

УДК 004.82

НЕЧЁТКИЙ АППРОКСИМАТОР АТМОСФЕРНЫХ ТЕМПЕРАТУРНЫХ ПОЛЕЙ

М. Ю. Катаев, А. В. Лавыгина, И. А. Ходашинский, Д. А. Эпштейн

*Государственное образовательное учреждение высшего профессионального образования
«Томский государственный университет систем управления и радиоэлектроники»,
634050, г. Томск, просп. Ленина, 40
E-mail: hodashn@rambler.ru*

Рассмотрены алгоритмы построения нечёткого аппроксиматора атмосферных температурных полей. Для определения параметров консеквентов нечётких правил аппроксиматора используется метод наименьших квадратов, а для определения параметров antecedentes — генетический алгоритм. Представлены результаты численного эксперимента. Приведены оптимальные параметры работы аппроксиматора.

Ключевые слова: нечёткие системы, аппроксимация, генетический алгоритм, метод наименьших квадратов.

Введение. Изучение атмосферы было и остаётся актуальной проблемой, решение которой тесно связано с задачами экологии, сельского хозяйства, геологии, метеорологии и др. Одной из особенностей её изучения является пространственная дискретность данных (как правило, неравномерная), получаемых от различного рода приборов и анализаторов. Так, при исследовании атмосферы очень часто возникают задачи, для решения которых необходимы данные в пространственной точке, где отсутствуют приборы измерений. Для этого разрабатываются специализированные наборы данных атмосферных параметров (температура, ветер, влажность и др.), например ECMWF или NCEP [1, 2], которые структурированы на равномерной пространственной сетке.

Обычно пространственная сетка таких наборов данных имеет достаточно низкое пространственное разрешение (например, для NCEP — это $2,5 \times 2,5^\circ$). Для получения набора параметров конкретной географической точки и заданного времени исследователи вынуждены проводить так называемую пространственную интерполяцию с одной сетки на другую. Одним из подходов является применение оптимальной интерполяции Гандина [3] либо билинейного подхода объективного анализа [4]. Данные на основе пространственной сетки значений формируются теми или иными аппроксиматорами, однако использование стандартных подходов к аппроксимации либо не даёт нужной точности, либо требует больших вычислительных мощностей. В связи с этим интересна и важна задача изучения новых подходов к решению проблемы интерполяции с пространственной сетки низкого разрешения на сетку с более высоким разрешением.

Указанная задача относится к компетенции моделирования, управляемого данными [5]. Это направление — в некотором смысле антипод глобальному моделированию, где на основе таблицы наблюдений строится единственная модель и в дальнейшем используется только построенное модельное описание. Типичными представителями такого традиционного подхода являются нейронные сети, различные формы нелинейной статистической регрессии, методы машинного обучения.

Цель моделирования, управляемого данными, состоит не в том, чтобы построить модель, которая объясняет весь процесс или явление, но определить лучший вывод для заданного входа, т. е. найти приближение функции только в окрестностях точки предсказания. Задача моделирования, управляемого данными, — улучшить точность предсказания

за счёт ухудшения интерпретируемости построенной модели. Типичными представителями такого подхода являются метод ближайших соседей, радиальные базисные функции, нечёткие системы типа Такаги — Сугено. Применяя указанную технологию, необходимо решить следующие проблемы: какие данные использовать? как много данных требуется? как эти данные взять?

Цель предлагаемой работы — решение проблемы аппроксимации температурных полей с помощью нечётких систем.

1. Постановка задачи. Доказано, что нечёткая система является универсальным аппроксиматором [6, 7], однако остаётся открытым вопрос определения параметров нечётких систем по конкретной таблице наблюдений.

Рассматриваемый в работе нечёткий аппроксиматор (НА) строится на основе модели типа Такаги — Сугено, правила которой имеют следующий вид:

$$\text{правило } i: \text{ ЕСЛИ } x_1 = A_{1i} \text{ И } x_2 = A_{2i} \text{ И } \dots \text{ И } x_m = A_{mi} \text{ ТО } z = f_i(x_1, \dots, x_m), \quad (1)$$

где f_i — линейная функция, определённая на переменных x_1, \dots, x_m .

