«Оптика атмосферы и океана», **20**, № 6 (2007)

УДК 519.6

А.В. Протасов

Динамико-вероятностный метод четырехмерного анализа полей метеоэлементов

Институт вычислительной математики и математической геофизики СО РАН, г. Новосибирск

Поступила в редакцию 28.03.2007 г.

Рассматривается метод восстановления полей метеорологических элементов в узлах пространственно-временной регулярной сетки по данным наблюдений на станциях. Метод основан на представлении искомых полей в виде разложений в конечные ряды по функциям естественного ортогонального базиса, который вычисляется по ансамблю пространственно-временных реализаций, рассчитанных с помощью динамико-вероятностного метода при заданной реальной статистической структуре полей метеоэлементов. По существу, предлагается один из методов быстрого усвоения данных наблюдений.

Проблема анализа и интерпретации фактической информации относится к числу важнейших вопросов, возникающих при построении математических моделей физических процессов, решении задач прогноза погоды, общей циркуляции атмосферы и океана, теории климата, а также при изучении и оценке влияния деятельности человека на окружающую среду. Один из аспектов этой проблемы состоит в разработке методов «сжатия» информации и выделении ее наиболее информативной части в виде суммы конечного ряда Фурье с небольшим числом слагаемых.

В данной статье предлагается один из методов четырехмерного анализа данных на основе климатического ансамбля возможных реализаций соответствующих многомерных гидрометеорологических полей для выбранного интервала времени и заданного региона в виде [1, 2]:

$$\{\xi^i_{(n)}, i = 1, 2, \ldots\}, \qquad (1)$$

где

$$\xi^i_{(n)} = [\mathbf{U}^i(\mathbf{X}_j, t_k), T^i(\mathbf{X}_j, t_k), H^i(\mathbf{X}_j, t_k), \ldots]^{\mathrm{T}}$$

— вектор реализаций полей скорости, температуры, геопотенциала и т.д. в пространственно-временных точках $(\mathbf{X}_j, t_k)$ рассматриваемой сеточной области размерности n; индекс Т определяет операцию транспонирования.

Для построения этого ансамбля используется метод динамико-вероятностного моделирования, достаточно подробно описанный в [1–3]. Существенной особенностью данного подхода является то, что в рамках единой модели используются данные реальных измерений, статистическое моделирование и численная модель гидротермодинамики атмосферы. При этом основным связующим элементом является вариационное усвоение информации гидродинамической моделью.

В качестве реальных данных были использованы данные реанализа поля температуры NCEP/NCAR за 1948–2005 гг. на 10 стандартных уровнях для зимнего сезона года с дискретностью по времени 6 ч и 2,5×2,5° по горизонтали. Выборка осуществлялась для заданной локальной области Ω Северного полушария размером 10×10° с центром в точке с кородинатами 60,56° с.ш. и 77,7° в.д. Задача расматривалась в системе координат x, y, p в области, нижним основанием которой является прямоугольник на плоскости, касательной в этой центральной точке. Для построения сеточной области было выбрано разрешение 24×20 по x и y с шагами Δx = 23,85 км и Δy = 58,74 км соответственно.

Начальным этапом построения климатического ансамбля (1) является построение соответствующего ансамбля реализаций

$$\{\xi^i_{(n)_c}, i = 1, 2, \ldots\} \qquad (2)$$

с использованием одного из методов статистического моделирования [1], который заключается в следующем.

Пусть R — многомерная корреляционная матрица. В нашем случае корреляционная матрица R рассчитана по данным реанализа для области, описанной выше. Представим спектральное разложение этой матрицы в виде

$$R = W\Lambda W^{\mathrm{T}}, \qquad (3)$$

где W — матрица собственных векторов корреляционной матрицы R; Λ — диагональная матрица соответствующих собственных значений. Заметим, что представление (3) является не чем иным, как разложением ее по так называемым главным факторам, а задача (3) является соответствующей задачей определения главных факторов. Дальнейший шаг заключается в определении квадратного корня матрицы R в виде $R^{1/2} = W\Lambda^{1/2}W^{\mathrm{T}}$, где $\Lambda^{1/2}$ – диагональная матрица, на диагонали которой стоят квадратные корни соответствующих собственных значений матрицы Λ.

