ОСНОВЫ МЕТОДА ОПРЕДЕЛЕНИЯ СИСТЕМЫ ЭМПИРИЧЕСКИХ УРАВНЕНИЙ ДЛЯ ПРОГНОЗИРОВАНИЯ ПРОЦЕССОВ В АТМОСФЕРЕ

Степаненко С.Р.

Основой метода являются: 1) способ описания множества реализаций с помощью функции с разделенными переменными (метод Фурье), 2) метод «мягкого» сглаживания для фильтрации погрешностей эмпирических характеристик. Приведен пример получения системы эмпирических уравнений по данным 8-ми срочных метеорологических наблюдений. Начальное приближение решения имеет высокие статистические характеристики, позволяющие использовать его в оперативной практике Гидрометеорологического центра. Система уравнений для полушария может быть основой физической теории климата, альтернативной существующему подходу на основе уравнений глобальной циркуляции атмосферы. Изложенный подход можно использовать для разработки модели архивных данных ВНИИГМИ-МЦД, что позволило бы существенно повысить качество гидрометеорологического обеспечения субъектов народного хозяйства.

Введение. Факт существования глобальных (климатических) колебаний погодных условий на нашей планете в настоящее время признается всеми. Но, во-первых, нет однозначного мнения относительно того, что является главной причиной нынешнего аномального характера погоды на Северном полушарии. Во-вторых, отсутствует достаточно надежный метод прогнозирования глобальных изменений. Поэтому пока опасно предпринимать какие-либо действия, направленные на нейтрализацию предполагаемых изменений, поскольку такие действия могут создать другие, вполне реальные угрозы.

Основная проблема прогнозирования климатических изменений связана с тем, что климатическая система представляет сложную (открытую неоднородную неравновесную (нелинейную) самоорганизующуюся) динамическую систему. Это порождает широкий диапазон пространственно-временных взаимообусловленных колебаний [2]. Изменения климатического масштаба обусловлены флуктуациями синоптического масштаба. Синоптические процессы взаимосвязаны с мезомасштабными явлениями и т.д. Поэтому решение проблемы описания и прогнозирования климата, по мнению К. Нассельмана [12], следует рассматривать как проблему «конструирования связанных моделей», описывающих процессы всех основных пространственно-временных масштабов. Один и тот же процесс в этой системе в одном уравнении должен выступать в виде закономерных изменений (как момент первого порядка), а в другом – как динамический хаос (в виде момента второго порядка).

Цель настоящей работы – рассмотреть основы метода нахождения системы эмпирических моделей для прогнозирования процессов в атмосфере разных временных масштабов на примере анализа значений температуры воздуха, полученных по данным 8-ми срочных метеорологических наблюдений.

Основная гипотеза и краткое описание метода определения системы эмпирических уравнений. Предполагается, что реализации (решения) нелинейных природных процессов, полученные при разных начальных условиях, являются линейно зависимыми между собой.

Реализации процесса могут быть двух типов. Первый тип представляет законченный цикл изменений во времени, например суточные и сезонные колебания приземной температуры воздуха. Второй тип реализаций представляет распределение метеорологических переменной на полушарии в фиксированный момент времени.

Для выполнения основной гипотезы *необходимо и достаточно*, чтобы каждую реализацию можно было бы представить в виде функции с разделенными переменными.

Для реализаций первого типа функция с разделенными переменными имеет вид

$$u_k(\xi) = f(\xi,\omega_k) = a_k + b_k f(\xi), \quad k = 1,2,\dots, M, \quad \omega_k = \{a_k, b_k\}, \tag{1}$$

где k– номер реализации, a_k b_k– параметры, постоянные для k-той реализации, **ξ** -

время от начала реализации, $\xi = [0,T]$.

Функцию с разделенными переменными реализаций второго типа можно представить в виде

$$u_k(x,y) = f(x,y,\omega_k) = a_k + b_k(x,y), \quad k=1,2,\ldots, M, \quad \omega_k = \{a_k\ b_k\} \qquad (2)$$

где x,y – горизонтальные координаты.

Выражения (1), (2) формально не отличаются, поэтому для рассмотрения логики анализа совокупности реализаций будем использовать выражение (1).

