

В.С. Комаров, А.В. Лавриненко, Н.Я. Ломакина, Ю.Б. Попов, А.И. Попова,
С.Н. Ильин

Пространственная экстраполяция метеорологических полей в области мезомасштаба на основе четырехмерной смешанной динамико-стохастической модели и аппарата калмановской фильтрации

Институт оптики атмосферы СО РАН, г. Томск

Поступила в редакцию 27.04.2004 г.

Рассматриваются методика и алгоритм пространственной экстраполяции мезометеорологических полей на основе четырехмерной смешанной динамико-стохастической модели и аппарата калмановской фильтрации. Обсуждаются результаты статистической оценки качества предложенного алгоритма при его использовании в задаче пространственной экстраполяции мезомасштабных полей температуры и ветра на не освещенную данными наблюдений территорию.

Введение

Одной из основных проблем современной мезометеорологии можно назвать проблему, связанную с оценкой и прогнозом состояния атмосферы над не освещенной данными наблюдений территориями по результатам измерений в прилегающих районах. Длительное время подобная проблема решалась в рамках объективного анализа метеорологических полей, проводимого на основе метода оптимальной интерполяции [1, 2].

В последние годы в связи с увеличением потока и видов метеорологической информации традиционная процедура объективного анализа стала вытесняться процедурой четырехмерного усвоения данных. Такая процедура объединяет в едином контуре две традиционно различные задачи — объективный анализ и прогнозирование метеорологических полей. Обычно в качестве прогностической модели используют совокупность уравнений гидротермодинамики (гидродинамическая модель). Недостатком такого подхода является то, что прогностическая модель гидродинамического типа используется только в качестве временного экстраполанта, причем в ходе прогнозирования не производится уточнение параметров этой модели на следующем шаге по времени.

В настоящей статье предлагается иной подход к решению задачи пространственной экстраполяции метеорологических полей в области мезомасштаба, который базируется на использовании алгоритма фильтра Калмана и четырехмерной смешанной динамико-стохастической модели. Такая модель позволяет описать изменение метеорологического поля одновременно в пространстве и во времени без при-

влечения большого объема оперативной информации о состоянии атмосферы. В то же время применение фильтра Калмана обеспечивает оценку и пошаговую коррекцию параметров прогностической модели в темпе поступления совокупности измерений от локальной сети аэрологических станций.

Данная статья является продолжением наших ранних публикаций [3, 4], где в основу подхода закладывалась линейная малопараметрическая динамико-стохастическая модель на базе дифференциальных уравнений первого порядка.

Динамико-стохастический подход, примененный в данной статье, как и в [3, 4], направлен на снижение порядка матриц ковариации ошибок оценивания и прогноза за счет упрощения прогностической модели поведения метеополя в пространстве и во времени, т.е. уменьшения размерности вектора состояния. Подобные вопросы упрощения расчетов рассмотрены также в публикациях [5–7]. При этом для прогностической гидродинамической модели предлагается путь предварительного вычисления матриц ковариаций с использованием аппроксимирующих функций. Результаты вычислений соответствующих матриц в этом случае зависят только от начальных данных. Такое решение задачи, с математической точки зрения, не адекватно классическому подходу оптимальной фильтрации Калмана.

1. Постановка задачи и метод ее решения

Задача пространственной экстраполяции централизованного метеорологического поля ξ' в области мезомасштаба заключается в оценке его значения в точке с прямоугольными координатами (x_0, y_0, z_0)

по измерениям в точках с координатами (x_i, y_i, z_i) ($i = 1, 2, 3, \dots, n$) и некоторой математической модели, описывающей изменения этого поля в пространстве и времени. В нашем случае в качестве такой модели используется, как уже говорилось выше, четырехмерная смешанная динамико-стохастическая модель вида

$$\xi'_{i,h}(k) = \sum_{j=1}^K a_j \xi'_{i,h}(k-j) + \sum_{m=1}^M b_m \xi'_{i,m}(k) + \sum_{s=1}^S \frac{c_s \xi'_{s,h}(k)}{\rho_{is}}, \quad (1)$$

