

В.С. Комаров, А.В. Креминский

**ОПТИМАЛЬНОЕ КОМПЛЕКСИРОВАНИЕ АЛЬТЕРНАТИВНЫХ
МЕТОДОВ ПРОСТРАНСТВЕННОГО ПРОГНОЗА В ЗАДАЧАХ
АТМОСФЕРНО-ЭКОЛОГИЧЕСКОГО МОНИТОРИНГА
Ч. I. МЕТОДОЛОГИЯ И АЛГОРИТМЫ**

Предлагается оригинальный подход к решению задачи восстановления мезометеорологических полей на неосвещенных данными наблюдений территориях, основанный на оптимальном комплексировании альтернативных методов пространственного прогноза. Рассматриваются теоретические основы используемых с этой целью методов (полиномиальной аппроксимации, оптимальной экстраполяции и модифицированной версии МГУА), а также комплексный алгоритм для пространственного прогноза вертикальных профилей температуры и ветра.

1. Введение

Среди актуальных проблем современной мезометеорологии важное место занимает проблема надежного и адекватного восстановления высотной структуры мезометеорологических полей в районах с редкой сетью аэрологических станций (или с их полным отсутствием) по данным атмосферного зондирования, проведенного на прилегающих территориях. Это связано с тем, что оптимальное решение подобной проблемы необходимо для реализации многих прикладных задач и, в частности, таких задач, как:

- оценивание пространственного распространения техногенных загрязнений в воздушном бассейне ограниченных территорий (например, промышленной зоны или региона в целом), определяемого температурной стратификацией (она влияет на турбулентную диффузию загрязняющих веществ) и режимом ветра во многих точках этой территории;

- численный прогноз мезомасштабных процессов (с горизонтальным масштабом в десятки и сотни километров) и соответствующей им погоды в отдельных регионах и пунктах;

- оптимальное размещение сети лидарных станций, данные измерений которых должны быть использованы для атмосферного мониторинга ограниченных территорий локального или регионального масштаба, т.е. территорий с горизонтальным размером от 50 до 300 км [1].

Однако на практике решение проблемы оптимального восстановления (пространственного прогноза) высотной структуры мезометеорологических полей сталкивается с рядом трудностей.

Во-первых, существующая мировая сеть аэрологических станций характеризуется крайней неоднородностью и малой плотностью (даже на наиболее освещенных данными наблюдений территориях Европы и Северной Америки наименьшее расстояние между ближайшими станциями составляет около 300–400 км), что явно не удовлетворяет современным требованиям объективного анализа мезометеорологических полей (т.е. их построения на основе данных наблюдений и некоторого численного алгоритма). В частности, для регионального и локального прогнозов погоды данные аэрологических наблюдений должны быть представлены в узлах пространственной сетки с горизонтальным шагом в 50–200 км (для региональной модели) и 5–50 км или даже 1–5 км (для локальной модели) [1].

Во-вторых, применяемые в настоящее время алгоритмы объективного анализа метеорологических полей базируются главным образом на использовании методов полиномиальной аппроксимации [2–8] или оптимальной интерполяции (она наиболее детально описана в [9, 10]), которые обладают рядом существенных недостатков. Так, например, метод полиномиальной аппроксимации характеризуется произвольным выбором алгебраических полиномов, осуществляемым без учета свойств анализируемого метеорологического поля, и не дает достаточно удовлетворительных результатов в районах с малой плотностью наблюдательных станций. Метод же оптимальной интерполяции (экстраполяции) требует для своей реализации предварительного обобщения большого объема первичной информации и расчета (по данным много-

летних наблюдений) необходимых статистических характеристик, в том числе коэффициентов пространственной корреляции.

В-третьих, до сих пор в задачах объективного анализа высотных метеорологических полей не получили широкого распространения нетрадиционные методы (например, метод группового учета аргументов (МГУА) [11]), которые дают достаточно хорошие результаты восстановления в условиях информационной неопределенности и для построения прогностической модели не требуют предварительного обобщения данных многолетних наблюдений.

Учитывая все перечисленные обстоятельства, а также отсутствие публикаций по проблеме пространственной экстраполяции мезометеорологических полей на неосвоенных территориях, представляется целесообразным использовать для решения поставленной задачи некоторый комплексный подход, основанный на процедуре оптимального комплексования альтернативных методов пространственного прогноза, например, методов полиномиальной и оптимальной экстраполяции и модифицированной версии метода группового учета аргументов.

В настоящей статье рассматривается подобный подход. При этом главное внимание уделено постановке задачи, теоретическим основам ее численного решения, принципам построения комплексного алгоритма и, наконец, самому алгоритму пространственного прогноза высотной структуры метеорологических полей.

2. Постановка задачи и некоторые теоретические основы ее численного решения по данным экспериментальных наблюдений

Пусть в процессе эксперимента, проводимого в точках $r_i \in W_x \subset R^m$ (здесь r – радиус-вектор точки, а $i = 1, 2, \dots, n$ – число точек в некотором замкнутом множестве W_x конечномерного евклидова пространства R^m) получены значения какого-либо параметра (в нашем случае это метеорологическая величина) $f(r_i)$. Тогда процедура экстраполяции (пространственного прогноза) параметра f в точку $r_0 \in W_x \subset R^m$, т.е. определение его значения $f(r_0)$ вне множества W_x по известным значениям параметра f в точках r_1, r_2, \dots, r_n , относящихся к множеству W_x , может быть осуществлена в классе линейных моделей по выражению вида

$$\hat{f}(r_0) = \sum_{i=1}^n a_i f(r_i), \quad (1)$$

где a_1, a_2, \dots, a_n – некоторые весовые множители, не зависящие от наблюдаемых значений параметра f в точках r_1, r_2, \dots, r_n и подлежащие определению, причем таким образом, чтобы качество соответствующих прогнозов $\hat{f}(r_0)$ можно было бы гарантировать неравенством

$$E[\hat{f}(r_0) - f(r_0)]^2 \leq \varepsilon_i^2, \quad (2)$$

здесь $E(\cdot)$ – оператор математического ожидания, производящий усреднение по всем возможным реализациям $\Delta = \hat{f}(r_0) - f(r_0)$ (где $f(r_0)$ – значение параметра f в точке r_0), а ε_i^2 – оценка допустимой погрешности, зависящая от дисперсии ошибок наблюдения и от максимально возможной погрешности пространственного прогноза.