Модель осуществляет отображение $F: \mathfrak{R}^m \rightarrow \mathfrak{R}$, заменяя оператор нечёткой конъюнкции произведением, а оператор агрегации нечётких правил сложением. Отображение F для модели типа Такаги — Сугено определяется формулой

$$F(\mathbf{x}) = \frac{\sum_{i=1}^R \mu_{A_{1i}}(x_1) \mu_{A_{2i}}(x_2) \dots \mu_{A_{mi}}(x_m) (a_{0i} + a_{1i}x_1 + \dots + a_{mi}x_m)}{\sum_{i=1}^R \mu_{A_{1i}}(x_1) \mu_{A_{2i}}(x_2) \dots \mu_{A_{mi}}(x_m)},$$

где $\mathbf{x} = [x_1, \dots, x_m]^T \in \mathfrak{R}^m$; $a_{ki} \in \mathfrak{R}$; R — число правил нечёткой модели; m — число входных переменных модели; $\mu_{A_{ij}}$ — функция принадлежности нечёткой области A_{ij} .

Нечёткая система может быть представлена как

$$z = f(\mathbf{x}, \boldsymbol{\theta}),$$

где z — скалярный выход системы; $\boldsymbol{\theta} = \|\theta_1, \dots, \theta_N\|$ — вектор параметров, N — число параметров нечёткой системы.

Пусть тестовая выборка, сформированная на основе таблицы наблюдений, представлена множеством данных $\{(\mathbf{x}_p; t_p), p = 1, \dots, K\}$, тогда средняя абсолютная функция ошибки, которую необходимо минимизировать, будет иметь вид

$$E(\boldsymbol{\theta}) = \sum_{p=1}^K |t_p - f(\mathbf{x}_p, \boldsymbol{\theta})|. \quad (2)$$

Проблема построения нечёткого аппроксиматора сводится к проблеме поиска минимума заданной функции (2) в многомерном пространстве, координаты которого соответствуют параметрам нечёткой системы.

Оптимально построенный НА должен удовлетворять двум критериям: точности и компактности. Количественная мера первого критерия задана формулой (2), мера второго критерия определяется количеством правил вида (1).

Выделим две основные фазы построения нечётких аппроксиматоров: обучение и проверка правильности построенного аппроксиматора. Обучение подразумевает определение структуры и идентификацию параметров НА. Структурные характеристики НА — это

число нечётких правил в базе и число лингвистических термов, на которое разбиты входные переменные. Этап определения структуры НА может быть выполнен с использованием следующих средств: субъективного разделения данных, алгоритма последовательного разделения или нечёткого кластерного анализа [8].

Идентификация параметров antecedentov и consequentov НА выполняется путём оптимизации работы нечёткой системы по заданному критерию (2). Для решения этой задачи используются две группы методов. Первая группа — классические методы оптимизации, основанные на производных: метод наименьших квадратов, градиентный метод, фильтр Калмана. Эти методы дают точные результаты, но они имеют тенденцию сходиться к локальным оптимумам. Вторая группа методов — метаэвристические, такие как алгоритмы муравьиной колонии, роящихся частиц, имитации отжига, генетические. Достоинство метаэвристических методов заключается в большей устойчивости. Но это методы грубого определения параметров, требующие больших временных ресурсов. Кроме того, применение метаэвристик не гарантирует нахождения оптимального решения [9, 10].

Таблица наблюдений строится на основе информации из набора данных об атмосферных температурных полях NCEP (пространственная сетка $2,5 \times 2,5^\circ$, высотная сетка 17 уровней 1000–30 мбар, шаг по времени 6 ч в течение года). Такая информация формируется на основе стандарта хранения и работы с научными данными — netCDF (network Common Data Form) [11].

Для определения параметров consequentov аппроксиматора в данной работе используется метод наименьших квадратов, а для параметров antecedentov — генетический алгоритм.