где $k = 0, 1, 2, \dots$ – дискретное текущее время с интервалом дискретизации Δt , ($t_k = k\Delta t$); K – порядок запаздывания по времени, который определяет глубину авторегрессии; M – количество высотных уровней, участвующих в формировании оценки поля ξ' ; S – количество пунктов наблюдения; a_j , b_m и c_s – неизвестные и подлежащие оцениванию параметры, определяющие временную, высотную и пространственную зависимость между значениями измерений поля в разные моменты дискретного времени k , на различных уровнях и в разных точках мезомасштабного полигона соответственно; $\rho_{is} = \rho_0 / (\rho_0 - R_{is})$ – нормирующий коэффициент, который определяет взаимное расположение точек наблюдения на плоскости в пределах мезомасштабного полигона и отражает наличие пространственной корреляции между ними (здесь ρ_0 – радиус пространственной корреляции и $R_{is} = \sqrt{(x_i - x_s)^2 + (y_i - y_s)^2}$ – расстояние между точками i и s , км).

В соответствии с [1, 8] в планетарном пограничном слое (при $h < 1,5$ км) радиус пространственной корреляции $\rho_0 = 2000$ км (для температуры) и $\rho_0 = 750$ км (для ортогональных составляющих скорости ветра), а для свободной атмосферы он соответственно равняется 2500 и 1000 км.

Из выражения (1) следует, что значение поля ξ' в любой точке полигона на выбранном высотном уровне будет параметрическим образом зависеть: от значений этого поля в предыдущие отсчеты времени на глубину K (связанного с радиусом временной корреляции); от его значений на всех высотных уровнях вплоть до уровня M в момент времени k (момент наблюдения) и от всех измеренных значений поля на выбранном уровне h в других точках наблюдения.

Поскольку в выражении (1) в качестве исходного поля используется центрированное поле ξ' , то, прежде чем переходить к рассмотрению методики его оценивания в некоторой точке пространства с координатами (x_0, y_0) , следует остановиться на процедуре такого центрирования. С этой целью поле ξ представим в виде суммы регулярной $\bar{\xi}$ и флуктуационной ξ' составляющих, т.е. $\xi = \bar{\xi} + \xi'$.

Для оценки регулярной составляющей поля $\bar{\xi}$ в i -х точках заданного мезомасштабного полигона,

т.е. в точках измерений, и на h -м уровне используется среднее полигонное значение этого поля, определяемое из выражения

$$\bar{\xi}_n^{(h)} = \frac{1}{S} \sum_{i=1}^S \xi_i^{(h)}, \quad (2)$$

где $\xi_i^{(h)}$ – измеренное значение метеорологического поля на i -й станции и на h -м высотном уровне.

В то же время для оценки регулярной составляющей поля $\bar{\xi}$ в точке экстраполяции (x_0, y_0) на заданном высотном уровне h используется средневзвешенное значение, рассчитываемое по данным измерений трех ближайших (к точке экстраполяции) станций с использованием выражения вида

$$\bar{\xi}_0^{(h)} = \frac{\sum_{i=1}^3 q_i \xi_i^{(h)}}{\sum_{i=1}^3 q_i}, \quad (3)$$

где $\xi_i^{(h)}$ – измеренное значение поля ξ в i -й точке (или станции) и на h -м высотном уровне;

$q_i = 1 - (R_{i0} / \sum_{j=1}^3 R_{j0})$ – весовой коэффициент

(здесь R_{i0} – расстояние i -й станции от точки экстраполяции (x_0, y_0) , x и y – прямоугольные координаты станции).

Для получения значений центрированного поля ξ'_i в точках измерения на каждом шаге оценки ξ_i (здесь и далее для простоты записи индекс « h » опущен) можно воспользоваться выражением

$$\xi'_i = \xi_i - \bar{\xi}_n, \quad (4)$$

а для оценки поля ξ в точке экстраполяции (x_0, y_0) , проведенной с помощью выбранного алгоритма, можно использовать выражение вида

$$\xi_0 = \xi'_0 + \bar{\xi}_0. \quad (5)$$

Рассмотрим теперь собственно методику пространственной экстраполяции, основанную на использовании фильтра Калмана и модели вида (1).

Задача оценки поля ξ' в некоторой точке с координатами (x_0, y_0) в случае применения четырехмерной разностной динамико-стохастической модели распадается на два этапа. На первом этапе по значениям поля ξ' , полученным в точках измерений, производится оценка коэффициентов модели a_j , b_m , c_s , которые вследствие предположения об однородности и изотропности поля для заданного мезомасштабного полигона будут постоянными в смысле среднего. На втором этапе, по оцененным коэффициентам на основе математической модели (1), восстанавливаются значения центрированного поля в заданной точке пространства и на различных высотных уровнях.