Исходя из поставленной задачи, рассмотрим коротко теоретические основы ее решения, производимого с помощью одного из альтернативных методов (полиномиальной аппроксимации, оптимальной интерполяции и МГУА), которые, как уже было сказано выше, были использованы для нахождения оптимального комплексного алгоритма экстраполяции мезометеорологических полей. С этой целью воспользуемся работами [8–11].

а. Метод полиномиальной аппроксимации

Согласно этому методу, определение значений какой-либо метеорологической величины f в точке с координатами x_0 и y_0 , расположенной на некоторой плоскости (или изобарической поверхности), производится по данным измерений той же величины f в i -х точках ($i = 1, 2, \dots, N$), находящихся в окрестностях точки прогнозирования. Если предположить, что поле значений f

может быть описано многочленом (обычно в качестве него используется алгебраический полином) по некоторым заданным функциям координат $F_k(x, y)$, то оно может быть представлено в виде [7]:

$$f_i = f(x_i, y_i) = \sum_{k=1}^K a_k F_{ki}, \quad (i = 1, 2, \dots, N), \quad (3)$$

где $F_{ki} = F_k(x_i, y_i)$, причем для алгебраического полинома $F_k(x_i, y_i)$ имеет следующий вид:

$$F_1(x_i, y_i) = 1; \quad F_2(x_i, y_i) = x; \quad F_3(x_i, y_i) = y; \quad (4)$$

$$F_4(x_i, y_i) = x^2; \quad F_5(x_i, y_i) = xy; \quad F_6(x_i, y_i) = y^2 \dots,$$

a_k – некоторые весовые множители, подлежащие определению с помощью N уравнений (3), где $k = 1, 2, \dots, K$ – число коэффициентов a_k в используемом полиноме.

Сразу же следует подчеркнуть, что поскольку данные наблюдений содержат некоторые ошибки, то для их учета необходимо, чтобы количество наблюдений N всегда превышало число коэффициентов полинома K . Но тогда равенство (3) удовлетворяется лишь приближенно, и поэтому на практике для нахождения коэффициентов a_k с помощью уравнений (3) обычно используется метод наименьших квадратов (наиболее подробно этот метод описан в [12]), причем таким образом, чтобы сумма квадратов невязок уравнений (3), т.е. сумма квадратов разностей между левыми и правыми частями

$$E = \sum_{i=1}^N \left(f_i - \sum_{k=1}^K a_k F_{ki} \right)^2 \quad (5)$$

была минимальной. Это условие соблюдается лишь тогда, когда производные функции E по всем a_k будут равны нулю, т.е.

$$\frac{\partial E}{\partial a_1} = 0; \quad \frac{\partial E}{\partial a_2} = 0; \dots; \quad \frac{\partial E}{\partial a_k} = 0. \quad (6)$$

В настоящей статье экстраполяция величин f осуществляется в условиях неоднородного размещения i -х точек на плоскости, что требует дополнительного учета (путем придания соответствующих весов) различий в их расстоянии от точки прогнозирования. Исходя из этого условия, в данном случае вместо минимума величины E , определяемого формулой (5), ищется минимум другой величины, а именно:

$$E' = \sum_{i=1}^N S_i \left(f_i - \sum_{k=1}^K F_{ki} \right)^2, \quad (7)$$

где множители S_i принимаются заданным образом, зависящими от расстояния

$$r_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}. \quad (8)$$

Кроме того, поскольку в соответствии с решаемой задачей требуется определять величины f_0 (а не f_i), то для ее оценки обычно помещают начало координат в точку прогнозирования, т.е. $x = y = 0$. Тогда из выражения (3) следует, что

$$f_i(0, 0) = a_1, \quad (9)$$

и поэтому из всех коэффициентов алгебраического полинома a_k нужно искать лишь первый из них.

б. Метод оптимальной экстраполяции

В отличие от метода полиномиальной аппроксимации данный метод основывается на анализе статистической структуры метеорологических полей. Поэтому прежде чем излагать существо метода, приведем некоторые характеристики статистической структуры метеорологических полей.

Пусть f – значение метеорологической величины в точках i или j ; \bar{f} – среднее значение этой величины (норма); $f' = f - \bar{f}$ – отклонение от среднего значения. В таком случае средняя величина квадрата функции $D = \overline{f'^2}$ называется дисперсией метеорологической величины f ; средняя величина произведения значений f' в точках i и j , т.е. $m_{ij} = \overline{f'_i f'_j}$, – ковариационной функцией, а ее отношение к корню квадратному из произведения дисперсий в этих точках, т.е.

$$\mu_{ij} = \frac{f'_i f'_j}{\sqrt{\overline{f'^2_i} \overline{f'^2_j}}} = \frac{m_{ij}}{\sqrt{D_i D_j}} \quad (10)$$

– нормированной автокорреляционной функцией (или автокорреляционным моментом).