2. Нечёткий аппроксиматор. На вход аппроксимирующей системы поступают следующие данные: координаты точки предсказания (b^*, l^*) , в которой необходимо определить значение температуры; координаты узлов пространственной сетки 4×4 вместе с известной температурой в этих узлах $(b^i, l^j, t^{i,j})$ (рис. 1). Выбор такой сетки обусловлен критерием компактности базы правил. По данным пространственной сетки формируется обучающая и тестовая выборки. Тестовая выборка состоит из двух строк: $(b^{i-1}, l^j, t^{i-1,j})$

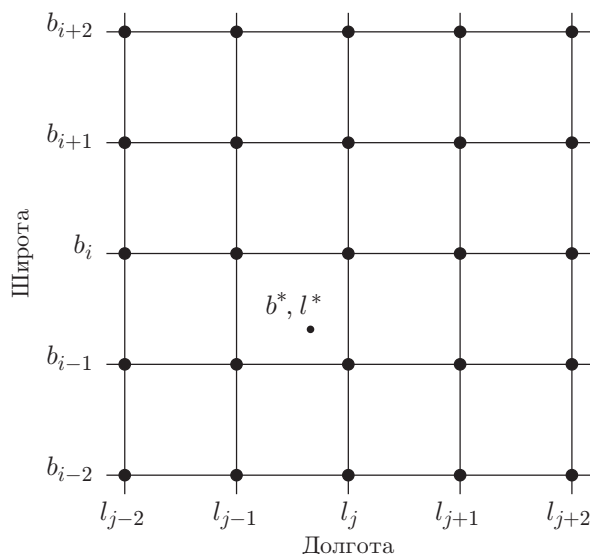


Рис. 1. Входные данные аппроксиматора в виде фрагмента пространственной сетки

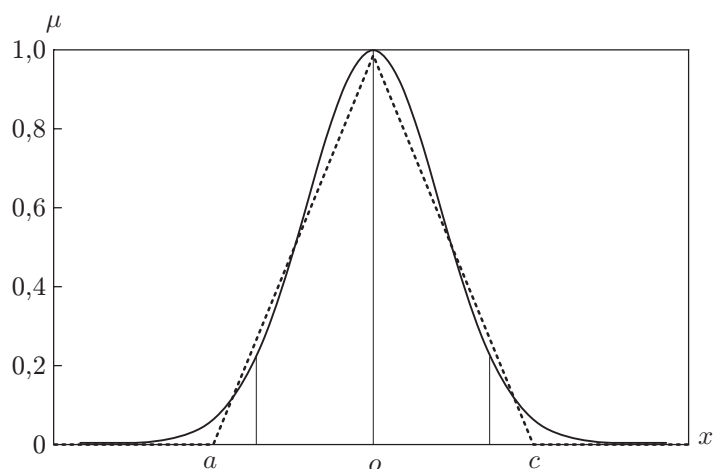


Рис. 2. Треугольная (пунктирная кривая) и гауссова (сплошная кривая) функции принадлежности терма «около o »

и $(b^i, l^{j-1}, t^{i,j-1})$. На обучающей выборке формируется нечёткая система типа Такаги — Сугено нулевого порядка, которая содержит 14 правил следующего вида:

- правило 1: ЕСЛИ $b = \text{«около } b^{i+1}\text{»}$ И $l = \text{«около } l^{j-2}\text{»}$ ТО $t^{i+1,j-2}$;
- правило 2: ЕСЛИ $b = \text{«около } b^{i+1}\text{»}$ И $l = \text{«около } l^{j-1}\text{»}$ ТО $t^{i+1,j-1}$;
- правило 3: ЕСЛИ $b = \text{«около } b^{i+1}\text{»}$ И $l = \text{«около } l^j\text{»}$ ТО $t^{i+1,j}$;
- правило 4: ЕСЛИ $b = \text{«около } b^{i+1}\text{»}$ И $l = \text{«около } l^{j+1}\text{»}$ ТО $t^{i+1,j+1}$;
- правило 5: ЕСЛИ $b = \text{«около } b^i\text{»}$ И $l = \text{«около } l^{j-2}\text{»}$ ТО $t^{i,j-2}$;
- правило 6: ЕСЛИ $b = \text{«около } b^i\text{»}$ И $l = \text{«около } l^j\text{»}$ ТО $t^{i,j}$;
- правило 7: ЕСЛИ $b = \text{«около } b^i\text{»}$ И $l = \text{«около } l^{j+1}\text{»}$ ТО $t^{i,j+1}$;
- правило 8: ЕСЛИ $b = \text{«около } b^{i-1}\text{»}$ И $l = \text{«около } l^{j-2}\text{»}$ ТО $t^{i-1,j-2}$;
- правило 9: ЕСЛИ $b = \text{«около } b^{i-1}\text{»}$ И $l = \text{«около } l^{j-1}\text{»}$ ТО $t^{i-1,j-1}$;
- правило 10: ЕСЛИ $b = \text{«около } b^{i-1}\text{»}$ И $l = \text{«около } l^{j+1}\text{»}$ ТО $t^{i-1,j+1}$;
- правило 11: ЕСЛИ $b = \text{«около } b^{i-2}\text{»}$ И $l = \text{«около } l^{j-2}\text{»}$ ТО $t^{i-2,j-2}$;
- правило 12: ЕСЛИ $b = \text{«около } b^{i-2}\text{»}$ И $l = \text{«около } l^{j-1}\text{»}$ ТО $t^{i-2,j-1}$;
- правило 13: ЕСЛИ $b = \text{«около } b^{i-2}\text{»}$ И $l = \text{«около } l^j\text{»}$ ТО $t^{i-2,j}$;
- правило 14: ЕСЛИ $b = \text{«около } b^{i-2}\text{»}$ И $l = \text{«около } l^{j+1}\text{»}$ ТО $t^{i-2,j+1}$,