Обозначим через Y(t), E(t) два процесса с двумя временными масштабами T_0, T_1, $T_1 >> T_0$, t – непрерывное время. Пусть исследуемый процесс (временной ряд) U(t) есть сумма этих процессов, т.е. U(t) = Y(t) + E(t). Тогда каждую реализацию $u_i(\xi)$ процесса U можно представить в виде

$$u_k(\xi) = y_k(\xi) + e_k(\xi) = a_k + b_k f(\xi) + e_k(\xi), \qquad (3)$$

где ξ – время от начала соответствующих реализаций, $\xi = [0, T_1]$, $t = (k - 1)T_1 + \xi$, функция $\hat{u}_k(\xi) = a_k + b_k f(\xi)$ является k-той реализацией процесса Y(t), а функция $e_k(\xi)$ есть временной ряд процесса E(t) на $[0,T_1]$.

Из условия $T_1 >> T_0$ следует, что компоненты $y_k(\xi)$, $e_k(\xi)$ будут статистически независимыми. Тогда легко показать, что функцию $f(\xi)$ в (2) можно заменить средней реализацией $\bar{u}(\xi)$ и выражение (3) принимает вид

$$u_k(\xi) = a_k + b_k\bar{u}(\xi) + e_k(\xi). \qquad (4)$$

Параметры a_k, b_k в (4) можно найти, например, методом наименьших квадратов.

Предположим, что дискретность наблюдений сравнима с временным масштабом T_0 и все значения $e_k(\xi)$ статистически независимы между собой по ξ. Тогда для каждой k–той реализации можно найти среднеквадратическое значение σ_k отклонений $e_k(\xi)$, которое можно считать характеристикой *режима* колебаний процесса E(t).

Режим колебаний процесса Y(t) можно определить следующим образом. Найдем амплитуду колебаний средней реализации

$$\bar{A}_y = [\max\{\hat{u}_i(\xi)\} - \min\{\hat{u}_i(\xi)\}]/2, \qquad (5)$$

где $\hat{u}_i(\xi) = a_k + b_k\bar{u}(\xi)$.

Амплитуда текущей реализации $y_i(\xi) = \hat{u}_i(\xi)$ будет равна

$$A_{yk} = b_k \bar{A}_y. \qquad (6)$$

Тогда режим процесса Y(t) можно характеризовать функцией $\bar{u}(\xi)$, а его изменения значениями параметра A_{yk}. Следовательно, для прогноза режима процесса Y(t) достаточно знать будущие значения параметра b_k.

Взаимосвязь процессов разного масштаба можно записать в виде выражения

$$F(a_k, b_k, \sigma_k) = 0, \qquad (7)$$

из которого легко получить уравнение параметризации мелкомасштабного процесса E(t) через параметры крупномасштабного процесса Y(t).

Для прогнозирования a_k, b_k, σ_k можно использовать модели авторегрессий

$$\breve{a}_k = p_m \breve{a}_{k-m} + p_{m-1} \breve{a}_{k-m-1} + \ldots + p_1 \breve{a}_{k-1} + \delta_{ak}, \qquad (8)$$

$$\breve{b}_k = q_m \breve{b}_{k-m} + q_{m-1} \breve{b}_{k-m-1} + \ldots + q_1 \breve{b}_{k-1} + \delta_{bk}, \qquad (9)$$

$$\breve{\sigma}_k = r_m \breve{\sigma}_{k-m} + r_{m-1} \breve{\sigma}_{k-m-1} + \ldots + r_1 \breve{\sigma}_{k-1} + \delta_{\sigma k}, \qquad (10)$$

где m - порядок авторегрессии, p, q, r – неизвестные параметры, δ_{ak}, δ_{bk}, $\delta_{\sigma k}$ - случайные величины.

В модели (8) искомые значения $\breve{a}_k$, $\breve{b}_k$, $\breve{\sigma}_k$ не зависят от значений δ_{ak}, δ_{bk}, $\delta_{\sigma k}$, которые здесь играют ту же роль, что и погрешности измерений при изучении детерминированных процессов. Следовательно, значения δ_{ak}, δ_{bk}, $\delta_{\sigma k}$ могут быть сколь угодно близкими к нулю: чем меньше дисперсия δ, тем точнее можно определить неизвестные параметры. В этом состоит наиболее принципиальное отличие моделей (8) - (10) от авторегрессий в практических приложениях теории случайных процессов [1], где именно значения δ *порождают* авторегрессионный процесс.