Известно, что на небольших расстояниях (до нескольких сотен километров [10]) поле любой метеорологической величины f может быть принято в качестве изотропного и однородного. В этом случае дисперсии f во всех точках поля одинаковы, т.е. $D_i = D_j = d^2$, а величины автокорреляционной функции зависят только от расстояния между точками i и j , и поэтому она может быть представлена в виде

$$\mu_{ij} = \mu(r_{ij}) = \overline{f'_i f'_j} / d^2. \quad (11)$$

Поскольку в практических расчетах истинные значения автокорреляционной функции и дисперсии заменяются на их эмпирические значения, рассчитанные по данным измерений метеорологической величины (а эти измерения осуществляются с ошибкой δ), то эмпирические значения автокорреляционной функции μ_{ij} можно записать следующим образом:

$$\mu_{ij} = \frac{1}{N d^2} \sum_{k=1}^N \hat{f}'_{ik} \hat{f}'_{jk}, \quad (12)$$

где $d^2 = \frac{1}{N} \sum_{k=1}^N \hat{f}'_k{}^2 - \delta^2$ – дисперсия метеорологической величины f ; $\hat{f}'_{ik}, \hat{f}'_{jk}$ – отклонения этой величины от нормы в точках i и j ; N – число случаев, используемых при статистических расчетах.

После этих предварительных выкладок перейдем теперь к изложению существа метода оптимальной экстраполяции. В соответствии с указанным методом нахождение метеорологической величины f в точке с координатами x_0 и y_0 осуществляется по данным ее измерений в точках i ($i = 1, 2, \dots, n$), расположенных в окрестностях прогнозируемой точки, с помощью соотношения [10]

$$f_0 = \bar{f}_0 + f'_0 = \bar{f}_0 + \sum_{i=1}^n p_i \hat{f}'_i, \quad (13)$$

где \bar{f}_0 и f'_0 – среднее значение и отклонение от него метеорологической величины f ; $\hat{f}'_i = \hat{f}_i - \bar{f}_i$ – отклонение измеренных значений величины \hat{f}_i от нормы \bar{f}_i в i -х точках; p_i – некоторый интерполяционный вес, подлежащий определению, причем таким образом, чтобы средний квадрат такой экстраполяции

$$E = \overline{(\hat{f}_0 - f_0)^2} = \overline{f_0'^2} - \left[2 \overline{f_0' \sum_{i=1}^n p_i \hat{f}'_i} + \overline{\left(\sum_{i=1}^n p_i \hat{f}'_i \right)^2} \right] \quad (14)$$

был бы минимальным.

Используя, как в случае с методом полиномиальной аппроксимации, условие минимума функции E , т.е. условие

$$\frac{\partial E}{\partial p_i} = 0 \quad (i = 1, 2, \dots, n), \quad (15)$$

дифференцируя каждый член выражения (14) и введя некоторую относительную ошибку $\eta = \delta^2/d^2$, т.е. отношение среднего квадрата ошибки измерения метеорологической величины к ее дисперсии, характеризующей изменчивость величины f , приходим к системе уравнений

$$\sum_{j=1}^n p_j \mu_{ij} + p_i \eta = \mu_{0i} \quad (i = 1, 2, \dots, n), \quad (16)$$

используемых для определения весов p_i .

Для расчета коэффициентов p_i по формуле (16) необходимо прежде всего определить соответствующие значения автокорреляционных моментов μ_{ij} и μ_{0i} . В силу сделанного выше предположения об изотропности и однородности поля автокорреляционный момент μ_{ij} зависит только от расстояния между рассматриваемыми i -й и j -й точками, а также расстояния каждой из влияющих точек до прогнозируемой точки с координатами x_0 и y_0 .

Для определения указанных расстояний в декартовой системе координат, связанных с плоскостью географической карты, можно использовать выражение

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}. \quad (17)$$

На практике для оценки нормированных автокорреляционных функций в зависимости от расстояния r между точками i и j используются различные аналитические выражения. Так, например, аналитическое выражение для нормированной автокорреляционной функции температуры на поверхности земли, полученное в [13], имеет вид

$$\mu(r) = \exp(-0,825 r^{0,92}), \quad (18)$$

а для этой же функции, полученной в [14] для составляющей скорости ветра, аналитическое выражение записывается в следующем виде:

$$\mu_v(r) = \mu_s(r) = (1 - 0,98 r) \exp(-0,98 r). \quad (19)$$

После расчета значений μ_{ij} составляется система уравнений (16) и определяется вес p_i .

На заключительном этапе (после нахождения веса p_i) на основе выражения (13) производится определение значения метеорологической величины f_0 в точке с координатами x_0 и y_0 , для чего предварительно находится норма $\overline{f_0}$ и значения \hat{f}_i , которые оцениваются путем вычитания норм $\overline{f_i}$ во всех i -х точках из соответствующих измеренных значений \hat{f}_i .

в. Метод группового учета аргументов

В отличие от методов полиномиальной и оптимальной экстраполяции данный метод, взятый в модифицированном виде, позволяет экстраполировать в пространстве не только величины f_p , но и k -мерные векторы (вертикальные профили) этой метеорологической величины $f_i(\kappa_k)$, заданной на конечном числе уровней, в качестве которых обычно берутся геометрические высоты h , где $h = 0, 1, \dots, h_k$, или давление p ($p = p, p, \dots, p_k$). Будем считать (в соответствии с [15]), что на территориях локального и регионального масштабов высотная структура метеорологического поля $f_i(h_k)$ является однородной и может быть описана одной ковариационной матрицей. Тогда при соблюдении этого условия процедура пространственного прогноза (методом группового учета аргументов) величины $f_0(h_k)$ в точке с координатами x_0 и y_0 по данным