где b и l — входные переменные (широта и долгота соответственно); t — выходная переменная (температура); «около b^i » или «около l^j » — функции принадлежности, определённые на переменных широты и долготы соответственно. Пример таких функций приведён на рис. 2.

После формирования структуры аппроксиматора необходимо идентифицировать параметры функций принадлежности (a_i, o_i, c_i для треугольной и o_i, σ_i для гауссовой) и параметры консеквентов.

3. Идентификация параметров функций принадлежности ведётся с помощью генетического алгоритма. Начинает работу генетический алгоритм с некоторого набора исходных решений, называемых популяцией. Каждое решение (элемент популяции) представляет собой строку символов, которая называется особью или хромосомой. Хромосома состоит из множества генов. В нашем случае ген кодирует один параметр нечёткой модели, т. е. хромосомы представлены в виде вектора действительных чисел. В нулевой ген помещается значение ошибки вывода нечёткой системы (2). Схема строения хромосомы приведена на рис. 3.

Ошибка	a_{b1}	o_{b1}	c_{b1}	a_{b2}	o_{b2}	c_{b2}	...	a_{l1}	o_{l1}	c_{l1}	...	a_{l4}	o_{l4}	c_{l4}
--------	----------	----------	----------	----------	----------	----------	-----	----------	----------	----------	-----	----------	----------	----------

Рис. 3. Пример кодирования решения

При формировании хромосомы должны быть соблюдены соответствующие требования к значениям параметров функций принадлежности. Для треугольной функции они следующие:

- 1) $a_{xi} \leq o_{xi} \leq c_{xi}$,
- 2) $o_{xi-1} \leq o_{xi} \leq o_{xi+1}$,

где x — индекс b или l .

В процессе работы алгоритма хромосомы изменяются под воздействием генетических операторов (скрещивание, мутация). Хромосомы с меньшей ошибкой при выполнении эволюционной операции селекции имеют больше шансов попасть в следующее поколение.

4. Идентификация параметров консеквентов проводится на основе метода наименьших квадратов.

Представим правило из приведённой выше базы следующим образом:

правило i : ЕСЛИ $b = A_{bi}$ И $l = A_{li}$ ТО t_i ,

где A_{bi} и A_{li} — нечёткие множества, заданные своими функциями принадлежности. Тестовая выборка сформирована множеством троек вида (b_k, l_k, t_k^*) . Необходимо минимизировать функцию

$$f = \sum_{k=1}^K (t_k^* - y_k)^2 \rightarrow \min,$$

где y_k — выходное значение, полученное в результате нечёткого вывода для входных значений (b_k, l_k) ; K — число строк в таблице наблюдений. После дифференцирования по параметру консеквента будем иметь:

$$\frac{\partial f}{\partial t_s} = 2 \sum_{k=1}^K \left[\left(t_k^* - \frac{\sum_{i=1}^R t_i \mu_{A_{bi}}(b_k) \mu_{A_{li}}(l_k)}{\sum_{i=1}^R \mu_{A_{bi}}(b_k) \mu_{A_{li}}(l_k)} \right) \frac{\mu_{A_{bs}}(b_k) \mu_{A_{ls}}(l_k)}{\sum_{i=1}^R \mu_{A_{bi}}(b_k) \mu_{A_{li}}(l_k)} \right] = 0,$$

