Lomakina N.Ya., Popov Y.B., Popova A.I., Suvorov S.S.* Estimation and Extrapolation of the Atmospheric State Parameters on the Mesoscale Level Using a Filter Algorithm // J. Atmos. and Ocean Technol. 2004. V. 21, N 3. P. 488–494.
5. *Комаров В.С., Лавриненко А.В., Попов Ю.Б., Ломакина Н.Я., Попова А.И., Ильин С.Н.* Пространственная экстраполяция метеорологических полей в области мезомасштаба на основе четырехмерной смешанной динамико-стохастической модели и аппарата калмановской фильтрации // Оптика атмосф. и океана. 2004. Т. 17, № 8. С. 651–656.
6. *Komarov V.S., Lavrinenko A.V., Lomakina N.Ya., Popov Yu.B., Popova A.I.* New Method of Spatial Extrapolation of Meteorological Fields on the Mesoscale Level Using a Kalman Filter Algorithm for a Four-Dimentional Dynamic-Stochastic Model // J. Atmos. and Opt. Technol. 2007. V. 24, N 2. P. 182–193.
7. *Сейдж Э.Р., Мэлса Дж.Л.* Теория оценивания и ее применение в связи и управлении. М.: Связь, 1976. 496 с.

V.S. Komarov, K.Yu. Dubovik, Yu.B. Popov, A.V. Lavrinenko. **Spatial interpolation of meteorological fields with the use of low-parametrical dynamic-stochastic model with a vertical component.**

A new method of spatial interpolation of fields of air temperature and wind velocity components in low troposphere is considered. The method is based on the linear Kalman filter algorithm and a low parametrical dynamic-stochastic model with a vertical component. The quality of statistical estimates of the suggested algorithm of spatial interpolation is considered. The results are compared with the algorithm, using 4D-dynamic stochastic model.