УДК 519.2:534

ТЕОРЕТИКО-ВОЗМОЖНОСТНЫЙ ПРОГНОЗ СРЕДНЕМЕСЯЧНЫХ ЗНАЧЕНИЙ ТЕМПЕРАТУРЫ

Ю. П. Пытьев, И. В. Мазаева

(*кафедра компьютерных методов физики*)

E-mail: prudnikov@phys.msu.su

В работе исследуется долгосрочный прогноз среднемесячных значений температуры теоретико-возможностным методом редукции измерений при наличии априорной информации о входных данных. Возможность события ориентирована на относительную оценку истинности данного события в ранговой шкале, в которой могут быть содержательно истолкованы лишь отношения «больше», «меньше» или «равно». Таким образом, результаты прогноза характеризуются на уровне терминов «более возможно» или «менее возможно» иметь определенную температуру в данном месяце.

Исследуется долгосрочный прогноз (на месяц или сезон) по известным среднемесячным значениям температуры за период 1958–1990 гг. на 98 станциях, распределенных по всей территории бывшего СССР. Вследствие отсутствия достаточно разработанной общей методики долгосрочных прогнозов, основывающейся на строгих теоретических положениях или на глубоком анализе физики атмосферных процессов большого масштаба, мы вынуждены принять факт, что выводы, полученные эмпирическим путем, не имеют достаточно полного физического истолкования.

Для решения задачи долгосрочного прогнозирования использовались методы теоретико-возможностной редукции измерения при наличии априорной информации о входных данных.

В теории возможностей [1] каждому событию S ставится в соответствие возможность $\mathrm{P}(S)$, ориентированая на относительную оценку его истинности, его предпочтительности в сравнении с любым другим событием, причем в ранговой (порядковой) шкале, в которой могут быть представлены и содержательно истолкованы лишь отношения «больше», «меньше» или «равно» и дуальная ей необходимость $N(S)$. Под событием S понимается любое подмножество множества X элементарных событий. Возможность $\mathrm{P}(\cdot)$ является мерой, определенной на σ-алгебре $\mathrm{P}(X)$ всех подмножеств X и принимающей значения в шкале $\mathcal{L}$. В отличие от вероятности, принимающей значения в «абсолютной» шкале, возможность принимает значения в шкале $\mathcal{L}=([0,1],\leqslant,+,\bullet)$, определенной как отрезок $[0,1]$, где сложение $+$ определено как $\max$, а умножение $\bullet$ — как $\min$. Шкала $\mathcal{L}$ инвариантна относительно группы непрерывных строго монотонных преобразований отрезка $[0,1]$ в себя, оставляющих неподвижными точки 0 и 1. Инвариантность шкалы $\mathcal{L}$ значений возможности определяет принцип относительности, согласно которому в теории возможностей каждый исследователь при построении модели может использовать «свою» шкалу, которая служит ему собственной системой «отсчета», в отличие от «абсолютной шкалы» значений вероятности. Нечеткий вектор в теории возможностей является прямым аналогом случайного вектора в теории вероятностей и может быть охарактеризован распределением возможностей его значений.

Пусть схема измерения среднемесячных значений температуры имеет вид равенства

$$\xi = \mathrm{A}\varphi + \nu, \tag{1}$$

где φ, ν — нечеткие элементы, $\varphi\in\mathcal{F}$ — входные данные, содержащие частично неизвестную прогнозируемую компоненту, $\nu\in\mathcal{R}$ — погрешность измерений, ξ — имеющиеся данные о среднемесячных значениях температуры; A — линейный оператор, связывающий входные и выходные данные.

