

В.С. Комаров, А.И. Попова, Ю.Б. Попов

Адаптивный алгоритм пространственного прогнозирования метеорологических полей на основе фильтра Калмана и полиномиальной модели второго порядка

Институт оптики атмосферы СО РАН, г. Томск

Поступила в редакцию 14.04.2004 г.

Рассматривается адаптивный алгоритм пространственного прогнозирования метеорологических полей, основанный на использовании фильтра Калмана и полиномиальной модели второго порядка с изменяющимися во времени коэффициентами, учитывающей меняющуюся конфигурацию локальной сети аэрологических станций. Обсуждаются результаты оптимизации количества используемых станций (с точки зрения минимума ошибки прогнозирования), а также оценка качества разработанного адаптивного алгоритма при его применении в процедуре объективного анализа мезомасштабных полей температуры и ветра.

Введение

Среди многих проблем современной метеорологии важное место занимает проблема объективного анализа метеорологических полей (и в первую очередь, полей геопотенциала, температуры и ветра) в области мезомасштаба, проводимого по данным локальной сети аэрологических станций. При этом решение подобной проблемы, в отличие от объективного анализа макромасштабных метеорологических полей, в значительно большей степени зависит от регулярности поступления данных от всех имеющихся аэрологических станций, поскольку их число в пределах мезомасштабных полигонов ограничено.

Однако на практике по различным причинам (например, из-за выхода из строя технического оборудования или штормового характера погоды) радиозондовые наблюдения на той или иной станции заданного полигона могут не проводиться. Кроме того, при передаче по каналам связи аэрологических телеграмм могут возникать грубые ошибки, обусловленные неправильным набором оператором символов в телеграмме, искажениями при передаче данных и т.п., поэтому, как следствие, телеграммы могут быть декодированы частично или вовсе не подлежат декодированию.

В любом из перечисленных случаев, для корректной работы алгоритмов пространственной интерполяции (основы объективного анализа), возникает необходимость восстановления утраченных данных для той или иной станции. В ранее приведенных работах [1–3] алгоритмы пространственной интерполяции были разработаны в предположении, что данные наблюдений для заданной конфигурации локальной измерительной сети присутствуют всегда. Уточним, что под конфигурацией локальной измерительной сети мы понимаем количество

аэрологических станций и их пространственное расположение в рамках мезомасштабного полигона.

В отличие от [1–3], в настоящей статье для решения задачи объективного анализа метеорологического поля в области мезомасштаба предлагается несколько иная методика. Она также основана на применении фильтра Калмана и полиномиальной модели второго порядка с изменяющимися во времени коэффициентами, однако учитывает изменяющуюся (от срока к сроку наблюдений) конфигурацию локальной измерительной сети. Кроме того, в этой методике используется дополнительно процедура расщепления исходных измерений на регулярную и флуктуационную составляющие. Рассмотрим эту методику подробнее.

Пусть в i -х точках заданного мезомасштабного полигона ($i = 1, 2, \dots, S$) с координатами (x_i, y_i) проводятся измерения метеорологического поля ξ . Требуется по данным этих измерений определить значение поля ξ_0 в некоторой точке (или узле регулярной сетки) с координатами (x_0, y_0) , находящейся на не освещенной метеорологической информацией территории того же мезомасштабного полигона. С этой целью представим поле ξ в виде суммы его регулярной $\bar{\xi}$ и флуктуационной ξ' составляющих (т.е. $\xi = \bar{\xi} + \xi'$), и для каждой из них рассмотрим методику восстановления.

Так, для оценки регулярной составляющей поля $\bar{\xi}_0$ в точке или узле (x_0, y_0) по данным трех ближайших (к этой точке) станций используется следующее выражение:

$$\xi(x_i, y_i) = a_0 + a_1(x_i - x_0) + a_2(y_i - y_0), \quad (1)$$

или в матричном виде

$$\xi(x, y) = \mathbf{H}(x, y) \cdot \mathbf{A}, \quad (2)$$

где $\xi(\mathbf{x}, \mathbf{y}) = |\xi(x_1, y_1), \xi(x_2, y_2), \xi(x_3, y_3)|^T$ – вектор измерений, T – оператор транспонирования;

$$\mathbf{H}(\mathbf{x}, \mathbf{y}) = \begin{bmatrix} 1 & (x_1 - x_0) & (y_1 - y_0) \\ 1 & (x_2 - x_0) & (y_2 - y_0) \\ 1 & (x_3 - x_0) & (y_3 - y_0) \end{bmatrix} - \text{переходная матрица};$$