$$\sum_{k=1}^K t_k^* \frac{\mu_{A_{bs}}(b_k) \mu_{A_{ls}}(l_k)}{\sum_{i=1}^R \mu_{A_{bi}}(b_k) \mu_{A_{li}}(l_k)} = \sum_{i=1}^R \sum_{k=1}^K \frac{t_i \mu_{A_{bi}}(b_k) \mu_{A_{li}}(l_k) \mu_{A_{bs}}(b_k) \mu_{A_{ls}}(l_k)}{\left(\sum_{j=1}^R \mu_{A_{bj}}(b_k) \mu_{A_{lj}}(l_k) \right)^2},$$

$$\sum_{k=1}^K t_k^* \frac{\mu_{A_{bs}}(b_k) \mu_{A_{ls}}(l_k)}{\sum_{i=1}^R \mu_{A_{bi}}(b_k) \mu_{A_{li}}(l_k)} =$$

$$= \sum_{i=1}^R t_i \frac{\sum_{k=1}^K \left[\mu_{A_{bi}}(b_k) \mu_{A_{li}}(l_k) \mu_{A_{bs}}(b_k) \mu_{A_{ls}}(l_k) \prod_{m=1, m \neq k}^K \left(\sum_{j=1}^R \mu_{A_{bj}}(b_m) \mu_{A_{lj}}(l_m) \right)^2 \right]}{\prod_{k=1}^K \left(\sum_{j=1}^R \mu_{A_{bj}}(b_k) \mu_{A_{lj}}(l_k) \right)^2}.$$

Пусть

$$\mathbf{B} = \begin{bmatrix} \sum_{k=1}^K t_k^* \frac{\mu_{A_{b1}}(b_k) \mu_{A_{l1}}(l_k)}{\sum_{i=1}^R \mu_{A_{bi}}(b_k) \mu_{A_{li}}(l_k)} \\ \sum_{k=1}^K t_k^* \frac{\mu_{A_{b2}}(b_k) \mu_{A_{l2}}(l_k)}{\sum_{i=1}^R \mu_{A_{bi}}(b_k) \mu_{A_{li}}(l_k)} \\ \vdots \\ \sum_{k=1}^K t_k^* \frac{\mu_{A_{bR}}(b_k) \mu_{A_{lR}}(l_k)}{\sum_{i=1}^R \mu_{A_{bi}}(b_k) \mu_{A_{li}}(l_k)} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_R \end{bmatrix},$$

$$\mathbf{A} = [A_{wi}] = \left[\frac{\sum_{k=1}^K \left[\mu_{A_{bi}}(b_k) \mu_{A_{li}}(l_k) \mu_{A_{bw}}(b_k) \mu_{A_{lw}}(l_k) \prod_{m=1, m \neq k}^K \left(\sum_{j=1}^R \mu_{A_{bj}}(b_m) \mu_{A_{lj}}(l_m) \right)^2 \right]}{\prod_{k=1}^K \left(\sum_{j=1}^R \mu_{A_{bj}}(b_k) \mu_{A_{lj}}(l_k) \right)^2} \right],$$

где R — число правил нечёткой модели; $w, i = 1, \dots, R$.

Тогда $\mathbf{A} \times \mathbf{X} = \mathbf{B}$. Если \mathbf{A} — невырожденная ($\det(\mathbf{A}) \neq 0$), то $\mathbf{X} = \mathbf{A}^{-1} \times \mathbf{B}$.

5. Эксперимент. Цель эксперимента — определение возможностей нечёткого аппроксиматора и исследование влияния параметров нечёткой системы и генетического алгоритма на качество аппроксимации.

Проведение эксперимента требовало решения следующих проблем: какие данные использовать? как много данных требуется? как эти данные взять?

Первая часть эксперимента проводилась на двух контрольных точках с координатами (80° с. ш., 10° з. д.) на уровне давления 1000 мбар и (50° с. ш., 50° з. д.) на уровне давления 600 мбар. Реальные значения температур в указанных точках известны.

Эксперимент осуществлялся на пространственной сетке 5×5 (см. рис. 1), контрольной была точка $(b^i, l^j, t^{i,j})$. Тестовая выборка состояла из четырёх точек: $(b^i, l^{j-1}, t^{i,j-1})$, $(b^{i+1}, l^j, t^{i+1,j})$, $(b^i, l^{j+1}, t^{i,j+1})$, $(b^{i-1}, l^j, t^{i-1,j})$. Ещё двадцать точек использовалось в качестве обучающих, на их основе строилась база из двадцати правил (принцип построения изложен в разд. 2).