Тогда мы можем определить случайный вектор

$$\xi^{(i)}_{(n)_c} = D_{\xi}R^{1/2}\psi^{(i)}(x_j, y_j, p_j, t_j)^{\mathrm{T}} + \bar{\xi}(x_j, y_j, p_j, t_j), \quad i = 1, 2, \ldots, \tag{4}$$

где $\psi^{(i)}(x_j, y_j, p_j, t_j)^{\mathrm{T}}$, $i = 1, 2, \ldots$ – гауссовский случайный вектор с единичной дисперсией и нулевым средним; D_{ξ} – диагональная матрица дисперсий; $\bar{\xi}(x_j, y_j, p_j, t_j)$ – соответствующий вектор средних. Нетрудно видеть, что корреляционная матрица случайного вектора $\xi^{(i)}_{(n)_c}$ в точности совпадает с матрицей R.

Заметим, что в общем случае реализации из ансамбля (2) определены на некотором множестве нерегулярных пространственно-временных точек, что полностью определяется данными измерений, используемыми для расчета корреляционной матрицы R. Кроме того, ансамбль (2) содержит, как правило, не все поля, необходимые для использования в численных моделях динамики атмосферы, а входящие в ансамбль (2) компоненты (температура, скорость, давление и т.д.) не являются согласованными относительно соответствующей численной модели динамики атмосферы.

Таким образом, возникает задача оптимального приближения ансамбля (2) соответствующим ансамблем реализаций, в котором каждая из реализаций удовлетворяла бы численной модели гидротермодинамики, а соответствующие статистические свойства были бы наиболее близкими.

В нашем случае для этой цели используется метод вариационного усвоения [3]. Следует заметить, что вариационное усвоение следует применять на всем рассматриваемом временном интервале и в пределах предсказуемости численной модели. В силу нелинейности исходной численной модели гидродинамики атмосферы и, как следствие этого, неединственности решения задачи минимума рассматриваемого функционала качества, а также в силу того, что в общем случае не существует доказательства сходимости и единственности получаемых решений, следует крайне аккуратно подходить к введению в модель дополнительных минимизируемых функционалов, таких как функционалы качества модели, функционала модели измерений и т.д. Поэтому в каждом конкретном случае необходимы дополнительные теоретические или численные исследования эффективности подобных введений функционалов, а также введение уравнений, определяющих соответствующие связи между рассматриваемыми компонентами.

Кроме того, для целей построения климатического ансамбля реализаций необходимо рассматривать задачу вариационного усвоения на всем временном интервале, поскольку применение так называемого последовательного пошагового усвоения не обеспечивает необходимую гладкость решений и соответствующий тренд для дальнейшего использования, например, полученного поля в режиме прогноза.

Таким образом, для построения конечного климатического ансамбля реализаций (1) мы применим вариационное усвоение, что, по-видимому, наиболее соответствует поставленной выше задаче. С этой целью для каждой реализации из ансамбля (2) решается задача вариационного усвоения с помощью математической модели гидротермодинамики атмосферы [1, 2], в результате чего получается ансамбль новых реализаций, отличающихся от исходных точностью решения задачи усвоения и удовлетворяющих свойствам, присущим математической модели.

Подробное описание совместной динамико-вероятностной модели и некоторые ее характеристики даны в работе [3]. Для решения этой задачи используется итерационный метод градиентного спуска, основанный на методе Лагранжа и решении соответствующих прямых и сопряженных задач. В численных расчетах ансамбль (2) был представлен только реализациями поля температуры. Однако ансамбль (1) содержит уже все поля метеоэлементов в соответствии с использованной моделью [1, 2]. В этом смысле модель является не только пространственно-временным интерполянтом, но и позволяет воспроизводить недостающие поля метеоэлементов. Анализ статистической структуры ансамбля (1) показывает, что этот ансамбль может быть использован в качестве климатического при решении прикладных задач, и в том числе при решении задач распространения примеси в атмосфере и при исследовании процессов выбросов в атмосфере.

В данной статье мы используем полученный климатический ансамбль (1) для решения задачи четырехмерного анализа гидрометеорологических данных в атмосфере. Один из алгоритмов такого использования предложен в работе [4] и основан на представлении искомого гидрометеорологического поля в виде соответствующего ряда по естественным ортогональным функциям, рассчитанным по реальным данным только для поля геопотенциала для зимнего периода и по достаточно ограниченной выборке.