$\mathbf{A} = |a_0, a_1, a_2|^T$ – вектор неизвестных коэффициентов, для оценки которого использовано выражение вида

$$\hat{\mathbf{A}} = \mathbf{H}^{-1}(\mathbf{x}, \mathbf{y}) \cdot \bar{\xi}(\mathbf{x}, \mathbf{y}). \quad (3)$$

Решая (3), можно легко найти значение \hat{a}_0 , которое соответствует значению регулярной составляющей $\bar{\xi}_0$ в точке (x_0, y_0) .

В то же время оценка флуктуационной составляющей поля ξ' базируется на применении алгоритма фильтрации Калмана и малопараметрической модели, описывающей пространственную изменчивость этого поля. В качестве такой модели используется полином второй степени с изменяющимися во времени коэффициентами, а именно:

$$\xi'_i(k) = a_0(k) + a_1(k)x_i + a_2(k)y_i + a_3(k)x_i y_i + a_4(k)x_i^2 + a_5(k)y_i^2, \quad (4)$$

где $\xi'_i(k)$ – значение флуктуационной составляющей метеорологического поля в точке i с координатами (x_i, y_i) в момент времени k ; $a_0(k), a_1(k), \dots, a_5(k)$ – неизвестные коэффициенты полинома.

Для оценивания коэффициентов полинома введем вектор состояний динамической системы следующего вида:

$$\mathbf{X}(k) = |a_0(k), a_1(k), \dots, a_5(k)|^T = |X_1(k), X_2(k), \dots, X_6(k)|^T. \quad (5)$$

В этом случае динамику изменения составляющих вектора состояний (5) можно записать в виде системы разностных уравнений [1]:

$$\begin{cases} X_1(k+1) = X_1(k) + \omega_1(k) \\ X_2(k+1) = X_2(k) + \omega_2(k) \\ \dots \\ X_6(k+1) = X_6(k) + \omega_6(k), \end{cases} \quad (6)$$

где $\boldsymbol{\omega}(k) = |\omega_1(k), \omega_2(k), \dots, \omega_6(k)|$ – вектор случайных возмущений системы (порождающие шумы, шумы состояния).

В векторной форме система уравнений (6) принимает вид

$$\mathbf{X}(k+1) = \mathbf{X}(k) + \boldsymbol{\omega}(k). \quad (7)$$

Будем считать, что флуктуационное поле ($\xi' = \xi - \bar{\xi}$) измеряется в i -х точках и в дискретные моменты времени k с некоторой погрешностью. Тогда модель его текущих измерений $Y_i(k)$ может

быть задана в виде суммы истинного значения $\xi'_i(k)$ и ошибки измерения $\varepsilon_i(k)$, т.е.

$$Y_i(k) = \xi'_i(k) + \varepsilon_i(k). \quad (8)$$

Вся совокупность измерений $Y_i(k)$ объединяется в вектор измерений $\mathbf{Y}(k) = |Y_1(k), Y_2(k), \dots, Y_S(k)|^T$, где S – число используемых станций, на которых проводятся измерения.

Представим модель (8) через переменные состояния (5), для чего введем переходную матрицу наблюдений $\mathbf{H}^*(\mathbf{x}, \mathbf{y})$ и запишем связь вектора измерений $\mathbf{Y}(k)$ с вектором состояний $\mathbf{X}(k)$ в матричной форме

$$\mathbf{Y}(k) = \mathbf{H}^*(\mathbf{x}, \mathbf{y}) \cdot \mathbf{X}(k) + \mathbf{E}(k). \quad (9)$$

Отметим, что размерность вектора $\mathbf{E}(k)$ также определяется числом станций S .