Итак, выражения (4), (7), (8), (9), (10) представляют систему уравнений *для описания и прогнозирования* процессов с характерными временными масштабами $T_0 << T_1$. Зависимость $\hat{u}_k(\xi) = a_k + b_k\bar{u}(\xi)$ при фиксированном значении k отражает динамику процесса T_1. Уравнения (8), (9) позволяют прогнозировать эти параметры, если они изменяются во времени с характерным масштабом колебаний $T_2 >> T_1$. С помощью модели (10) можно прогнозировать значение во времени энтропии процесса T_0, а уравнение (7) отражает взаимосвязь процессов T_0, T_1.

Обоснование основной гипотезы. В работе [5] в качестве реализаций рассматривались значения атмосферного давления в срок наблюдения на

Европейской территории Советского Союза, и среднемесячные многолетние значения гидрометеорологических характеристик, полученные на кораблях погоды в Северной Атлантике. Во всех случаях гипотеза о линейной зависимости реализаций не отвергалась.

В работе [6] показана линейная взаимосвязь полей среднемесячной приземной температуры воздуха и полей высоты геопотенциала H_{500} в узлах регулярной сетки на Северном полушарии. Получены оригинальные выводы о величине и характере колебаний климатического масштаба.

В [7] использование линейной связи среднемесячных значений метеорологических переменных позволило разработать эффективный алгоритм обнаружения систематических ошибок в архивных данных.

В статье [8] функция с разделенными переменными использовалась для оценки погрешностей климатических характеристик, полученных по данным судовых метеорологических наблюдений за период 1961-1987 на небольшой акватории Северной Атлантики.

В работе [10] было выдвинуто предположение, что поля радиоактивного загрязнения почвы ^{137}Cs после Чернобыльской аварии на территории Калужской также являются линейно зависимыми между собой. Это позволило получить ряд важных выводов, частности, о том, что степень загрязнения в некоторых населенных пунктах может возрастать во времени. Этот вывод полностью согласуется с более ранними экспериментальными исследованиями. В этой же работе выполнен анализ минимальной за сутки приземной температуры воздуха на территории России за период 1984-1998 годы в предположении, что сезонные колебания минимальной за разные годы линейно взаимосвязаны. Показана возможность использования результатов анализа в долгосрочном прогнозе погоды по методу аналогов.

Два обоснования использования моделей авторегрессии.

<u>Обоснование 1</u>. Известно, что для однозначного описания сложного процесса всегда используется система уравнений для m переменных, $m \geq 1$. Число m называют размерностью системы. Допустим, что $m > 1$, но мы имеем результаты измерений только одной переменной. Вопрос: содержит ли ряд наблюдений этой

переменной информацию о поведении всей системы в целом, т.е. является ли этот ряд достаточным для однозначного описания траектории состояния системы. Если ответ положительный, то это может служить основанием считать, что прогноз каждой переменной можно осуществлять независимо. Если ответ отрицательный, то необходимо искать полный набор переменных.

Положительный ответ основывается на теореме Такенса [4]. Ее смысл заключается в следующем. Пусть состояние системы полностью описывается m переменными: $x_1(t)$, $x_2(t),\ldots, x_m(t)$. С интервалом T производятся измерения какой-либо одной из них, например, $x_1(t)$. Тогда, согласно теореме, вместо последовательности, состоящей из m переменных $x_1(t)$, $x_2(t),\ldots, x_m(t)$, можно рассматривать последовательность $x_1(t+T)$, $x_1(t+2T),\ldots, x_1[t+(m-1)T]$, т.е. в каждый момент времени состояние системы может быть описано m значениями одной переменной, взятыми со сдвигом T. Это и есть основное свойство моделей авторегрессий (8) – (10).

<u>Обоснование 2</u>. В качестве практического, рассмотрим нелинейную систему уравнений Лоренца [4]

$$dx/dt = \sigma(y - x), \qquad dy/dt = rx - y - xz, \qquad dz/dt = xy - bz. \qquad (11)$$

Пусть значения параметров равны: $\sigma = 10$, $r = 28$, $b = 8/3$. Проинтегрируем численно (11) на интервале $t = [0,1]$ с начальными шагами $d_x = d_y = d_z = 0.00002$ при начальных значениях $x_0 = 5$, $y_0 = 5$, $z_0 = 10$. Для определения моделей авторегрессий возьмем численные значения решения с шагом $\Delta = 0.001 >> d_x$. Применяя шаговый метод включения переменных, получим для всех переменных уравнения авторегрессии 3-го порядка с коэффициентом детерминации 100.00% и абсолютной погрешностью не более 0.001 от значений координат x, y, z.