Проблема «как взять данные» решалась путём создания оригинального адаптивного программного комплекса для работы с данными netCDF.

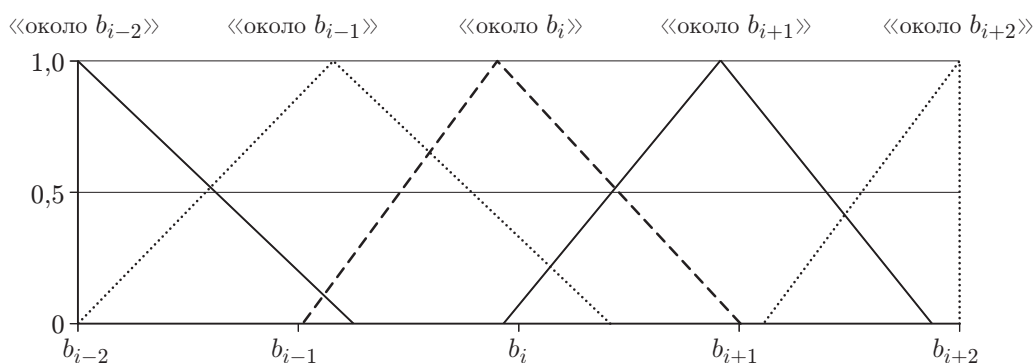


Рис. 4. Нечёткие значения входной переменной «широта»

Исследовалась зависимость точности и скорости аппроксимации от параметров системы. На каждом этапе изменялся один параметр работы генетического алгоритма, остальные ставились по умолчанию: тип функции принадлежности — гауссова, размер популяции — 10, тип селекции — случайная, тип скрещивания — арифметическое, тип мутации — одноточечная, вероятность мутации — 0,7, количество итераций — 100. Для каждого испытания проводилось 50 запусков алгоритма и находились усреднённые значения ошибки аппроксимации и времени работы алгоритма. Каждая входная переменная представлена пятью термами, равномерно покрывающими всю область определения переменной на первой итерации. Пример переменной «широта» и пять её термов показаны на рис. 4. Количественной мерой критерия эффективности и оптимальности является ошибка и время поиска решения.

Результаты влияния типа функций принадлежности на качество и скорость аппроксимации приведены в табл. 1. Из полученных данных видно, что скорость работы алгоритма мало зависит от типа функций принадлежности. Треугольные функции принадлежности обрабатываются несколько медленнее из-за более сложного алгоритма проверки включения хромосомы в популяцию. Ошибка аппроксимации существенно зависит от типа функций принадлежности. Применение гауссовых функций принадлежности улучшает точность работы алгоритма более чем в 2 раза, поэтому можно сделать вывод о необходимости их использования.

Для определения оптимального размера популяции были взяты четыре разных размера популяции: 10, 30, 50 и 100 хромосом. Результат исследования представлен в табл. 2. При значительном увеличении размера популяции усреднённая ошибка уменьшается незначительно, а время работы системы увеличивается пропорционально количеству хромосом в популяции. На основании анализа полученных данных можно сделать вывод о том, что наиболее эффективно можно использовать размер популяции около 10 хромосом.

В данной работе исследовалось четыре типа селекции хромосом: случайный отбор,

Таблица 1

Влияние типа функций принадлежности на работу системы

Точка	Параметры системы	Тип функции принадлежности	
		гауссова	треугольная
80° с. ш., 10° з. д.	Ошибка	0,20471096	0,58500511
	Время, с	4,12	4,48
50° с. ш., 50° з. д.	Ошибка	0,15930731	0,34780164
	Время, с	4,09	4,22

Таблица 2

Влияние размера популяции на работу системы

Точка	Параметры системы	Размер популяции				
		5	10	30	50	100
80° с. ш., 10° з. д.	Ошибка	0,3829436	0,1958879	0,1846638	0,1702530	0,1553917
	Время, с	3,51	4,44	9,51	15,6	29,5
50° с. ш., 50° з. д.	Ошибка	0,2914332	0,1518879	0,1466638	0,1342530	0,1193917
	Время, с	3,45	4,46	9,45	15,2	29,6