известных (измеренных) значений $f_i(h_k)$ в ближайшей i -й точке осуществляется с помощью системы линейных регрессионных моделей вида [11]

$$Y_0(h, N+1) = \sum_{\tau=1}^{N^*} A(h, \tau) Y_i(h, N+1-\tau) + \sum_{j=1}^{h-1} B(h, j) Y_0(j, N+1) + \varepsilon(h, N+1)$$

$$(h = \bar{h} + 1; \bar{h} + 2; \dots, h_k), \quad (20)$$

генерируемых на основе исходных экспериментальных данных, представляющих собой пространственно-временные наблюдения вида

$$\{Y_i(h, t), h = 0, 1, \dots, h_k; t = 1, \dots, N\}; \quad (21)$$

$$\{Y_0(h, t), h = 0, 1, \dots, \bar{h} \leq h_k; t = N+1\}, \quad (22)$$

где h – высота; t – время наблюдений; N^* – порядок запаздывания по времени ($N^* < [N - h - 1]/2$); $A(h, 1), \dots, A(h, N)$ и $B(h, 0), \dots, B(h, h-1)$ – неизвестные параметры модели; $\varepsilon(h, N+1)$ – невязка модели.

Алгоритм выбора наилучшей прогностической модели рассматривается в разделе 4 данной статьи.

3. Основные принципы построения комплексного алгоритма пространственного прогноза

Рассмотрим общие принципы построения комплексного алгоритма, т.е. алгоритма, получаемого в результате комплексирования альтернативных прогностических методик, который использован в настоящей работе для восстановления (пространственного прогноза) вертикальных профилей метеорологических величин (в нашем случае это температура воздуха (T_v), зональная (U) и меридиональная (V) составляющие скорости ветра) в точке с координатами x_0 и y_0 , по данным наблюдений на близрасположенных аэрологических станциях N (или пунктах дистанционного зондирования). Мы руководствовались следующими основными принципами:

1. Оперативный пространственный прогноз метеорологических величин должен осуществляться после предварительного оценивания соответствующих средних значений как в исходных i -х точках, так и в прогнозируемой точке с координатами x_0 и y_0 , где могут быть использованы модельные (фоновые) значения.

2. Прогностическая модель должна синтезироваться на основе ограниченного объема оперативной информации, т.е. в условиях полного отсутствия данных многолетних наблюдений.

3. Используемый алгоритм должен основываться на процедуре оптимального (с точки зрения качества экстраполяции) комплексирования альтернативных методов пространственного прогноза и учитывать особенности высотной структуры мезометеорологических полей, определяющих погодные условия на ограниченных территориях.

4. Процедура построения комплексного алгоритма пространственного прогноза должна состоять из двух этапов. Вначале осуществляется выбор из всех используемых уровней K такого уровня (или уровней) k_i , где отмечается минимум ошибки восстановления ε . С этой целью используется либо метод полиномиальной аппроксимации, либо метод оптимальной экстраполяции. А уже на втором этапе осуществляется сама процедура пространственного прогноза, когда на основе восстановленных (с помощью методов полиномиальной аппроксимации или оптимальной экстраполяции) значений метеорологической величины f_0 , взятых на уровне (уровнях) минимума погрешности прогноза ε и в срок наблюдений $t = N+1$, и на основе статистической выборки пространственно-временных наблюдений $f_i(h_k, t)$, где $t = 1, 2, \dots, N$ (согласно [16], $N = 7 \div 15$ наблюдений), производится оценивание, но уже с помощью ММГУА, всех недостающих компонент профиля $f_0(h_k, N+1)$.

Все эти постулаты и были использованы авторами при построении общей блок-схемы и комплексного алгоритма пространственного прогноза высотной структуры мезометеорологических полей, реализованных применительно к полям температуры, зональной и меридиональной составляющих скорости ветра.