В соответствии с [9] для оценивания неизвестных параметров модели (1), т.е. a_j , b_m , c_s , необхо-

димом задать систему разностных уравнений в матричном виде:

$$\mathbf{X}(k+1) = \mathbf{F}(k) \cdot \mathbf{X}(k) + \mathbf{\Omega}(k), \quad (6)$$

где $\mathbf{X}(k+1) = |a_1(k+1), a_2(k+1), \dots, a_K(k+1); b_1(k+1), b_2(k+1), \dots, b_M(k+1); c_1(k+1), c_2(k+1), \dots, c_S(k+1)|^T = |X_1(k+1), X_2(k+1), \dots, X_K(k+1), \dots, X_{K+M}(k+1), \dots, X_{K+M+S}(k+1)|^T$ – вектор-столбец размерностью $(n \times 1) = (K + M + S) \times 1$, включающий в себя неизвестные и подлежащие оцениванию переменные состояния динамической системы (вектор состояний); \mathbf{T} – оператор транспонирования; $\mathbf{F}(k)$ – матрица перехода для дискретной системы размерностью $(n \times n)$; $\mathbf{\Omega}(k) = |\omega_1, \omega_2, \dots, \omega_n|^T$ – вектор-столбец случайных возмущений системы (вектор шумов состояния).

Если предположить, что рассматриваемое метеорологическое поле изотропно и стационарно, а на заданном интервале времени неизвестные и подлежащие оцениванию параметры $\mathbf{X}(k)$ в среднем не изменяются, то

$$\mathbf{X}(k+1) = \mathbf{X}(k), \quad (7)$$

или

$$\mathbf{X}(k+1) = \mathbf{F}(k) \mathbf{X}(k). \quad (8)$$

Таким образом, матрица перехода $\mathbf{F}(k)$ для нашего случая будет соответствовать единичной матрице \mathbf{I} размерностью $(n \times n)$:

$$\mathbf{F}(k) = \mathbf{I} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}. \quad (9)$$

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

$$\mathbf{Y}'(k) = \xi'(k) = \mathbf{H}(k) \cdot \mathbf{X}(k) + \mathbf{E}(k), \quad (10)$$

где $\mathbf{Y}'(k)$ – вектор фактических измерений размерностью $(s \times 1)$; $\mathbf{H}(k)$ – матрица наблюдений $(s \times n)$, определяющая функциональную связь между истинными значениями переменных состояния и фактическими измерениями; $\mathbf{E}(k)$ – вектор ошибок измерений (шум измерений).

Вектор фактических измерений $\mathbf{Y}'(k)$ в момент времени k включает в себя всю совокупность измеренных значений интересующего нас метеорологического поля на S станциях и всех высотных уровнях. Этот вектор может быть записан в виде

$$\mathbf{Y}'(k) = |Y'_{11}(k), Y'_{12}(k), Y'_{13}(k), \dots, Y'_{1,M}(k), Y'_{21}, Y'_{22}(k), \dots, Y'_{2,M}(k), \dots, Y'_{S,M}(k)|^T, \quad (11)$$

где первый элемент индекса показывает номер точки наблюдения (станции), а второй – номер высотного уровня, участвующего в рассмотрении. Таким образом, в векторе измерений последовательно друг за другом записаны высотные профили измерений со всех точек наблюдения, участвующих в рассмотрении, при этом $s = S \times M$.

Зададим матрицу наблюдений $\mathbf{H}(k)$. При сопоставлении выражений для базисной функции (1) и математической модели измерений (10) становится очевидным, что элементами матрицы $\mathbf{H}(k)$ являются значения измерений поля ξ' в точках наблюдения, на всех высотах и в данный и в предыдущие моменты времени (на глубину K). Матрица $\mathbf{H}(k)$ имеет структуру, состоящую из трех блоков. Первый блок имеет размерность $(n \times K)$ и включает значения метеовеличины для предыдущих отсчетов времени на глубину K . По мере изменения дискретного времени k и поступления новых данных от измерительных станций, элементы матрицы $\mathbf{H}(k)$ пошагово перемещаются внутри первого блока слева направо, тем самым образуя скользящее окно процесса авторегрессии с эффективной шириной K .