Поскольку ансамбль (1) уже содержит статистически независимые пространственно-временные реализации, включающие в себя полный набор взаимно согласованных относительно численной модели динамики атмосферы гидрометеорологических компонентов (температура, геопотенциал, скорость ветра), то представляется вполне естественным использовать эту методику не только для анализа отдельных гидрометеорологических компонентов, но и для четырехмерного анализа и усвоения соответствующих реальных данных в целом. Естественно, что статистическая значимость полученных результатов полностью определяется ансамблем (1).

Этот подход имеет ряд преимуществ. Во-первых, базис естественных ортогональных функций, построенный по достаточно представительной выборке, обладает необходимыми свойствами статистической структуры полей метеоэлементов, что имеет особенно большое значение при редкой сети станций. Во-вторых, количество базисных функций предполагается сравнительно небольшим, что позволяет построить эффективный вычислительный алгоритм для его реализации. Кроме того, из методов построения естественного ортогонального базиса следует, что каждая из его функций имеет статистически согласованные компоненты, поэтому результат восстановления с помощью предлагаемого метода имеет ту же степень согласования, что и базисные функции.

Для расчета базисных функций естественного ортогонального базиса (главных факторов)

$$\{\varphi_i\}(i = \overline{1,m}) \tag{5}$$

по ансамблю (1) была использована одна из модификаций алгоритма, описанного в работе [5] для обобщенной ковариационной матрицы R_a ансамбля (1). Логарифмы собственных значений, характеризующих информативность рассчитанных базисных функций, представлены на рис. 1, из которого видно, что при 1615 реализациях в ансамбле (1) 50 базисных функций вполне достаточно для описания с хорошей точностью рассматриваемых метеополей ($m = 50$).

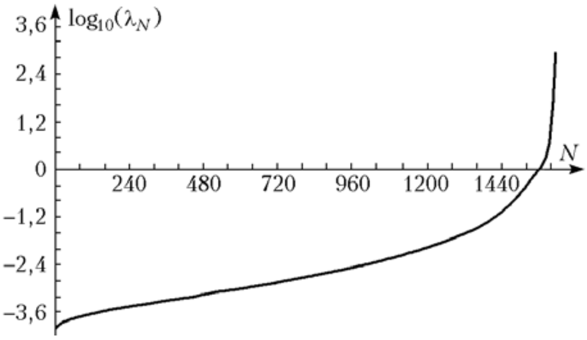

Рис. 1. Логарифмы собственных значений корреляционной матрицы, рассчитанной по ансамблю реализаций (1)

Таким образом, в соответствии с работой [4] рассмотрим подпространство векторов $\tilde{R}_m$ вещественного векторного пространства R_N, компонентами которых являются значения полей метеоэлементов в узлах регулярной пространственно-временной сеточной области $\Omega^{ht} \subset \Omega$. Пусть вектор-функции (5) являются базисом этого подпространства $\tilde{R}_m$. Тогда любой вектор $\varphi \in \tilde{R}_m$ мы можем представить в виде ряда Фурье

$$\varphi = \Phi\mathbf{a}, \tag{6}$$

где Φ – матрица размерности $N \times m$, составленная из базисных векторов $\{\varphi_i\}(i = \overline{1,m})$; $\mathbf{a} = (a_1,\ldots,a_m)^{\mathrm{T}}$ – вектор коэффициентов Фурье.

Пусть в рассматриваемой области Ω задана нерегулярная сеть Θ, в узлах которой известны данные измерений исследуемых полей метеоэлементов. Рассмотрим подпространство векторов G евклидова пространства, определенных на нерегулярной сеточной области Θ, причем в качестве компонентов векторов возьмем значения исследуемых полей одного или нескольких метеоэлементов (таких же, как и в $\tilde{R}_m$). В этом подпространстве введем скалярное произведение

$$(\varphi,\psi)_M = (M\varphi,\psi),$$

где $\varphi,\psi \in G$; символ $(,)$ обозначает скалярное произведение в евклидовом пространстве; M – положительно определенная симметричная матрица, выбор которой определяется целями исследования, физическими размерностями компонентов векторов и априорными сведениями о структуре рассматриваемых полей. В данном случае скалярное произведение является сеточным аналогом соответствующего скалярного произведения, определяющего интеграл полной энергии для решения используемой модели гидротермодинамики при решении задачи вариационного усвоения.