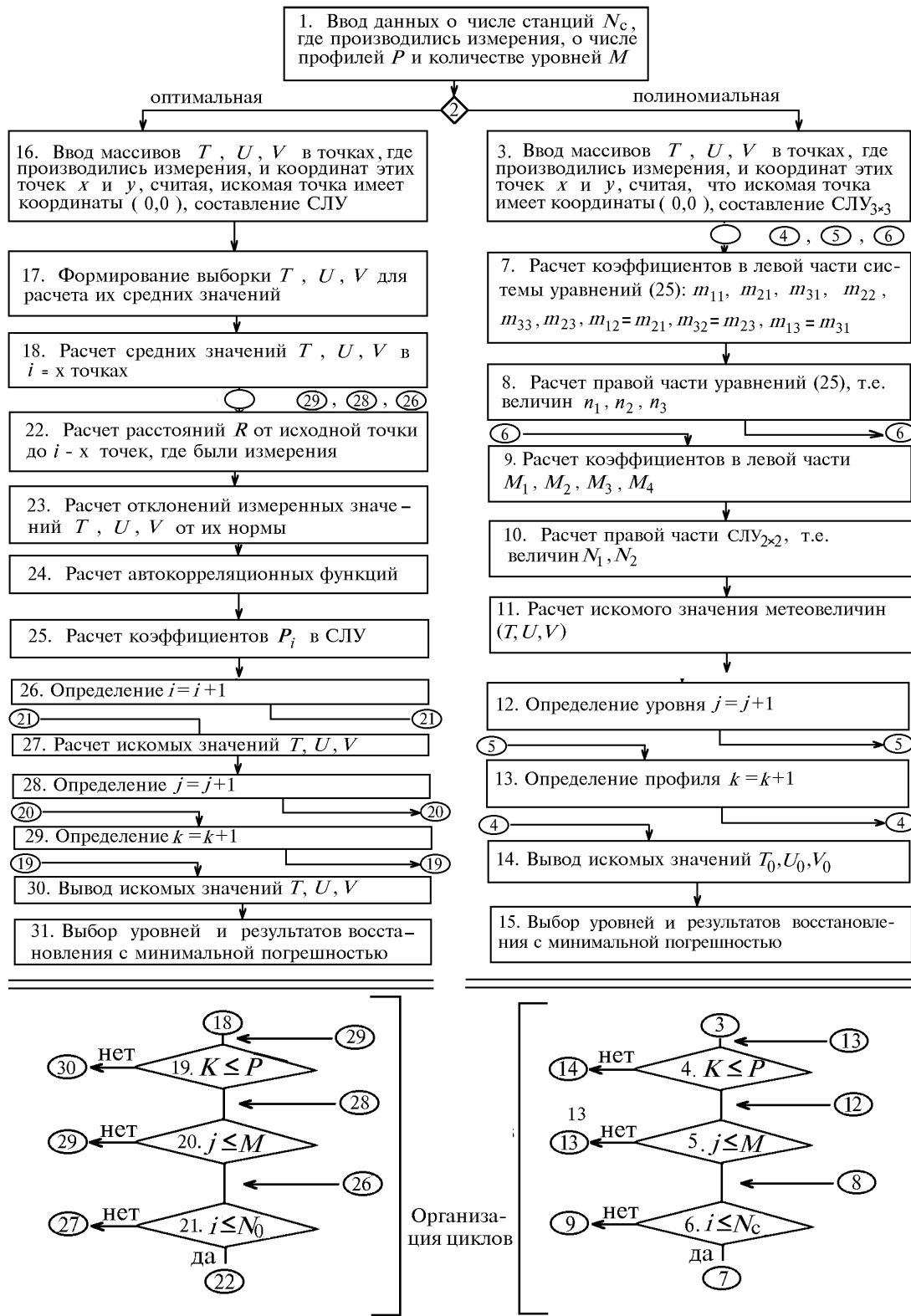


Рис. 1. Блок-схема пространственного прогноза методами полиномиальной и оптимальной экстраполяции

4. Блок-схема и комплексный алгоритм пространственного прогноза вертикальной структуры мезометеорологических полей

Для лучшего понимания процедуры комплексного алгоритма, используемого для пространственного прогноза вертикальной структуры мезометеорологических полей, рассмотрим более подробно ее отдельные этапы. Остановимся вначале на блок-схеме выбора уровня (или уровней) с минимумом ошибки прогноза (рис. 1), осуществляемого с помощью метода полиномиальной аппроксимации или метода оптимальной экстраполяции. Этот алгоритм может быть представлен в следующей последовательности.

1. Ввод данных о числе станций N_c , в которых проводились измерения, и количестве высотных уровней M .

2. Выбор метода пространственного прогноза из двух возможных методов (полиномиальной аппроксимации или оптимальной экстраполяции).

При выборе метода полиномиальной аппроксимации расчет температуры (T), зональной (U) и меридиональной (V) составляющих скорости ветра базируется на процедуре представления участка поля этих метеорологических величин в окрестностях точки алгебраическим полиномом

$$P_1(x, y) = a_0 + a_1 x + a_2 y, \quad (23)$$

где x, y – координаты; a_0, \dots, a_2 – коэффициенты. Поскольку начало координат помещено в рассматриваемую точку, то при $x = y = 0$ имеем, что

$$f_i(0, 0) = a_0, \quad (24)$$

где f_i – искомая величина, т.е. значения T, U, V .

3. Ввод массивов T, U, V , полученных для точек, где производились измерения, а также координат этих точек x и y , считая, что искомая точка имеет координаты $(0, 0)$, и составление (с учетом условий минимума суммы квадратов ошибок по всем коэффициентам a) системы линейных уравнений третьего порядка (СЛУ_{3*3})

$$\left. \begin{aligned} a_0 m_{11} + a_1 m_{21} + a_2 m_{31} &= n_1 \\ a_0 m_{12} + a_1 m_{22} + a_2 m_{32} &= n_2 \\ a_0 m_{13} + a_1 m_{23} + a_2 m_{33} &= n_3 \end{aligned} \right\}, \quad (25)$$

используемых для определения коэффициентов a в искомой точке.

4 – 6. Организация циклов для перебора уровней $j \leq M$, профилей $k \leq P$ и станций $i \leq N_c$.

7. Расчет коэффициентов в левой части СЛУ_{3*3}, с использованием следующих соотношений:

$$\begin{aligned} m_{11} &= 1; \quad m_{21} = \frac{1}{N_c} \sum_{i=1}^{N_c} x_i; \quad m_{31} = \frac{1}{N_c} \sum_{i=1}^{N_c} y_i; \quad m_{22} = \frac{1}{N_c} \sum_{i=1}^{N_c} x_i x_i; \quad m_{32} = \frac{1}{N_c} \sum_{i=1}^{N_c} x_i y_i; \quad m_{33} = \frac{1}{N_c} \sum_{i=1}^{N_c} y_i y_i; \\ m_{12} &= m_{21}; \quad m_{13} = m_{31}; \quad m_{23} = m_{32}. \end{aligned}$$

8. Расчет правой части СЛУ_{3*3} с использованием формул

$$n_1 = \frac{1}{N_c} \sum_{i=1}^{N_c} f_i; \quad n_2 = \frac{1}{N_c} \sum_{i=1}^{N_c} x_i f_i; \quad n_3 = \frac{1}{N_c} \sum_{i=1}^{N_c} y_i f_i.$$