Вторая часть матрицы $\mathbf{H}(k)$ имеет размерность $n \times (M + 1)$ и представляет собой блоки угловых матриц, верхняя часть которых заполнена нулями, а нижняя заполнена значениями поля, измеренными на соседних уровнях. Третья часть матрицы $\mathbf{H}(k)$ $(n \times S)$ заполняется значениями поля, измеренными во всех точках наблюдения в текущий момент времени k .

После определения всех элементов, входящих в выражения (6) и (10), задача оценивания решается с помощью линейного фильтра Калмана, обеспечивающего оценку элементов вектора состояния с минимальными среднеквадратическими ошибками.

При этом уравнения оценивания имеют следующий вид [9]:

$$\hat{\mathbf{X}}(k+1) = \hat{\mathbf{X}}(k+1|k) + \mathbf{G}(k+1) \cdot [\mathbf{Y}(k+1) - \mathbf{H}(k+1) \cdot \hat{\mathbf{X}}(k+1|k)], \quad (12)$$

где

$$\hat{\mathbf{X}}(k+1) = |\hat{X}_1(k+1), \hat{X}_2(k+1), \dots, \hat{X}_n(k+1)|^T$$

– оценка вектора состояния на момент времени $(k+1)$, $\hat{\mathbf{X}}(k+1|k)$ – вектор предсказанных оценок на момент времени $(k+1)$ по данным на шаге k ; $\mathbf{G}(k+1)$ – матрица весовых коэффициентов размерностью $(n \times s)$.

Отметим, что векторы $\hat{\mathbf{X}}(k+1)$ и $\hat{\mathbf{X}}(k+1|k)$ имеют размерность $(n \times 1)$, а матричное уравнение для расчета вектора предсказания

$$\hat{\mathbf{X}}(k+1|k) = \mathbf{F}(k) \cdot \hat{\mathbf{X}}(k). \quad (13)$$

Расчет весовых коэффициентов в фильтре Калмана производят с помощью стандартных уравнений вычисления ковариации ошибок оценивания [9].

Для начала работы алгоритма фильтрации (12) в условиях отсутствия априорной информации следует начальные значения коэффициентов задать равными $a_i = 0$, $b_i = 0$, $c_i = 0$. Таким образом, $\hat{\mathbf{X}}(0) = 0$. Другие начальные условия, связанные с оценкой элементов матриц шумов состояния и наблюдений, задаются исходя из известных значений стандартных отклонений и ошибок радиозондовых данных.

Таким образом, по мере обработки данных измерений в моменты времени k , с использованием уравнений (12) и заданных начальных условий, формируется оценка вектора состояния

$$\hat{\mathbf{X}}(k) = \left[\hat{a}_1, \hat{a}_2, \dots, \hat{a}_K, \hat{b}_1, \hat{b}_2, \dots, \hat{b}_M, \hat{c}_1, \hat{c}_2, \dots, \hat{c}_S \right], \quad (14)$$

с помощью которой осуществляется непосредственное восстановление метеорологического поля ξ в точке экстраполяции на основе выражения вида

$$\hat{\xi}_0(k) = \bar{\xi}(k) + \hat{\mathbf{Y}}'_0(k) = \bar{\xi} + \hat{\mathbf{H}}(k) \times \hat{\mathbf{X}}(k), \quad (15)$$

где $\hat{\xi}_0(k)$ и $\bar{\xi}$ – векторы восстановленных в точке (x_0, y_0) и средневзвешенных значений метеовеличины; $\hat{\mathbf{X}}(k)$ – полученная оценка вектора состояния в момент k ; $\hat{\mathbf{Y}}'_0(k)$ – вектор оценок значений флуктуационной составляющей поля в точке экстраполяции; $\hat{\mathbf{H}}(k)$ – матрица перехода размерностью $(M \times n)$ для восстановления того же поля в точке (x_0, y_0) .

2. Результаты исследований алгоритма фильтра Калмана

Рассмотренный выше алгоритм, базирующийся на использовании четырехмерной динамико-стохастической модели и фильтра Калмана, был подвержен качественной оценке при использовании в задаче пространственной экстраполяции мезомасштабных полей температуры и ветра.