Задача восстановления полей метеоэлементов в узлах регулярной сетки Ω^{ht} по измеренным их значениям на нерегулярной сети станций при данном подходе сводится к нахождению вектора коэффициентов разложения в формуле (6) таких, чтобы проинтерполированные значения вектор-функции $\varphi \in \tilde{R}_m$ наименее отклонялись от соответствующих измеренных значений в узлах данной нерегулярной сети Θ.

Пусть $\psi_{\text{изм}}$ – вектор, составленный из измеренных значений в узлах нерегулярной сеточной области $\Theta \subset \Omega$, а φ – вектор из подпространства $\tilde{R}_m$, который требуется построить по заданному вектору $\psi_{\text{изм}}$. Обозначим через $\psi_{\text{изм}} = L\varphi$ образ вектора φ в подпространстве G, полученный с помощью линейного оператора интерполирования L с регулярной сетки на нерегулярную. Так как вектор $\varphi \in R_m$ представляется в виде (6), то $\psi = L\Phi\mathbf{a}$. Рассмотрим функционал, характеризующий меру отклонения вектора измеренных значений $\psi_{\text{изм}}$ в точках нерегулярной сети станций от значений вектор-функции $\boldsymbol{\varphi} \in R_m$, проинтерполированных на нерегулярную сетку с помощью линейного оператора L:

$$J = (\boldsymbol{\psi}_{\text{изм}} - L\Phi\mathbf{a}, \boldsymbol{\psi}_{\text{изм}} - L\Phi\mathbf{a})_M. \qquad (7)$$

Из условия экстремума функционала J получим для определения коэффициентов a_i, $i = \overline{1, m}$, линейную неоднородную алгебраическую систему уравнений

$$(L\Phi)^{\text{T}} M L\Phi\mathbf{a} = (L\Phi)^{\text{T}} M\boldsymbol{\psi}. \qquad (8)$$

Эту систему можно записать в виде

$$B\mathbf{a} = \mathbf{f}, \qquad (9)$$

где $B = (L\Phi)^{\text{T}} ML\Phi$ – симметричная, неотрицательно-определенная матрица; $\mathbf{f} = (L\Phi)^{\text{T}} M\boldsymbol{\psi}_{\text{изм}}$ – вектор правой части системы (8).

Заметим, что система (9) в некоторых случаях взаимного расположения узлов нерегулярной сети станций может оказаться плохо обусловленной. Поэтому для ее решения используем следующий алгоритм.

Матрицу B представим в виде

$$B = W_B \Lambda_B W_B{}^{\text{T}}, \qquad (10)$$

где Λ_B – диагональная матрица собственных значений; W_B – ортогональная матрица преобразования, столбцы которой составлены из собственных векторов матрицы B.

Тогда, учитывая (10), решение системы (9) получим по формуле

$$\mathbf{a} = W_B \Lambda_B^+ W_B{}^{\text{T}} \mathbf{f},$$

где

$$\Lambda_B^+ = \operatorname{diag}\{\lambda_i^+\}, \ (i = \overline{1, m})$$

– диагональная матрица, построенная по аналогии с псевдообратной матрицей, т.е.

$$\lambda_i^+ = \begin{cases} 1/\lambda_i, \ \text{при}\, \lambda_i > \varepsilon \\ 0, \ \text{при}\, \lambda_i \le \varepsilon, \end{cases}$$

ε – некоторое достаточно малое число.

Окончательно с помощью полученного вектора коэффициентов $\mathbf{a}$ можно восстановить поле $\boldsymbol{\varphi}$ на регулярной сеточной области Ω^{ht} по формуле (6).

Для иллюстрации эффективности описанной выше методики были смоделированы по формуле (4) данные поля температуры в области Θ для моментов времени $t = 0$ и 6 ч на 10 стандартных уровнях. Эти данные использованы в качестве входных для задачи вариационного усвоения с помощью численной модели и для четырехмерного анализа по формулам (6)–(9). На рис. 2 приводятся соответствующие сравнительные результаты расчетов для уровня 500 мбар и момента времени $t = 0$ ч, которые показывают достаточно хорошее качественное совпадение соответствующих полей изолиний.