Используя численные значения интеграла (11), легко показать, что с помощью методов регрессионного анализа можно также практически точно восстановить исходную дифференциальную форму. Если погрешность значительная, то необходимо применять другие методы, которые минимизируют невязки по всем переменным, включенным в уравнение.

Основные масштабы колебаний приземной температуры воздуха. Климатическая система является неоднородной, нелинейной, самоорганизующейся системой. Самоорганизация означает, что движения некоторого временного масштаба, представляющие динамический хаос, могут самопроизвольно образовать упорядоченные движения большего временного масштаба. Именно поэтому в атмосфере наблюдаются движения (структуры) разных временных масштабов. Таких масштабов, представляющих наибольший практический интерес, в атмосфере наблюдается пять основных типов [2]:

1) мезомасштабные колебания с характерным периодом несколько часов,

2) суточные колебания с постоянным периодом 24 часа,

3) нерегулярные колебания синоптического масштаба со средним периодом около 5-7 дней,

4) сезонные колебания с постоянным периодом 365 дней,

5) нерегулярные глобальные колебания с периодом несколько лет, несколько десятилетий, несколько столетий и т.д.

Сезонные колебания приземной температуры воздуха можно разделить на две составляющие. Одна составляющая зависит от радиационного баланса при некоторых постоянных условиях в атмосфере. Вторая составляющая колебаний имеет адвективную природу. Она обусловлена циклическим характером сезонных колебаний глобальной циркуляции атмосферы, который и определяет когерентность движения воздушных масс. Формально когерентность проявляется в том, что средние многолетние значения температуры воздуха даже при осреднении за 100 лет нельзя адекватно описать функцией, зависящей только от склонения солнца.

Таким образом, необходимо показать, что по данным метеорологических наблюдений можно найти модели для описания взаимосвязи и прогнозирования колебаний всех указанных процессов.

Пример определения системы эмпирических уравнений.

1. Исходные данные. Пусть исходными данными является ряд 8-ми срочных наблюдений приземной температуры воздуха, полученный на станции Москва ВДНХ за период 1984-2001 годы. Представим их в виде матрицы $u(i, j, k)$, где $i = 1,2,\ldots,8$, $j = 1,2,\ldots,365$, $k = 1,2,\ldots,M$, M – количество лет. Примем в качестве реализаций законченный цикл изменений температуры в течение суток при фиксированном склонении солнца, т.е. значения u(i, j = const, k). Тогда средней реализацией для заданного склонения солнца (j-тый день) будут значения $\bar{u}(i, j)$, полученные после осреднения u(i, j, k) по k.

Определим методом наименьших квадратов параметры a_{jk}, b_{jk} модели (4). Значения a_{jk} являются обычными среднесуточными значениями температуры воздуха за j-тый день k-того года. Параметр b_{jk} показывает, во сколько раз размах суточных изменений в j-тый день k-того года отличается от размаха средних многолетних значений $D\bar{u}_k = \max\{\bar{u}_k(\xi)\} - \min\{\bar{u}_k(\xi)\}$. Невязки (остатки) e_{jk} можно считать оценкой мезомасштабных колебаний, а σ_{jk} – их среднеквадратическим отклонением в текущий день года.

2. Определение радиационной составляющей сезонных колебаний параметров суточных колебаний. Радиационную составляющую можно определить в три этапа. Вначале надо найти для каждого дня средние многолетние значения $\bar{a}(j)$, $\bar{b}(j)$, $\bar{\sigma}(j)$. Затем эти значения аппроксимировать функциями $P_a(j)$, $P_b(j)$, $P_\sigma(j)$. И, наконец, представить значения a_i, b_i, σ_{ei} в виде

$$a(j,k) = c_{ak} + d_{ak}P_a(j) + \varepsilon_{ak}(j), \qquad (12)$$

$$b(j,k) = c_{bk} + d_{bk}P_b(j) + \varepsilon_{bk}(j), \qquad (13)$$

$$\sigma(j,k) = c_{sk} + d_{\sigma k}P_\sigma(j) + \varepsilon_{\sigma k}(j), \qquad (14)$$

где $C = \{c_{ak}, d_{ak}, c_{bk}, d_{bk}, c_{sk}, d_{\sigma k}\}$ – эмпирические параметры.