Поскольку пространственная экстраполяция в настоящей статье рассматривается применительно к прогнозу распространения облака загрязнения веществ промышленного происхождения, то нами были взяты не данные измерений температуры и ветра на отдельных уровнях, а их средние в слое значения в некотором интервале высот $h_k - h_0$, где $h_0 = 0$ совпадает с уровнем земной поверхности, а h_k – высота верхней границы исследуемого k -го слоя атмосферы. При этом расчет средних в слое (или просто средних) значений температуры, зональной и меридиональной составляющих скорости ветра проводился с помощью выражения

$$\langle \xi \rangle_{h_k - h_0} = \sum_{i=1}^k \left\{ \left(\frac{\xi_{h_{i-1}} + \xi_{h_i}}{2} \right) \left(\frac{h_i - h_{i-1}}{h_k} \right) \right\}, \quad (16)$$

где $\langle \cdot \rangle$ обозначает процедуру осреднения данных наблюдений в некотором слое атмосферы, а ξ – измеренное значение метеорологической величины на различных атмосферных уровнях.

Для оценки качества алгоритма калмановской фильтрации был использован архив данных двухлетних (2000–2001 гг.) и двухсрочных (0 и 12 ч по Гринвичу) радиозондовых наблюдений восьми аэрологических станций: Москва (55°45' с.ш., 37°57' в.д.), Смоленск (54°45' с.ш., 32°04' в.д.), Бологое (57°54' с.ш., 34°03' в.д.), Вологда (59°19' с.ш., 39°55' в.д.), Нижний Новгород (56°16' с.ш., 44°00' в.д.), Рязань (54°38' с.ш., 39°42' в.д.), Сухиничи (54°06' с.ш., 35°21' в.д.), Курск (51°46' с.ш., 36°10' в.д.), представляющих собой типичный мезомасштабный полигон (рис. 1). При этом все данные наблюдений за температурой и ветром, представленные для зимы и лета на стандартных изобарических поверхностях и уровнях особых точек, были приведены с помощью линейной интерполяции к единой системе следующих геометрических высот: 0 (уровень земной поверхности); 0,2; 0,4; 0,8; 1,2; 1,6; 2,0; 3,0; 4,0; 5,0; 6,0 и 8,0 км. Такая система геометрических высот позволяет описать почти всю тропосферу и особенно пограничный слой с большим вертикальным разрешением.

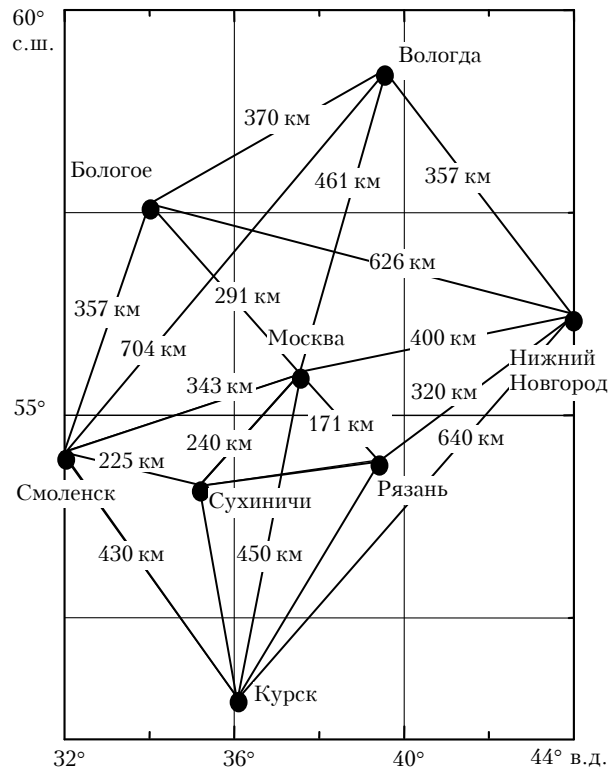


Рис. 1. Схема мезометеорологического полигона

Для оценки точности алгоритма фильтра Калмана (ее результаты приведены на рис. 2) в качестве контрольной точки (в нее осуществлялась пространственная экстраполяция) была использована ст. Смоленск, находящаяся на расстоянии 225 км от ближайшей ст. Сухиничи.