Таблица 3

Влияние типа оператора скрещивания на работу системы

Точка	Параметры системы	Тип популяции				
		арифметическая	одно-точечная	двух-точечная	много-точечная	унифицированная
80° с. ш., 10° з. д.	Ошибка	0,2078317	0,3745043	0,4126968	0,205355	0,22184
	Время, с	4,29	5,48	5,30	7,02	5,08
50° с. ш., 50° з. д.	Ошибка	0,1478317	0,3145043	0,3291344	0,161355	0,17384
	Время, с	4,18	5,39	5,28	6,95	5,24

Таблица 4

Влияние типа мутации на работу системы

Точка	Параметры системы	Тип мутации	
		одноточечная	многоточечная
80° с. ш., 10° з. д.	Ошибка	0,19187503	0,22490686
	Время, с	3,78	5,28
50° с. ш., 50° з. д.	Ошибка	0,14587503	0,16490686
	Время, с	3,81	5,22

Таблица 5

Влияние количества итераций на работу системы

Точка	Параметры системы	Размер популяции			
		50	100	300	500
80° с. ш., 10° з. д.	Ошибка	0,32149891	0,20569980	0,19695792	0,19443969
	Время, с	2,15	3,73	10,8	16,6
50° с. ш., 50° з. д.	Ошибка	0,24149891	0,16369980	0,15895792	0,15443969
	Время, с	2,06	3,64	11,1	16,8

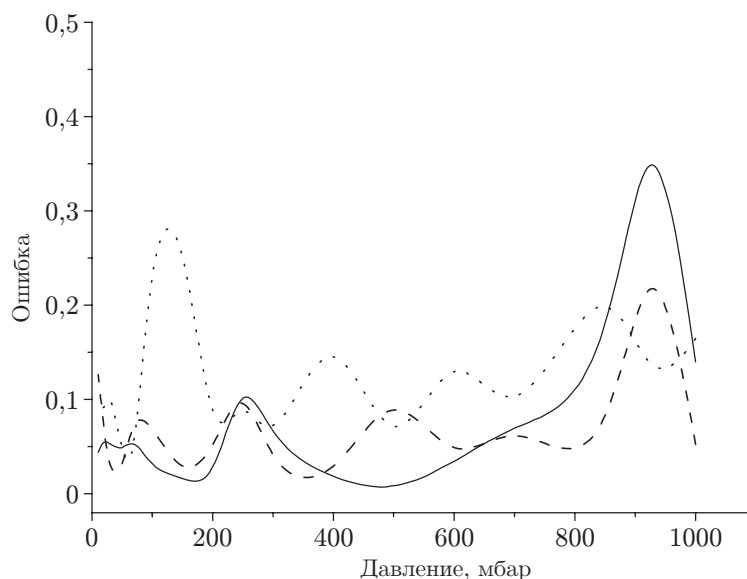


Рис. 5. Зависимости ошибки от давления (сплошная кривая — полюс, пунктирная — средние широты, штриховая — экватор)

турнирный, рулеточный, по стратегии элитаризма. Эксперименты показали, что тип селекции не оказывает существенного влияния на ошибку аппроксимации и время вывода.

В эксперименте использовались операторы пяти видов скрещивания: арифметическое, одноточечное, двухточечное, многоточечное и унифицированное. В табл. 3 сведены результаты влияния типа оператора скрещивания на ошибку и скорость аппроксимации. По полученным данным установлено, что арифметическое, многоточечное и унифицированное скрещивания дают практически одинаковую ошибку, однако арифметическое скрещивание превосходит конкурентов по скорости работы и потому является оптимальным.

Результаты исследования влияния типа мутации на время и точность аппроксимации приведены в табл. 4. Из таблицы видно, что одноточечная мутация предпочтительнее многоточечной как по времени расчёта, так и по ошибке.

Количество итераций является одним из критериев останова, поэтому его значение существенно влияет на качество работы аппроксиматора. Работа алгоритма была исследована на 50, 100, 300 и 500 шагах. Результат исследования представлен в табл. 5. По полученным данным видно, что при увеличении количества итераций с 50 до 100 точность аппроксимации возрастает соизмеримо с уменьшением времени работы. При дальнейшем увеличении числа итераций значительного повышения точности не наблюдается, в то время как скорость работы падает существенно. На основании анализа полученных данных можно сделать вывод о том, что в алгоритме нечёткой аппроксимации эффективно можно использовать около 100 итераций.