Переходная матрица измерений $\mathbf{H}^*(\mathbf{x}, \mathbf{y})$ имеет размерность $(S \times 6)$ и задается через координаты станций следующим образом:

$$\mathbf{H}^*(\mathbf{x}, \mathbf{y}) = \begin{bmatrix} 1 & x_1 & y_1 & x_1 y_1 & x_1^2 & y_1^2 \\ 1 & x_2 & y_2 & x_2 y_2 & x_2^2 & y_2^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_s & y_s & x_s y_s & x_s^2 & y_s^2 \end{bmatrix}. \quad (10)$$

Модель динамической системы (7) и модель измерений (9) являются линейными, поэтому задача оценивания решается с помощью линейного фильтра Калмана–Бьюси [4, 5], обеспечивающего оценку вектора состояния $\mathbf{X}(k)$ с минимальной дисперсией. Отметим, что при этом на элементы моделей (7) и (9) накладываются следующие условия:

$$M[\boldsymbol{\Omega}(k) \boldsymbol{\Omega}^T(j)] = \mathbf{Q}_k \delta_{kj}$$

– ковариационная матрица шума состояний (δ_{kj} – символ Кронекера);

$$M[\mathbf{E}(k) \mathbf{E}^T(j)] = \mathbf{R}_k \delta_{kj}$$

– ковариационная матрица ошибок измерений;

$$M[\boldsymbol{\Omega}(k) \mathbf{E}(k)^T] = 0$$

– случайные процессы $\boldsymbol{\Omega}(k), \mathbf{E}(k)$ взаимно некоррелированы;

$$M[\mathbf{X}_0 \boldsymbol{\Omega}^T(k)] = M[\mathbf{X}_0 \mathbf{E}(j)^T] = 0$$

– начальное состояние \mathbf{X}_0 некоррелировано с возмущениями $\boldsymbol{\Omega}(k)$ и $\mathbf{E}(k)$.

Алгоритм оценки коэффициентов полинома имеет следующий вид:

$$\hat{\mathbf{X}}(k+1) = \hat{\mathbf{X}}(k+1|k) + \mathbf{G}(k+1) \cdot \mathbf{J}_0(k+1) \cdot [\mathbf{Y}(k+1) - \mathbf{H}^*(\mathbf{x}, \mathbf{y}) \cdot \hat{\mathbf{X}}(k+1|k)], \quad (11)$$

где

$$\hat{\mathbf{X}}(k+1) = |\hat{X}_1(k+1), \hat{X}_2(k+1), \dots, \hat{X}_6(k+1)|^T$$

– оценка вектора состояния на момент времени $(k+1)$;

$$\hat{\mathbf{X}}(k+1|k) = \hat{\mathbf{X}}(k)$$

– вектор предсказанных оценок на момент времени $(k+1)$ по данным на шаге k ; $\mathbf{G}(k+1)$ – матрица весовых коэффициентов размерностью $(6 \times S)$; $\mathbf{J}_0(k+1)$ – диагональная матрица, учитывающая наличие или отсутствие данных от измерительной станции.

В отличие от классического фильтра Калмана [6] в (11) введен дополнительный оператор – матрица $\mathbf{J}_0(k+1)$. Ее формирование осуществляется на этапе предварительного анализа текущих измерений. При этом $\mathbf{J}_0(k+1)$ имеет размерность $(S \times S)$ и элементы ее главной диагонали принимают значения:

0 – при отсутствии или низком качестве данных от измерительной станции i ;

1 – когда данные от станции i не несут аномальных ошибок.

Таким образом, при выполнении математических операций (11) компоненты невязки измерений

$$[\mathbf{Y}(k+1) - \mathbf{H}^*(\mathbf{x}, \mathbf{y}) \cdot \hat{\mathbf{X}}(k+1|k)],$$

соответствующие некорректным или отсутствующим измерениям, исключаются из формирования оценки $\hat{\mathbf{X}}(k+1)$.