9. Приведение СЛУ_{3*3} к СЛУ_{2*2}, т.е. получение системы линейных уравнений вида

$$\begin{aligned} a_0 M_1 + a_1 M_2 + N_1 &; \\ a_0 M_3 + a_1 M_4 + N_2 & \end{aligned} \quad (26)$$

и расчет коэффициентов в левой части СЛУ_{2*2} с использованием выражений

$$\begin{aligned} M_1 &= m_{12} - (m_{13} m_{32} / m_{33}); \quad M_2 = m_{22} - (m_{23} m_{32} / m_{33}); \\ M_3 &= m_{11} - (m_{13} m_{31} / m_{33}); \quad M_4 = m_{21} - (m_{23} m_{31} / m_{33}). \end{aligned}$$

10. Расчет правой части СЛУ_{2*2} с использованием соотношений

$$N_1 = n_2 - (n_3 m_{32} / m_{33}); \quad N_2 = n_1 - (m_{31} n_3 / m_{33}).$$

11. Расчет значений a_0 на основе выражения вида

$$a_0 = [N_1 - (M_2 N_2 / M_4)] / [M_1 - (M_2 M_3 / M_4)], \quad (27)$$

получаемого путем решения СЛУ_{2*2}, и нахождение на его основе значений T, U, V в искомой точке путем подставления коэффициента a_0 в выражение (24).

12. Повторение циклов расчета для всех используемых уровней $j = j + 1$.

13. Повторение циклов расчета для всех профилей $k = k + 1$.

14. Вывод искомых значений в виде таблицы.

15. Выбор уровней и результатов восстановления с минимальной погрешностью.

При выборе метода оптимальной экстраполяции определение тех же значений T, U, V в искомой точке осуществляется по иной схеме (см. тот же рис. 1), а именно:

16. Ввод массива T, U, V в точках, где производились измерения, а также координат этих точек x и y , считая, что искомая точка имеет координаты $(0, 0)$.

17. Преобразование данных и формирование исходной выборки T, U, V для расчета их средних значений.

18. Расчет средних значений $\bar{T}, \bar{U}, \bar{V}$ во всех i -х точках.

19–21. Организация циклов для перебора профилей $k \leq P$, уровней измерения T, U, V $j \leq M$, и перебора станций $i \leq N_c$.

22. Расчет расстояний R от исходной точки до точки, где производились измерения метеорологических величин T, U, V с помощью выражения

$$R = \sqrt{X^2 + Y^2}. \quad (28)$$

23. Расчет отклонения измеренного значения T, U, V от их среднего значения, т.е. определение

$$f = f'_{\text{изм}} - \bar{f}. \quad (29)$$

24. Расчет автокорреляционных функций μ для температуры с помощью аналитического выражения

$$\mu(R) = \exp(-0,825 r^{0,92}), \quad (30)$$

а для ветра

$$\mu_U(R) = \mu_V(R) = (1 - 0,98 R) \exp(-0,98 R). \quad (31)$$

25. Расчет коэффициентов p_i в системе уравнений (16) методом Гаусса, алгоритм которого состоит из двух этапов.

Первый этап заключается в последовательном исключении неизвестных из уравнений, с помощью выражения:

$$p_i = \left(\mu_{0i} - \sum_{j=1}^n p_j \mu_{ij} - p_i \eta \right) / \mu_{ij}, \quad (32)$$

где μ_{0i}, μ_{ij} – автокорреляционные функции; p_i – вес, а η – так называемая мера ошибки измерений.

Согласно (32), необходимо каждый элемент первой строки расширенной матрицы СЛУ поделить на диагональный элемент:

$$\mu_{ij} = \mu_{ij} / \mu_{ii} \cdot \quad (33)$$

Затем подставляем выражение (32) во все остальные уравнения системы, исключая тем самым соответствующие p_i в каждом уравнении. А элементы расширенной матрицы преобразуются по формуле:

$$\mu_{0i} = \mu_{ij} - \mu_{ii} \mu_{jj} \cdot \quad (34)$$

В результате проделанных вычислений получим СЛУ с верхней треугольной матрицей, у которой все элементы ниже главной диагонали равны нулю.

Второй этап решения СЛУ состоит в последовательном определении неизвестных коэффициентов p_i по формуле (16).

26. Повторение циклов расчета для используемых станций $i = i + 1$.

27. Расчет искомого значения T, U, V в точке с координатами x_0 и y_0 с использованием климатической нормы \bar{a}_0 (например, взятой из таблиц региональных моделей высотного распределения температуры и составляющих скорости ветра или полученной иным способом) с помощью формулы вида

$$a_0 = \bar{a}_0 + \sum_{i=1}^{N_c} p_i a_i \cdot \quad (35)$$

28. Повторение циклов расчета для всех используемых уровней $j = j + 1$.

29. Повторение циклов расчета для всех профилей $k = k + 1$.

30. Вывод результатов (искомых значений T, U, V) в виде таблицы.

31. Выбор уровней и результатов восстановления с минимальной погрешностью.

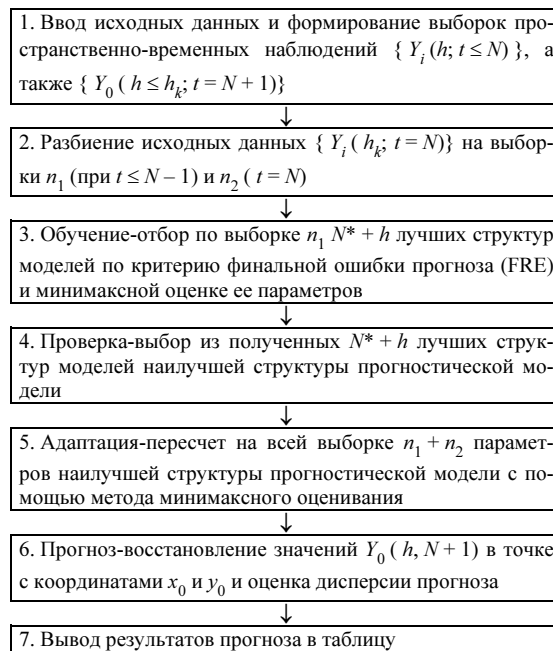


Рис. 2. Блок-схема восстановления высотных профилей метеорологических величин с помощью модифицированной версии МГУА

После нахождения уровня (уровней) с минимумом ошибки восстановления T, U, V осуществляется процедура определения (с помощью ММГУА [11] значений этих метеорологических величин для всех недостающих уровней (блок-схема на рис. 2). Данный алгоритм реализуется в следующей последовательности.