Анализ рис. 2 показывает, что предложенный алгоритм, основанный на методе калмановской фильтрации и четырехмерной динамико-стохастической модели и примененный для пространственной

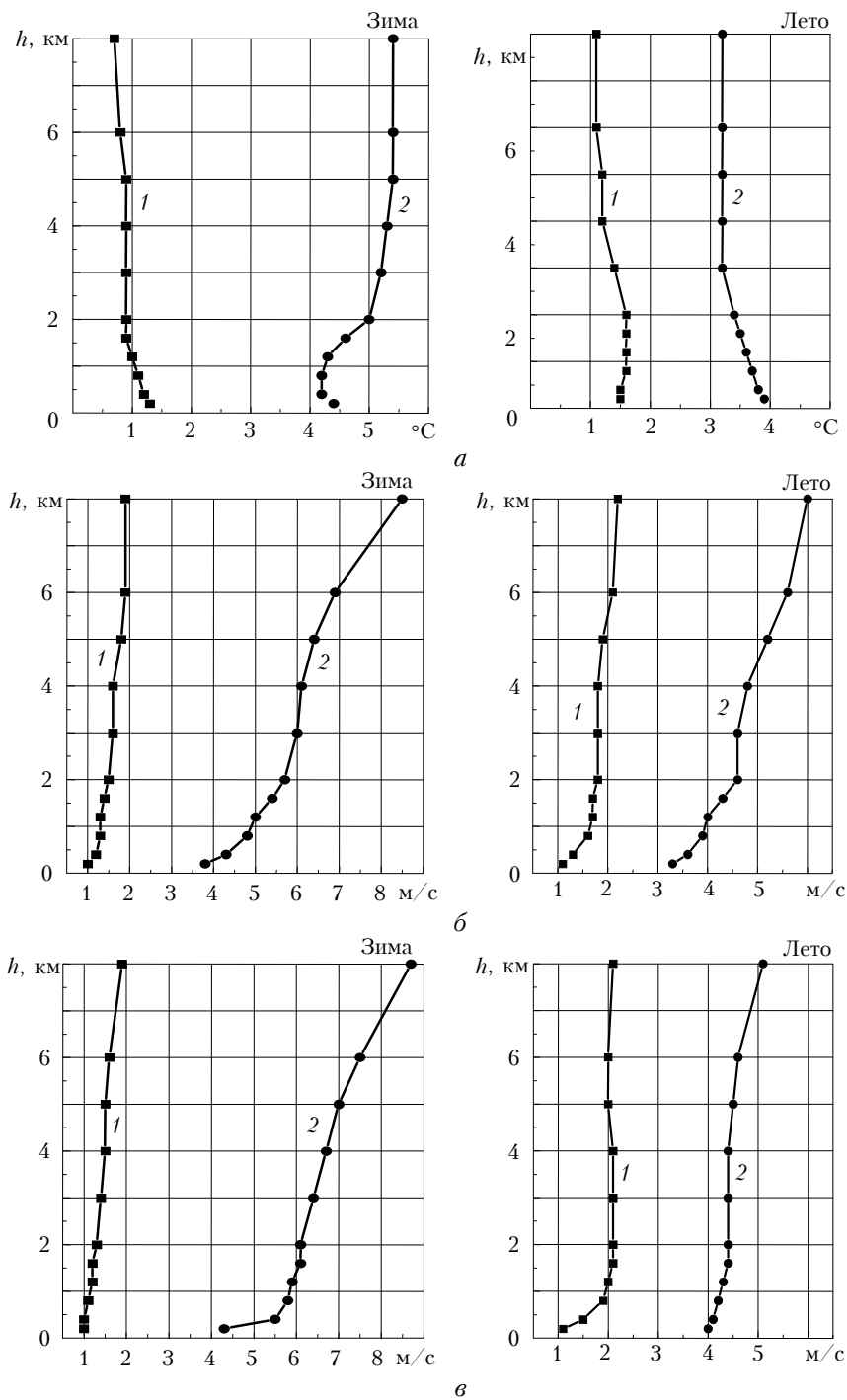


Рис. 2. Зависимости от высоты среднеквадратических погрешностей экстраполяции средних в слое значений температуры (а), зональной (б) и меридиональной (в) составляющих скорости ветра до расстояния 225 км, проведенной с помощью алгоритма фильтра Калмана и четырехмерной динамико-стохастической модели (1), а также соответствующие стандартные отклонения (2)