В работе рассмотрена теоретико-возможностная модель измерений (1) среднемесячных значений температуры, заданная совместным распределением возможностей значений следующих нечетких элементов: выходного сигнала ξ (имеющихся данных о среднемесячных значениях температуры), входного сигнала φ (содержащего прогнозируемые данные) и погрешности измерений ν. В задаче прогнозирования, рассматриваемой как задача интерпретации измерения (1), зададим линейный оператор $U:\ \mathcal{F}\to\mathcal{U}$, определяющий представляющие интерес прогнозируемые параметры $u\in\mathcal{U}$ данных φ: $u=U\varphi$, $\mathcal{R}$, $\mathcal{F}$ и $\mathcal{U}$ — евклидовы пространства. В задаче прогнозирования требуется определить стратегию прогнозирования $d(\cdot):\mathcal{R}\to\mathcal{U}$ так, чтобы элемент $d(\xi)$ можно было считать самой точной версией (оценкой) нечеткого элемента $U\varphi$ [1–2].

Рассмотрим решение задачи прогнозирования. Пусть заданы линейный оператор A и распределения π^{φ}, π^{ν} нечетких элементов φ, ν, которые будем считать независимыми: $\pi^{\varphi}(f)$ — возможность равенства $\varphi=f\in\mathcal{F}$, $\pi^{\nu}(n)$ — возможность равенства $\nu=n\in\mathcal{R}$. Тем самым будет задана теоретико-возможностная модель $[\mathrm{A},\pi^{\varphi}(\cdot),\pi^{\nu}(\cdot)]$ схемы измерения (1) [1].

Качество стратегии оценивания характеризуем величиной необходимости ошибки прогнозирования [1]:

$$N(d(\cdot)) =$$
$$= \theta\left(\sup_{x\in\mathcal{R}}\min_{f\in\mathcal{F}}(\pi^{\nu}(x-\mathrm{A}f), \pi^{\varphi}(f), \theta\circ l(Uf, d(x)))\right) =$$
$$= \inf_{x\in\mathcal{R}}\max_{f\in\mathcal{F}}\left(\theta\circ\pi^{\nu}(x-\mathrm{A}f), \theta\circ\pi^{\varphi}(f), l(Uf, d(x))\right),$$

где $l(x,y)$ — возможность ошибки, обусловленной заменой $x\in X$ на $y\in X$; $\theta(\cdot):\ [0,1]\to[0,1]$ — инволюция; $\pi^{\nu}(x-\mathrm{A}f)$ — условная возможность равенства $\xi=x$ при условии $\varphi=f$.

N-оптимальная стратегия $d^*(\cdot)$ определяется условием $\mathrm{N}(d^*(\cdot)) = \min\limits_{d(\cdot):\mathcal{R}\to\mathcal{U}}\mathrm{N}(d(\cdot))$, согласно которому при использовании стратегии $d^*(\cdot)$ необходимость ошибки прогнозирования $U\varphi$ посредством $d^*(\xi)$ минимальна.

Решение задачи в рамках этой модели получено в [1] и дается равенством

$$d^*(x) = U\tilde{\varphi},$$

где $\tilde{\varphi} = f_0 + F\mathrm{A}^*(\mathrm{A}F\mathrm{A}^* + \omega(x)\Sigma)^{-1}(x-\mathrm{A}f_0)$, $\omega=\omega(x)$ — корень уравнения $\omega||(BB^* + \omega I)^{-1}y|| = ||B^*(BB^*+\omega I)^{-1}y||$, $B=\Sigma^{1/2}\mathrm{A}F^{1/2}$, $y=\Sigma^{-1/2}(x-\mathrm{A}f_0)$.

Вектор x состоит из данных о возможных значениях температуры по группе станций. Группу станций подбираем, исходя из периодичности изменений температуры в каждой из станций. Для того чтобы выявить периодичность изменений температуры, к последовательности данных по каждому месяцу и по каждой станции за период с 1958 по 1990 г. применим комплексное дискретное преобразование Фурье. В результате такого преобразования получим комплексный вектор T, который состоит из спектральных составляющих комплексных отсчетов последовательности. Теперь для отображения мощности спектральных составляющих возведем действительную часть вектора T в квадрат и построим график зависимости T от n, где $n=1,\ldots,16$, причем не учитываются значения больше 16, поскольку становится очевидным, что в данной последовательности периодичность не наблюдается. Наиболее мощный (максимальный по величине) спектр по оси ординат будет соответствовать периоду по оси абсцисс. В качестве примера на рис. 1 приведен график определения периода изменений температуры для Москвы на ноябрь. Из графика видно, что период изменения температуры ноября в Москве равен приблизительно 5 годам. Рассмотрим распределение периодов по всей территории СССР (рис. 2) на примере сентября.