В остальном алгоритм соответствует классическому линейному фильтру Калмана–Бьюси. Расчет весовых коэффициентов $\mathbf{G}(k+1)$ представляет собой независимую от (11) рекуррентную процедуру, связанную с решением матричных уравнений для ковариаций ошибок оценивания [4]:

$$\mathbf{G}(k+1) = \mathbf{P}(k+1|k) \mathbf{H}^{*T}(\mathbf{x}, \mathbf{y}) \times$$

$$\times [\mathbf{H}^*(\mathbf{x}, \mathbf{y}) \mathbf{P}(k+1|k) \mathbf{H}^{*T}(\mathbf{x}, \mathbf{y}) + \mathbf{R}_E(k+1)]^{-1}; \quad (12)$$

$$\mathbf{P}(k+1|k) = \mathbf{P}(k|k) + \mathbf{R}_\Omega(k); \quad (13)$$

$$\mathbf{P}(k+1|k+1) =$$

$$= [\mathbf{I} - \mathbf{G}(k+1) \mathbf{H}^*(\mathbf{x}, \mathbf{y})] \cdot \mathbf{P}(k+1|k), \quad (14)$$

где $\mathbf{P}(k+1|k)$ – апостериорная матрица ковариаций ошибок предсказания размерностью (6×6) ; $\mathbf{P}(k+1|k+1)$ – априорная матрица ковариаций ошибок оценивания размерностью (6×6) ; $\mathbf{R}_E(k+1)$ – диагональная ковариационная матрица шумов наблюдения размерностью $(S \times S)$; $\mathbf{R}_\Omega(k)$ – диагональная ковариационная матрица шумов состояния размерностью (6×6) ; \mathbf{I} – единичная матрица размерностью (6×6) .

Окончательный расчет прогнозируемого значения метеорологической величины $\hat{\xi}_0(k+1)$ в точке (x_0, y_0) и момент времени $(k+1)$ выполняется по формуле

$$\hat{\xi}_0(k+1) = \bar{\xi}_0 + \hat{X}_1(k+1) + \hat{X}_2(k+1) x_0 + \hat{X}_3(k+1) y_0 +$$

$$+ \hat{X}_4(k+1) x_0 y_0 + \hat{X}_5(k+1) x_0^2 + \hat{X}_6(k+1) y_0^2. \quad (15)$$

Для начала работы алгоритма (11), (15) в момент $k=0$ (момент инициации) необходимо задать следующие начальные условия: начальный вектор оценивания $\hat{\mathbf{X}}(0) = M[\mathbf{X}(0)]$, начальную матрицу ковариаций ошибок оценивания

$$\mathbf{P}(0|0) = M\{[\mathbf{X}(0) - \hat{\mathbf{X}}(0)] \cdot [\mathbf{X}(0) - \hat{\mathbf{X}}(0)]^T\},$$

а также значения элементов ковариационных матриц шумов $\mathbf{R}_E(k)$ и $\mathbf{R}_\Omega(k)$.

На практике значения $\hat{\mathbf{X}}(0)$ и $\mathbf{P}(0|0)$ могут быть заданы, исходя из минимального объема сведений о реальных свойствах системы, а в случае полного отсутствия полезной информации задаются $\hat{\mathbf{X}}(0) = 0$, а $\mathbf{P}(0|0) = \mathbf{I}$.

Рассмотренный выше адаптивный алгоритм калмановской фильтрации, основанный на использовании полиномиальной модели второго порядка, был исследован на предмет его эффективности при применении в процедуре объективного анализа мезомасштабных полей температуры и ветра. С этой целью были использованы данные двухлетних (2002–2003 гг.) наблюдений тринадцати аэрологических станций с географическими координатами: Шлезвиг (54°32' с.ш., 09°33' в.д.), Эмден (53°23' с.ш., 07°14' в.д.), Грайфсвальд (54°06' с.ш., 13°24' в.д.), Берген (52°49' с.ш., 09°56' в.д.), Линденберг (52°13' с.ш., 14°07' в.д.), Эссен (51°24' с.ш., 06°58' в.д.), Майнинген (50°34' с.ш., 10°23' в.д.), Идар-Оберштайн (49°42' с.ш., 07°20' в.д.), Штуттгарт (48°50' с.ш., 09°12' в.д.), Кюммесбрукк (49°26' с.ш., 11°54' в.д.), Прага (50°00' с.ш., 14°27' в.д.), Мюнхен (48°15' с.ш., 11°33' в.д.), Бруно-Сокольнице (49°07' с.ш., 16°45' в.д.), представляющими типичный мезомасштабный полигон (рис. 1). Этот полигон выбран по той причине, что он представлен наибольшим числом аэрологических станций, расположенных друг от друга на минимальных (по сравнению с другими полигонами) расстояниях.