После выявления эффективных параметров системы были произведены серии экспериментов по проверке качества аппроксимации в различных климатических поясах, которые составили вторую часть эксперимента. В качестве тестовых климатических зон выбирались следующие: экваториальная (5° с. ш.), средних широт (40° с. ш.) и полярных широт (80° с. ш.). Для каждой точки проводилась аппроксимация по всем имеющимся уровням высоты (давления). Значения ошибки аппроксимации для каждой из трёх климатических зон показаны на рис. 5. Анализ результатов аппроксимации, представленных на рисунке, показывает, что её ошибка не превышает 0,4 К.

Заключение. В результате проведённых в данной работе экспериментов сделаны следующие выводы.

Скорость работы алгоритма мало зависит от типа функций принадлежности. Применение гауссовых функций принадлежности повышает точность работы алгоритма более чем в 2 раза.

Для построения нечёткого аппроксиматора рекомендуется использовать около 100 итераций генетического алгоритма, размер популяции около 10 хромосом, арифметическое скрещивание и одноточечную мутацию.

Ошибка аппроксимации менее 0,4 К удовлетворяет требованиям к качеству моделей для глобального кратко- и среднесрочного прогнозирования [12], а разработанный нечёткий аппроксиматор может использоваться для проведения численных расчётов, например, полей излучения, проходящих сквозь толщу атмосферы.

СПИСОК ЛИТЕРАТУРЫ

1. **Uppala S., Kallberg P., Hernandez A. et al.** ERA-40: ECMWF 45-year reanalysis of the global atmosphere and surface conditions 1957–2002 // ECMWF Newsletter. Reading, United Kingdom. 2004. N 101. P. 2–21.
2. **Kalnay E., Kanamitsu M., Kistler R. et al.** The NCEP/NCAR 40-year reanalysis project // Bull Am. Meteorol. Soc. 1996. **77**, N 3. P. 437–471.
3. **Катаев М. Ю., Катаев С. Г., Кусков А. В.** Интерполяция метеорологической информации базы данных NCEP на произвольную пространственную точку // Тр. Междунар. конф. по измерениям, моделированию и информационным системам для изучения окружающей среды (ENVIROMIS-2008). Томск: Изд-во Томского ЦНТИ, 2008. С. 55.
4. **Yang Q., Parvin B.** High-resolution reconstruction of sparse data from dense low-resolution spatio-temporal data // IEEE Trans. Image Process. 2003. **12**, N 6. P. 671–677.
5. **Solomatine D. P.** Applications of data-driven modelling and machine learning in control of water resources // Computational Intelligence in Control /Eds. M. Mohammadian, R. A. Sarker, X. Yao. London: Idea Group Publishing, 2002. P. 197–217.
6. **Kosko B.** Fuzzy systems as universal approximators // IEEE Trans. Comput. 1994. **43**, N 11. P. 1329–1333.
7. **Wang L. X., Mendel J. M.** Fuzzy basis functions, universal approximation, and orthogonal least-squares learning // IEEE Trans. Neural Networks. 1992. **3**, N 5. P. 807–814.
8. **Ходашинский И. А.** Идентификация нечетких систем: методы и алгоритмы // Проблемы управления. 2009. № 4. С. 15–23.
9. **Ходашинский И. А., Дудин П. А.** Параметрическая идентификация нечетких моделей на основе гибридного алгоритма муравьиной колонии // Автометрия. 2008. **44**, № 5. С. 24–35.
10. **Ходашинский И. А., Гнездилова В. Ю., Дудин П. А., Лавыгина А. В.** Основанные на производных и метаэвристические методы идентификации параметров нечетких моделей // Тр. VIII Междунар. конф. «Идентификация систем и задачи управления» (SICPRO'08). М.: Институт проблем управления им. В. А. Трапезникова РАН, 2009. С. 501–529.
11. **Rew R., Davis G., Emmerson S. et al.** The NetCDF users guide. Data model, programming interfaces, and format for self-describing, portable data. NetCDF version 4.0 // Unidata Program Center, 2008. 74 p.
<http://www.unidata.ucar.edu/software/netcdf/>(дата обращения 01.03.2010).
12. **Наставление** по глобальной системе обработки данных и прогнозирования: технический регламент № 485. М.: ВМО, 2005. 173 с.

Поступила в редакцию 21 июля 2009 г.