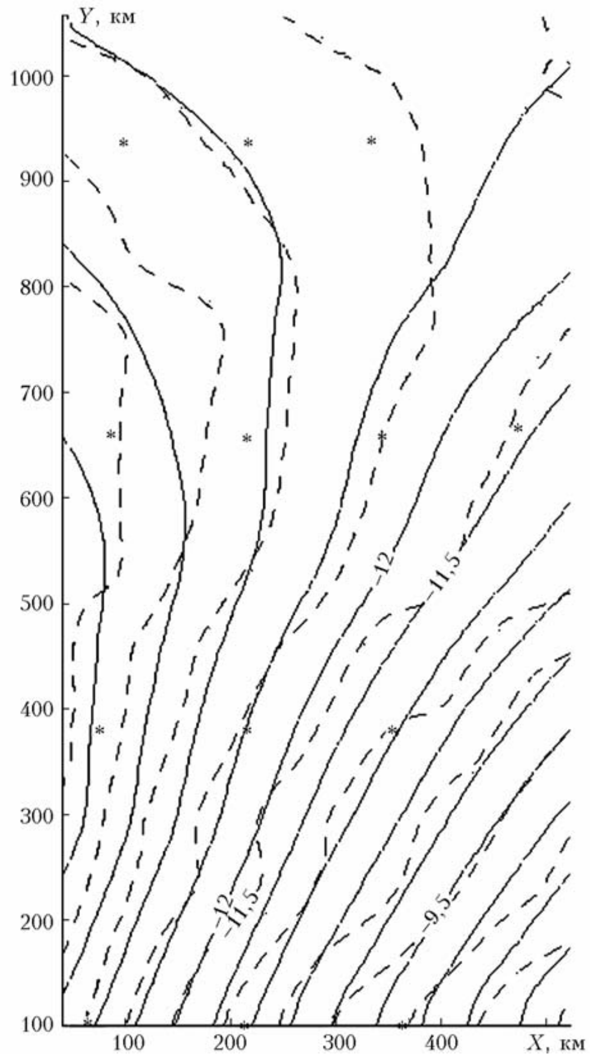

Рис. 2. Изолинии поля температуры на уровне 500 мбар в момент времени $t = 0$, полученного после вариационного усвоения (сплошные линии) данных, заданных в точках, обозначенных символом *, и изолинии поля температуры, полученного в результате четырехмерного анализа (штриховые линии) по главным факторам

Максимальная разность между значениями этих полей, а также данными составляет 0,93°.

Работа выполнена при поддержке РФФИ, грант № 05-05-98000-р_обь_а.

1. *Протасов А.В.* Динамико-вероятностное моделирование климатических полей метеоэлементов в локальной области на основе данных реанализа // Международ. научн. конгресс «ГЕО-Сибирь-2006». Новосибирск: Ред.-изд. отдел СГГА, 2006. Т. 3. Ч. 2. С. 112–118.

2. *Protasov A.V.* Dynamic-stochastic modeling of climatic fields of meteorological elements in local area on basis of the reanalysis data // 5 Int. Symp. «Mathematical Modeling of Dynamic Processes in Atmosphere, Ocean, and Solid Earth». Novosibirsk, 2006. P. 50–55.
3. *Ogorodnikov V.A., Protasov A.V.* Dynamic probabilistic model of atmospheric processes and the variational methods of data assimilation //Russ. J. Numer. Anal. Math. Modelling. 1997. VSP. V. 12. N 5. P. 461–479.
4. *Протасов А.В., Чекурова В.В.* Использование естественных ортогональных базисов для восстановления полей метеорологических элементов по данным измерений на редкой сети станций // Метеорол. и гидрол. 1983. № 1. С. 105–109.
5. *Пененко В.В., Протасов А. В.* Построение естественных ортогональных базисов для представления полей метеоэлементов // Изв. АН СССР. Физ. атмосф. и океана. 1978. Т. 14. № 12. С. 1361–1365.

A.V. Protasov. **Dynamic-stochastic method for four-dimensional analysis of fields of meteorological elements.**

The reconstruction method for fields of meteorological elements in the spatial-temporary regular grid using observation data at stations is considered. This method is based on expansion in series of required fields in finite number of natural orthogonal basis. The basis is determined on ensemble of spatial-temporary realizations received with the help of the dynamic-stochastic method at given real statistical structure of fields of meteorological elements. One of the methods of fast assimilation of the given observational data is in essence suggested.