Функции $\breve{a}(j,k) = c_{ak} + d_{ak}P_a(j)$, $\breve{b}(j,k) = c_{bk} + d_{bk}P_b(j)$, $\breve{\sigma}(j,k) = c_{sk} + d_{\sigma k}P_\sigma(j)$ характеризуют сезонные колебания в течение k-того года, обусловленные радиационным балансом, отклонения $\varepsilon_a(i, k)$, $\varepsilon_b(i, k)$, $\varepsilon_\sigma(i, k)$ характеризуют изменения, обусловленные адвективными факторами, а совокупность параметров С отражает межгодовые колебаний радиационной составляющей.

3. Определение дисперсий адвективных составляющих сезонных колебаний. Согласно выражению (7), среднеквадратические значения S_a, S_b, S_σ адвективных составляющих $\varepsilon_a(j, k)$, $\varepsilon_b(j, k)$, $\varepsilon_\sigma(j, k)$ и совокупность параметров С должны изменяться взаимосвязано. Однако эти составляющие не являются случайными величинами по j. Поэтому для вычисления S_a, S_b, S_σ необходимо выбрать такие значения в каждом k-том временном ряду $\varepsilon_a(i, k)$, $\varepsilon_b(i, k)$, $\varepsilon_\sigma(i, k)$, которые бы, с одной стороны, отражали его изменчивость, а, с другой стороны, были бы независимыми между собой. Для первого приближения в качестве таких значений далее будем использовать значения локальных экстремумов соответствующего временного ряда.

4. Фильтрация погрешностей параметров суточных колебаний. Параметры a_{jk}, b_{jk}, σ_{jk} получены с существенной погрешностью, поскольку длина реализации не велика. Поэтому полученные значения необходимо «очистить» от погрешностей. Для этого целесообразно использовать метод «мягкого» сглаживания [9], сущность которого состоит в следующем. Участки временного ряда, не содержащие локальных экстремумов остаются без изменения. В точках локальных экстремумов значения ряда заменяются на среднее значение, полученное из интерполированного значения по соседним точкам и фактического значения. Отсюда следует, что сглаживаются только большие значения погрешностей, которые в среднем существенно больше первой последовательной разности закономерной составляющей. Для значений параметров после сглаживания сохраним те же обозначения, что и до сглаживания.

5. Общий вид системы эмпирических уравнений. Из предыдущих пунктов видно, что состав эмпирических уравнений для атмосферы «богаче» системы уравнений (4), (7) – (10). Поэтому приведем ее в полном объеме

$$u(i, j, k) = a(j, k) + b(j, k)\bar{u}(i, j) + e(i, j, k), \qquad (15)$$

$$F_1[a(j, k), b(j, k), \sigma(j, k)] = 0, \qquad (16)$$

$$a(j, k) = c_a(k) + d_a(k)P_a(j) + \varepsilon_a(j, k), \qquad (17)$$

$$b(j, k) = c_b(k) + d_b(k)P_b(j) + \varepsilon_b(j, k) \qquad (18)$$

$$\sigma(j, k) = c_\sigma(k) + d_\sigma(k)P_\sigma(j) + \varepsilon_\sigma(j, k), \qquad (19)$$

$$F_2[c_a(k), d_a(k), c_b(k), d_b(k), c_\sigma(k), d_\sigma(k), S_a(k), S_{ab}(k), S_\sigma(k)] = 0, \qquad (20)$$

$$F_3[\varepsilon_a(j), \varepsilon_b(j), \varepsilon_b(j), \varepsilon_\sigma(j), P_a(j), P_b(j), P_\sigma(j), E_a(j), E_b(j), E_\sigma(j)] = 0, \qquad (21)$$

$$\varphi_1(\overset{\sqcup}{a}) = 0, \quad \varphi_2(\overset{\sqcup}{b}) = 0, \quad \varphi_3(\overset{\sqcup}{\sigma}) = 0, \qquad (22)$$

$$\psi_{11}(\overset{\sqcup}{c}_a) = 0, \quad \psi_{12}(\overset{\sqcup}{d}_a) = 0, \quad \psi_{13}(\overset{\sqcup}{S}_a) = 0, \qquad (23)$$

$$\psi_{21}(\overset{\sqcup}{c}_b) = 0, \quad \psi_{22}(\overset{\sqcup}{d}_b) = 0, \quad \psi_{23}(\overset{\sqcup}{S}_b) = 0, \qquad (24)$$

$$\psi_{31}(\overset{\sqcup}{c}_\sigma) = 0, \quad \psi_{22}(\overset{\sqcup}{d}_\sigma) = 0, \quad \psi_{23}(\overset{\sqcup}{S}_\sigma) = 0, \qquad (25)$$