1. Ввод массивов T, U, V , взятых для ближайшей (к искомой точке) i -й точки (станции), а также данных восстановления этих величин в искомой точке на уровнях минимума величины

ε, проведенного с помощью полиномиальной или оптимальной экстраполяции, и формирование выборок пространственно-временных наблюдений вида (21) и (22).

2. Разбиение исходной выборки (21) на подвыборку n_1 , содержащую все наблюдения до момента времени $t = N - 1$ включительно, и подвыборку n_2 , включающую в себя лишь наблюдения в момент времени $t = N$ (контрольная выборка).

3. Блок <обучение>. По подвыборке n_1 отбирается $N^* + h$ лучших структур моделей вида (20) по критерию Х. Акаике (FPE):

$$FPE = \frac{(N - N^* - 1) + s}{(N - N^* - 1) + s} RSS(s), \quad (36)$$

где

$$RSS(s) = \sum_{j=1}^{N-N^*-1} [Y_{h,N-j}^{(i)} - \hat{Y}_{h,N-j}^{(i)}(s)]^2$$

– остаточная сумма квадратов для текущей модели $\hat{Y}_{h,N-j}^{(i)}(s)$, содержащей s -ненулевых оценок параметров,

$$\hat{Y}_{h,N-j}^{(i)} = X\hat{\theta}, \quad X \in M_{(N-N^*-1)(N^*+h)}, \quad \hat{\theta} \in R^{N^*+h},$$

где

$$\hat{\theta} = [\hat{A}_{h,1} \dots \hat{A}_{h,N^*} \hat{B}_{h,0} \dots \hat{B}_{h,h-1}]^T \quad (37)$$

– минимаксная оценка параметров по подвыборке n_1 , вычисляемая по формулам, приведенным ниже; R^k – евклидово пространство k -мерных векторов; $M_{m \times p}$ – пространство матриц размера $m \times p$.

4. Блок <Проверка>. Из полученных $N^* + h$ лучших структур моделей отбирается единственная структура (в дальнейшем, называемая <наилучшая> структура) с минимальной ошибкой прогноза на контрольной подвыборке n_2 :

$$|Y_{h,N}^{(i)} - \hat{Y}_{h,N}^{(i)}(s)| \rightarrow \min,$$

где минимум берется по всем $N^* + h$ структурам, каждой из которых соответствует своя модель $\hat{Y}_{h,N}^{(i)}(s)$.

Матрица входных переменных (регрессоров) $X^*(s) \in M_{(N-N^*-1) \times s}$ наилучшей структуры имеет вид

$$X^*(s) = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ x_{p_1} & : & x_{p_2} & : & \dots & : & x_{p_s} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ Y_{h,N^*}^{(i)} & \dots & Y_{h,1}^{(i)} & \cdot & Y_{0,N^*+1}^{(i)} & \dots & Y_{h-1,N^*+1}^{(i)} \end{bmatrix}, \quad (38)$$

где x_{p_1}, \dots, x_{p_s} – столбцы матрицы X ,

$$X = \begin{bmatrix} Y_{h,N-2}^{(i)} & \dots & Y_{h,N-N^*-1}^{(i)} & \cdot & Y_{0,N-1}^{(i)} & \dots & Y_{h-1,N-1}^{(i)} \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ Y_{h,N^*}^{(i)} & \dots & Y_{h,1}^{(i)} & \cdot & Y_{0,N^*+1}^{(i)} & \dots & Y_{h-1,N^*+1}^{(i)} \end{bmatrix}, \quad (39)$$

а p_1, \dots, p_s – номера ненулевых компонент вектора оценок.

5. Блок <Адаптация>. Вектор параметров $\hat{q} \in R^s$ наилучшей структуры (37) пересчитывается (адаптируется) на всей выборке $n_1 + n_2$ исходных данных (21)–(22) по формулам, реализующим алгоритм мини-максного оценивания:

$$\theta^* = \gamma \tilde{\theta} \quad \tilde{\theta} = [\tilde{X}^T \tilde{X}]^{-1} \tilde{X}^T \tilde{Z} \quad (40)$$

вектор оценок

$$\gamma = \Delta_{h, N+1}^2 / (\Delta_{h, N+1}^2 + u_h), \quad u_h = \tilde{x}_h^T [\tilde{X}^T \tilde{X}]^{-1} \tilde{x}_h \sigma^2, \quad (41)$$

где $\Delta_{h, N+1}$ – максимально допустимые значения прогнозируемой величины $Y_{h, N+1}^0$, которые либо вычисляются по формуле

$$\Delta_{h, N+1} = \max_{t=1, \dots, N} |Y_{h, t}^{(i)}| \quad (h = 0, 1, \dots, h^*), \quad (42)$$