экстраполяции средних в слое значений температуры и ортогональных составляющих скорости ветра на расстояние до 225 км, дает достаточно хорошие результаты. Действительно, на этом расстоянии и независимо от сезона и взятого слоя атмосферы среднеквадратические погрешности такой экстраполяции варьируют в пределах 0,7–1,6 °C (для средней температуры) и 1,0–2,2 м/с (для зональной

и меридиональной составляющих скорости среднего ветра).

Кроме того, этот алгоритм позволяет экстраполировать поле ветра в пограничном слое атмосферы, т.е. в слое 0–1,6 км (здесь осуществляется согласно [10] основной перенос загрязняющих веществ промышленного происхождения) с достаточно высокой точностью (среднеквадратическая погрешность

порядка 1,0–2,0 м/с), которая близка к точности ветрового радиозондирования, когда его ошибка составляет 0,7–2,0 м/с [11].

В заключение следует подчеркнуть, что рассмотренный метод может быть улучшен. Для этого необходимо разработать адаптивный алгоритм, который позволит оценивать значение радиуса корреляции выбранной метеовеличины в каждой точке наблюдения и соответствующим образом подстраивать коэффициенты c_s в модели (1). Предметом отдельного исследования должно стать также изучение влияния параметров, определяющих временную и пространственную зависимость метеополя в точке прогноза от значений этого поля в точках наблюдения (в выражении (1) это первое и второе слагаемые).

1. Гандин Л.С., Коган Р.Л. Статистические методы интерпретации метеорологических данных. Л.: Гидрометеоздат, 1976. 359 с.
2. Гордин В.А. Математические задачи гидрометеорологического прогноза погоды. Л.: Гидрометеоздат, 1987. 264 с.
3. Комаров В.С., Попов Ю.Б. Оценка и прогнозирование параметров состояния атмосферы с помощью алгоритма фильтра Калмана. Часть 1. Методические

основы // Оптика атмосфер. и океана. 2001. Т. 14. № 4. С. 255–260.

4. Комаров В.С., Ильин С.Н., Кузнецов Б.П., Попов Ю.Б., Попова А.И. Динамико-стохастическое прогнозирование полей температуры и ветра применительно к оценке состояния загрязненности атмосферы на ограниченной территории // Оптика атмосфер. и океана. 2002. Т. 15. № 5–6. С. 479–482.
5. Dee D.P. Simplification of the Kalman filter for meteorological data assimilation // Quart. J. Roy. Meteorol. Soc. 1991. V. 117. P. 365–384.
6. Климова Е.Г. Методика усвоения данных метеонаблюдений на основе обобщенного субоптимального фильтра Калмана. // Метеорол. и гидрол. 1997. № 11. С. 55–65.
7. Климова Е.Г. Модель для расчета ковариаций однородных изотропных случайных полей ошибок прогноза // Метеорол. и гидрол. 2001. № 10. С. 24–33.
8. Панчев С. Случайные функции и турбулентность. Л.: Гидрометеоздат, 1967. 447 с.
9. Сейдж Э.П., Мэлса Дж.Л. Теория оценивания и ее применение в связи и управлении. М.: Связь, 1976. 496 с.
10. Владимиров А.М., Ляхин Ю.И., Матвеев Л.Т., Орлов В.Г. Охрана окружающей среды. Л.: Гидрометеоздат, 1991. 423 с.
11. Guide to meteorological instrument and observing practices. Paris: WMO, 1984. 130 p.

V.S. Komarov, A.V. Lavrinenko, N.Ya. Lomakina, Yu.B. Popov, A.I. Popova, S.N. Il'in. Spatial extrapolation of meteorological fields in the mesoscale region based on the four-dimensional mixed dynamic-stochastic model and the apparatus of Kalman filtering.

A technique and an algorithm of spatial extrapolation of mesometeorological fields based on the four-dimensional mixed dynamic-stochastic model and the apparatus of Kalman filtering are considered. The results of statistical estimation of the quality of the proposed algorithm are discussed as applied to spatial extrapolation of mesoscale wind and temperature fields on the territory not covered by observations.