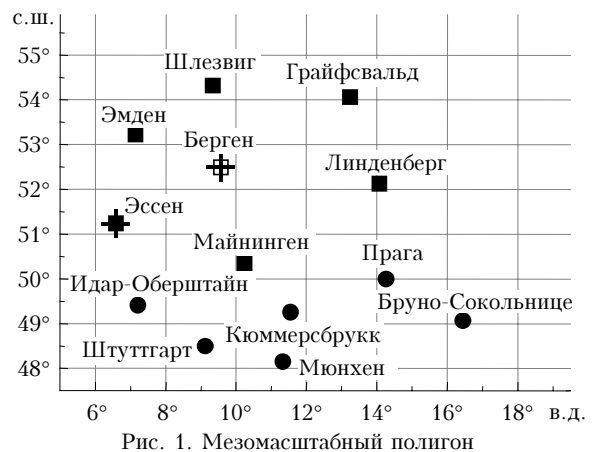


Рис. 1. Мезомасштабный полигон

Все отобранные данные были приведены (с помощью интерполяции и с учетом данных особых точек) к системе геометрических высот: 0 (уровень земной поверхности); 0,2; 0,4; 0,8; 1,2; 1,6; 2,0; 2,4; 3,0; 4,0; 5,0; 6,0 и 8,0 км.

Поскольку в пределах выбранного мезомасштабного полигона оказалось достаточно большое число аэрологических станций, то возник вопрос, какое количество станций обеспечит максимальную точность пространственного прогнозирования на основе предложенной методики. Поэтому предварительно, на первом этапе (т.е. до оценки качества адаптивного алгоритма) нами были проведены численные эксперименты по выбору оптимального количества аэрологических станций, обеспечивающего минимальную погрешность пространственной интерполяции полей температуры, зональной и меридиональной составляющих скорости ветра. С этой целью в качестве контрольной станции, куда осуществлялась интерполяция, была взята ст. Берген (см. рис. 1). Сама процедура оптимизации заключалась в последовательном сокращении числа используемых станций (в зависимости от их удаленности от ст. Берген) и в оценке на каждом шаге качества интерполяции метеорологического поля с помощью предложенного алгоритма. При этом число аэрологических станций последовательно уменьшалось от 12 до 3.

Результаты такой оценки приведены на рис. 2, где в качестве примера представлены значения среднеквадратической погрешности (δ_ξ) пространственной интерполяции температуры, зональной и меридиональной составляющих скорости ветра на уровнях 0,4; 3,0 и 8,0 км в зависимости от числа использованных станций N . Из графиков следует, что оптимальное число станций $N = 6$, причем независимо от высотного уровня и метеорологической величины. Именно при использовании данных измерений шести станций (они показаны на рис. 1 квадратами) отмечается минимальная среднеквадратическая погрешность интерполяции. Аналогичные результаты были получены и в случае про-

странственной экстраполяции поля ξ в контрольную точку ст. Эссен.

С учетом этих результатов на втором этапе исследований была реализована собственно процедура статистической оценки качества адаптивного алгоритма, базирующегося на использовании фильтра Калмана и полиномиальной модели второго порядка. При этом пространственная интерполяция и экстраполяция полей температуры и ортогональных составляющих скорости ветра проводилась на те же контрольные станции (Берген и Эссен) по данным шести станций, выявленных в процессе оптимизации.