либо задаются пользователем, а матрица входных переменных \tilde{X} и вектор выходных переменных \tilde{Z} определяются следующим образом:

$$\tilde{Z} = [\tilde{z}_i] = \begin{bmatrix} Y_{h, N-1}^{(i)} \\ \dots \\ Y_{h, N^*+1}^{(i)} \\ \dots \\ Y_{h, N}^{(i)} \end{bmatrix} \in R^{N-N^*}, \quad \tilde{X} = [\tilde{x}_{ij}] = \begin{bmatrix} X^*(s) \\ \dots \\ \tilde{x}_h^T \end{bmatrix} \in M_{(N-N^*)s}, \quad (43)$$

$\tilde{x}_h = [\tilde{x}_{p_1} \tilde{x}_{p_2} \dots \tilde{x}_{p_s}]^T \in R^s$, $\tilde{x}_{p_1}, \dots, \tilde{x}_{p_s}$ – наблюдения, стоящие в массиве наблюдений

$$\begin{bmatrix} Y_{h, N-1}^{(i)} & Y_{h, N-2}^{(i)} & \dots & Y_{h, N-N^*}^{(i)} & \vdots & Y_{0, N}^{(i)} & \dots & Y_{h-1, N}^{(i)} \\ & & & & & & & \end{bmatrix} \quad (44)$$

под номерами p_1, p_2, \dots, p_s соответственно,

$$\sigma^2 = \frac{1}{N - N^* - s} \sum_{i=1}^{N-N^*} \left(\tilde{z}_i - \sum_{j=1}^s \tilde{x}_{ij} \tilde{\theta}_j \right)^2. \quad (45)$$

6. Блок <Прогноз>. В этом блоке вычисляется прогноз $Y_{h, N+1}^{*0}$ в точке $(h, N + 1)$:

$$Y_{h, N+1}^{*0} = \tilde{x}_h^T \theta^* \quad (46)$$

и верхняя статистическая оценка дисперсии прогноза

$$E[E(Y_{h, N+1}^0 - Y_{h, N+1}^{*0})^2] \leq \delta_{h, N+1}, \quad \delta_{h, N+1} = u_h D_{h, N+1}^2 / (u_h + D_{h, N+1}^2), \quad (47)$$

где u_h вычисляется по формуле (41).

7. Вывод результатов пространственного прогноза вертикального профиля $Y_0(h, N + 1)$, т.е. недостающих компонент вертикального профиля f .

Таким образом, полностью решена проблема пространственного прогноза вертикальных профилей температуры и составляющих скорости ветра в районах с полным отсутствием данных температурно-ветрового зондирования. В заключение следует сказать, что об эффективности предлагаемого комплексного подхода можно судить лишь по результатам численных экспериментов, проведенных по данным высотного зондирования, осуществляемого в различных точках какого-либо мезометеорологического полигона. Но эта задача является предметом второй части данной статьи, помещенной в настоящем тематическом номере журнала.

1. Белов П.Н., Борисенков Е.П., Панин Б.Д. Численные методы прогноза. Л.: Гидрометеиздат, 1989. 376 с.
2. Gilchrist B., Cressman G.P. // *Tellus*. 1954. V. 6. N 4. P. 309–318.
3. Bushby F.H., Huckle V.M. // *Quart. Journ. Roy. Meteorol. Soc.* 1957. V. 83. N 356. P. 232–247.
4. Corby G.A. // *Quart. Journ. Roy. Meteorol. Soc.* 1961. V. 87. N 371. P. 34 – 42.
5. Corby G.A. // *Tellus*. 1963. V. 15. N 4. P. 432 – 438.
6. Быков В.В., Курбаткин Г.П. // *Изв. АН СССР. Сер. Геофиз.* 1961. N 2. С. 307–318.
7. Быков В.В., Курбаткин Г.П., Горелышева Н.В. // *Труды ММЦ*. 1964. Вып. 4. С. 56–72.
8. Белов П.Н. Практические методы численного прогноза погоды. Л.: Гидрометеиздат, 1967. 335 с.
9. Белоусов С.Л., Гандин Л.С., Машкович С.А. Обработка оперативной метеорологической информации с помощью электронных вычислительных машин. Л.: Гидрометеиздат, 1968. 282 с.
10. Гандин Л.С., Каган Р.Л. // *Статистические методы интерполяции метеорологических данных*. Л.: Гидрометеиздат, 1976. 359 с.
11. Кочерга Ю.Л. // *Автоматика*. 1991. N 5. С. 80–87.
12. Папазов М.Г., Могильный С.Г. Теория ошибок и способов наименьших квадратов. М.: Недра, 1968. 303 с.
13. Казакевич Д.И. Основы теории случайных функций и ее применение в гидрометеорологии. Л.: Гидрометеиздат, 1977. 319 с.
14. Либерман Ю.М. // *Применение статистических методов в метеорологии*. М.: Гидрометеиздат, 1978. С. 162–165.
15. Зуев В.Е., Комаров В.С. Статистические модели температуры и газовых компонент атмосферы. Л.: Гидрометеиздат, 1986. 264 с.
16. Комаров В.С., Акселевич В.И., Креминский А.В. // *Оптика атмосферы и океана*. 1994. Т. 7. N 2. С. 231–237.

Институт оптики атмосферы СО РАН,
Томск

Поступила в редакцию
29 декабря 1994 г.

V.S. Komarov, A.V. Kreminskii. Optimum Integration of Alternative Methods of Space Forecasting in Atmospheric and Ecological Monitoring. P. I. Methodology and Algorithms.

An original approach is proposed to solve the problem of reconstructing the mesometeorological fields on the territories devoid of observations data based on optimal integration of alternative methods of space forecasting. The basis of the methods used (polynomial approximation, optimum extrapolation, and MGAAM) as well as integrated algorithm for space forecasting the vertical profiles of temperature and wind are treated in the paper.