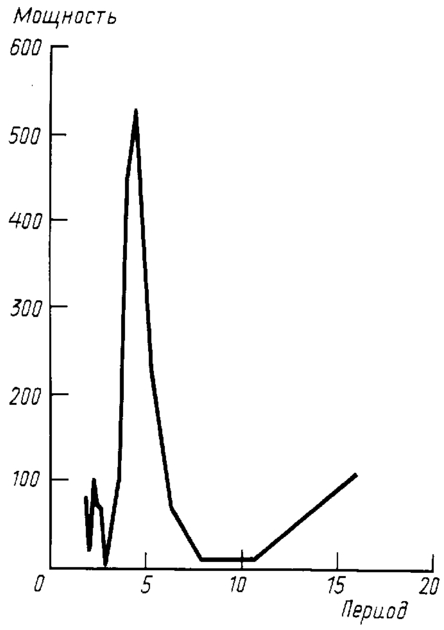

Рис. 1

На рис. 2 показано, что в большинстве регионов имеются свои области с одинаковыми значениями периода изменений температуры для группы станций. Для построения ковариационного оператора данные метеорологических станций группируются на основе подбора областей с одинаковым периодом изменения температуры. В работе рассмотрена интерпретация возможности на примере одного месяца московской метеорологической станции. На рис. 3 представлена теоретико-возможностная редукция измерения $\tilde{\varphi}$, минимизирующая необходимость ошибки оценивания, на декабрь 1990 г. (*а*) и 1978 г. (*б*) для метеорологической станции, расположенной в Москве. Зависимость значений возможности от числа лет, используемых для оценивания F, и от температуры показана в градациях серого. По оси X заданы возможные значения температуры, по оси Y — количество лет. Глубина цвета отображает распределение возможностей и варьируется от черного до белого. Чем «светлее» область, тем «более возможно», что температура попадет в данный интервал. Так, например, на рис. 3, *а* и *б* показано, что для декабря лучше делать прогноз на основе 3–5 лет, а результатом прогноза является интервал температур от -4 до $-2°$ для 1990 г. и от -14 до $-12°$ для 1978 г.

Соответствующие среднеквадратичные ошибки (*б*) при прогнозе на основе 5 лет отображены на рис. 4, *а* и *б*.

Аналогично для метеорологической станции, расположенной в Санкт-Петербурге (рис. 5, *а* и *б*), получим, что прогноз лучше делать на основе 2–4 лет, а результатом прогноза является интервал температуры для декабря 1990 г. от -2 до $+2°$, для 1978 г. от -13 до $-12°$.

Результатом прогноза могут также оказаться и два интервала температур. С метеорологической