Здесь следует подчеркнуть, что оценка качества пространственной интерполяции (экстраполяции) указанных метеополей проводилась применительно к прогнозу распространения облака загрязнения веществ промышленного происхождения. В этом случае согласно [7] используются усредненные в слое значения температуры и скорости ветра, рассчитанные для интервала высот $h - h_0$ (здесь h_0 соответствует уровню земной поверхности, а h — высота верхней границы выбранного слоя) с помощью выражения

$$\langle \xi \rangle_{h_0, h} = \frac{1}{h - h_0} \int_{h_0}^h \xi(h) dh, \quad (16)$$

где $\langle \cdot \rangle$ обозначает процедуру осреднения данных в заданном слое атмосферы, т.е. 0–0,2; 0–0,4; 0–0,8; 0–1,2; 0–1,6; 0–2,0; 0–2,4; 0–3,0; 0–4,0; 0–5,0; 0–6,0 и 0–8,0 км.

На рис. 3 в качестве примера представлены для летнего сезона графики зависимости среднеквадратических погрешностей пространственной интерполяции δ_1 (или экстраполяции δ_2) средних значений температуры $\langle T \rangle$, зональной $\langle U \rangle$ и меридиональной $\langle V \rangle$ составляющих скорости ветра от высоты и соответствующие стандартные отклонения σ , полученные по данным шести станций мезомасштабного полигона.

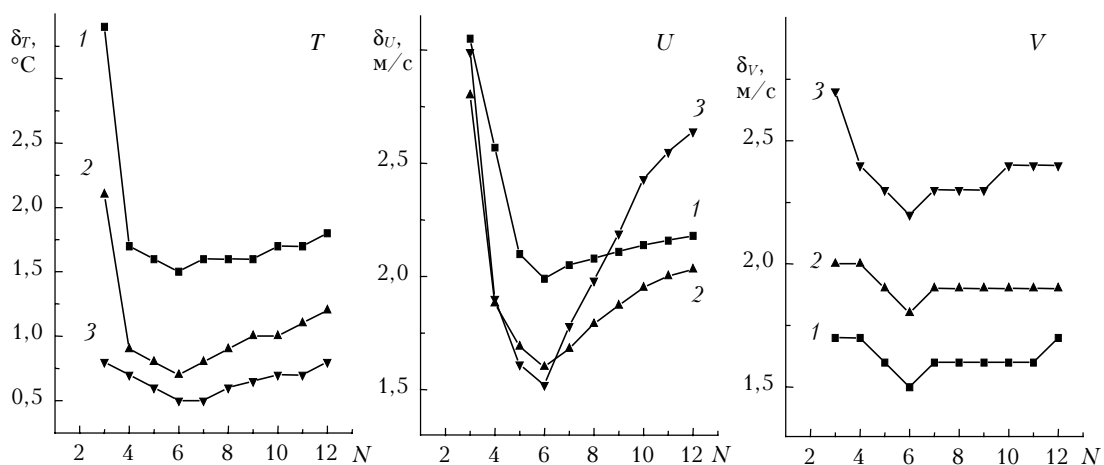


Рис. 2. Зависимость среднеквадратических погрешностей оценки температуры (T), зональной (U) и меридиональной (V) составляющих скорости ветра, от числа используемых станций N . Кривая 1 — 0,4; 2 — 3,0 и 3 — 8,0 км

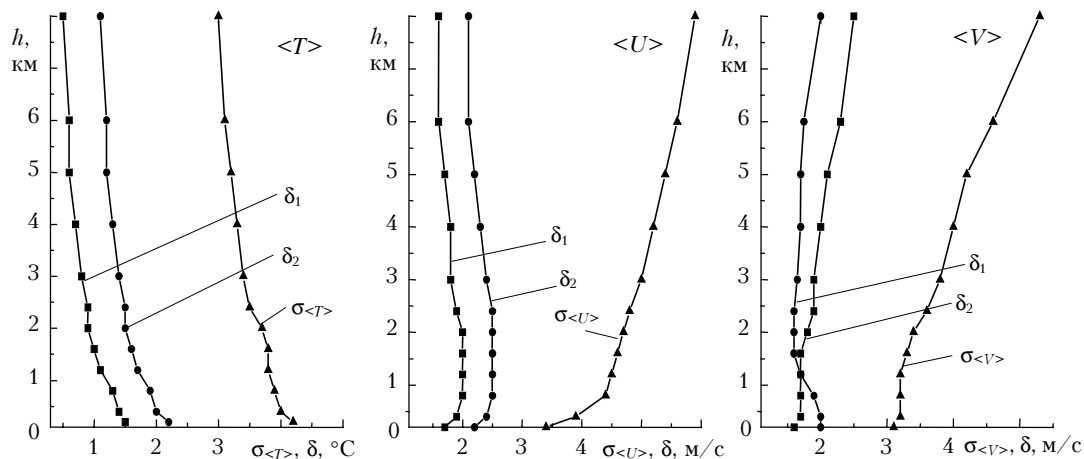


Рис. 3. Высотная зависимость среднеквадратических погрешностей пространственной интерполяции (δ_1) и экстраполяции (δ_2) средних в слое значений температуры $\langle T \rangle$, зональной $\langle U \rangle$ и меридиональной $\langle V \rangle$ составляющих скорости ветра, а также среднеполигонные стандартные отклонения σ

Анализ рис. 3 показывает, что, во-первых, предложенный адаптивный алгоритм обеспечивает лучшие результаты по точности в случае, когда он используется для пространственной интерполяции. При этом значения среднеквадратической погрешности независимо от слоя атмосферы варьируют в пределах 1,1–1,4 °C (для параметра $\langle T \rangle$) и 1,5–2,0 м/с (для параметров $\langle U \rangle$ и $\langle V \rangle$), во-вторых, тот же алгоритм дает вполне приемлемые для практики результаты в случае его применения для решения процедуры пространственной экстраполяции в области мезомасштаба, когда среднеквадратическая погрешность такой экстраполяции на расстояние 230 км составляет во всем пограничном слое атмосферы не более 1,5 °C (для $\langle T \rangle$) и 2,0 м/с (для $\langle U \rangle$ и $\langle V \rangle$).

Таким образом, рассмотренный адаптивный алгоритм, базирующийся на фильтре Калмана и полиномиальной модели с изменяющимися во времени коэффициентами, может быть с успехом применен для метеорологической поддержки (прогностическими данными) решения задач оценки и прогноза процессов загрязнения атмосферы на ограниченных территориях.

Komarov V.S., Popova A.I., Popov Yu.B. Adaptive algorithm for spatial prediction of meteorological fields based on the Kalman filter and the second-order polynomial model.

The adaptive algorithm for spatial prediction of meteorological fields based on the use of the Kalman filter and the second-order polynomial model with the coefficients varying in time is considered. This algorithm is capable of accounting for the variable configuration of the local network of aerological stations. The results of optimization (from the viewpoint of the minimal prediction error) of the number of used stations are discussed. The quality of the developed adaptive algorithm is evaluated for the case of its application to the procedure of objective analysis of mesoscale wind and temperature fields.

1. Комаров В.С., Кураков В.А., Ломакина Н.Я., Попова А.И. Применение метода калмановской фильтрации для прогнозирования параметров состояний атмосферы // Проблемы современной радиоэлектроники и систем управления: Матер. конф. Томск, 2002. Т. 1. С. 212.
2. Комаров В.С., Попов Ю.Б., Кураков В.А., Курушев В.В., Попова А.И. Алгоритмы пространственно-временного прогноза метеопараметров на основе фильтра Калмана с использованием полиномиальной модели второго порядка // Оптика атмосф. и океана. 2002. Т. 15. № 11. С. 1007–1011.
3. Kurakov V.A., Komarov V.C., Kreminsky A.V., Lomakina N.Ya., Popov Yu.B., Popova A.I., Suvorov S.S. Polynomial algorithm of the spatial forecast of atmospheric state parameters based on the Kalman filtering and its application // Proc. SPIE. 2002. V. 5027. P. 60–67.
4. Dee D.P. Simplification of the Kalman filter for meteorological data assimilation // Quart. J. Roy. Meteorol. Soc. 1991. V. 117. P. 365–384.
5. Фильтрация и стохастическое управление в динамических системах / Под ред. К.Т. Леондеса. М.: Мир, 1980. 407 с.
6. Сейдж Э.П., Мелса Дж.Л. Теория оценивания и ее применение в связи и управлении. М.: Связь, 1976. 496 с.
7. Брюхань Ф.Ф. Методы климатической обработки и анализа аэрологической информации. М.: Гидрометеоиздат, 1983. 112 с.