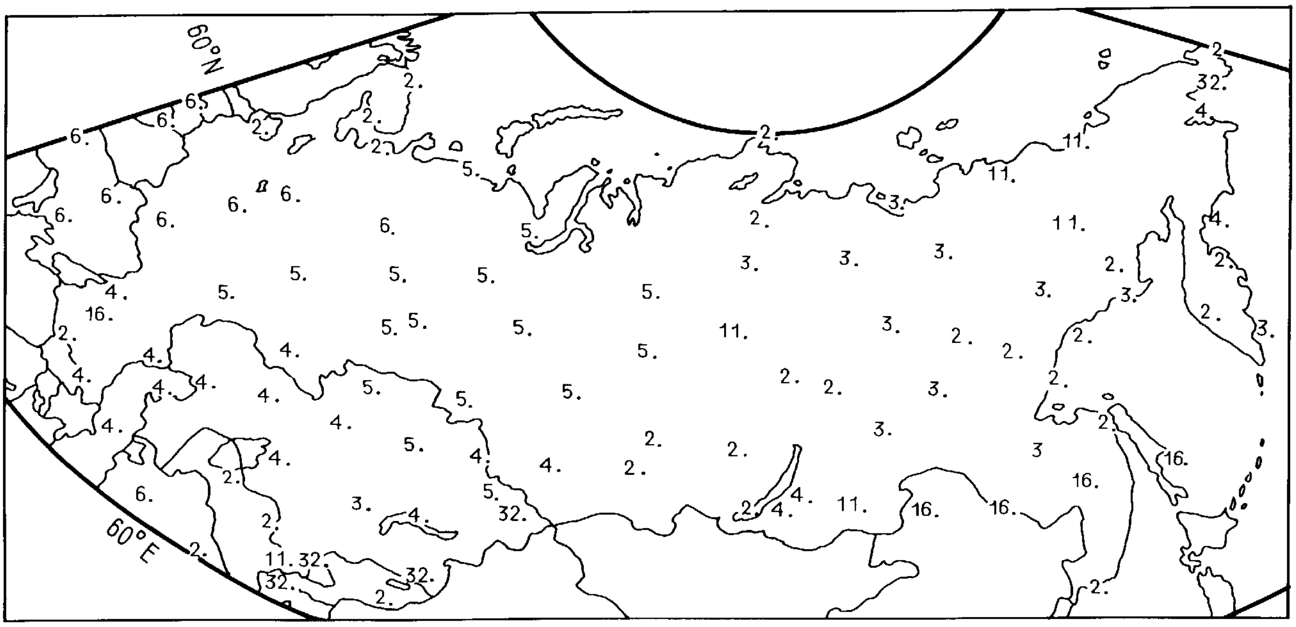

Рис. 2

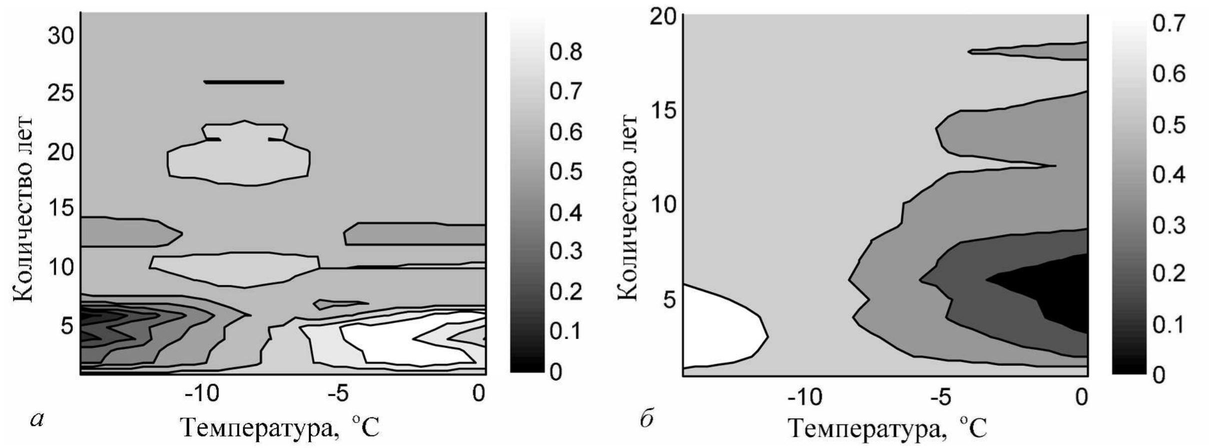

Рис. 3

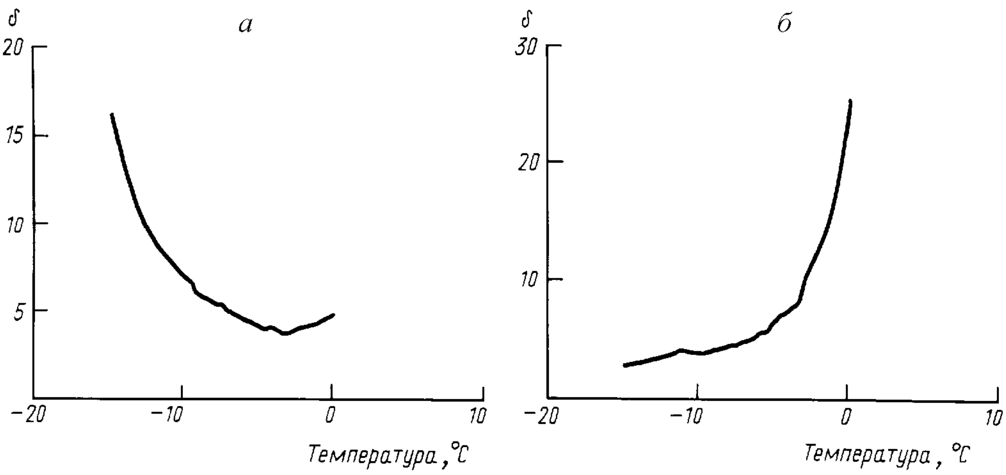

Рис. 4

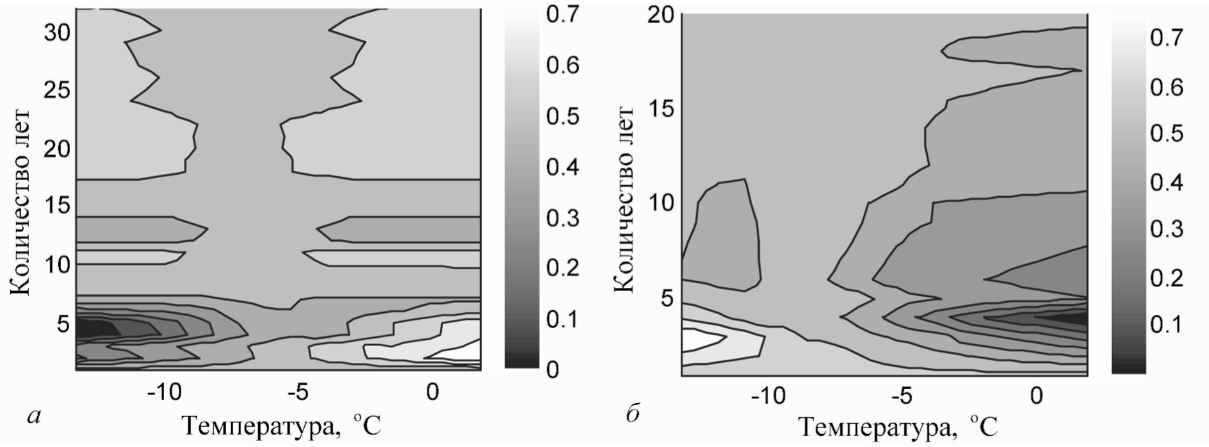

Рис. 5

точки зрения такие случаи можно объяснить расположением метеорологической станции на границе двух разных температурных фронтов, в связи с чем при малом изменении начальных условий (значений температуры в предшествующем месяце) возможно сильное изменение средней температуры месяца.

Таким образом, можно не использовать шкалу значений возможности от 0 до 1 и характеризовать результаты прогноза на уровне слов «более возможно» или «менее возможно».

Литература

1. *Пытьев Ю.П.* Возможность. Элементы теории и применения. М.: Эдиториал УРСС, 2000.
2. *Пытьев Ю.П.* Математические методы интерпретации эксперимента. М.: Высшая школа, 1989.

Поступила в редакцию
07.06.02