где:

$$\sigma(j, k) = \sqrt{\sum_i e^2(i, j, k)/8}, \quad S_a(k) = \sqrt{\sum_l \varepsilon_a^2(l, k)/n_a}, \qquad (26)$$

$$S_a(k) = \sqrt{\sum_l \varepsilon_b^2(l, k)/n_b}, \qquad S_\sigma(k) = \sqrt{\sum_l \varepsilon_\sigma^2(l, k)/n_\sigma}, \qquad (27)$$

$$E_a(j) = \bar{a}(j) - P_a(j), \quad E_b(j) = \bar{b}(j) - P_b(j), \quad E_\sigma(j) = \bar{\sigma}(j) - P_\sigma(j), \qquad (28)$$

где l под знаком суммы принимают такие значения j, при которых соответствующая величина является локальным экстремумом, n_a, n_b, n_σ – количество локальных экстремумов в рядах $\varepsilon_a(j, k)$, $\varepsilon_a(j, k)$, $\varepsilon_a(j, k)$, соответственно, при фиксированном значении k.

6. Особенности методов оценивания и статистические характеристики моделей (15) – (25), полученные по данным в одном пункте наблюдений. Эти модели представляют систему нелинейных уравнений, для определения которой необходимо использовать специализированный итерационный метод. Ниже рассмотрим методы и результаты определения начального приближения.

6.1. Статистические характеристики модели (15). Формулы для оценки статистических характеристик совокупности реализаций приведены в статье [7]. Однако, оценка статистической значимости для этой модели не является принципиальной. Отсутствие значимости при оценке параметров за какой-нибудь день означает лишь, что амплитуда суточных колебаний в данный день равна нулю. Такие случаи наблюдаются, в основном, в зимний период. В летний период при отсутствии сильной адвекции коэффициент детерминации модели (15) может достигать 98%.

6.2. Уравнение (16) отражает взаимосвязь параметров суточных колебаний за каждый день. Это уравнение можно найти, если использовать данные по большому количеству пунктов наблюдений (по региону или по полушарию). Поэтому здесь нахождение этой зависимости не приводится.

6.3. Уравнения (17) – (19). Здесь основная задача сводится к определению функций $P_a(\tau)$, $P_b(\tau)$, $P_\sigma(\tau)$. Эти функции должны иметь одинаковый вид для всех трех параметров. Их можно найти путем аппроксимации среднегодовых значений логистической функцией, которая в дифференциальной форме имеет вид

$$dy(t)/dt = -k_1/b(y(t) - a)[b - (y(t) - a)], \text{ при } t < t_x, \qquad (29)$$

$$dy(t)/dt = k_2/b(y(t) - a)[b - (y(t) - a)], \text{ при } t \geq t_x, \qquad (30)$$

где t_x – время достижения максимума среднесуточной температуры в течение года, k_1, k_2 - характерное время тепловой релаксации атмосферы при увеличении и уменьшении склонения солнца, соответственно; параметры a, b зависят от физико-географических условий.

Поскольку решение уравнений (27), (28) содержат экспоненту, то для первого приближения его можно заменить полиномом не высокой степени.

6.4. Уравнение (20) взаимосвязи параметров характеризует *межгодовые изменения радиационных и адвективных составляющих* сезонных колебаний параметров суточного хода. Для первого приближения это уравнение можно определить с помощью шагового метода включения переменных [3]. В качестве зависимой переменной $\mathcal{Y}$ были выбраны значения S_σ, т.е. $\mathcal{Y} = S_\sigma$. Остальные переменные: $x_1 = c_a(j)$, $x_2 = d_a(j)$, $x_3 = c_b(j)$, $x_4 = d_b(j)$, $x_5 = c_\sigma(j)$, $x_6 = d_\sigma(j)$, $x_7 = S_b(j)$, $x_8 = S_a(j)$ в этом выражении рассматривались как независимые. В результате получена модель

$$\mathcal{Y} = b_0 + b_1x_8f_1(x_2, x_4, x_5, x_6, x_7) + b_2x_2f_2(x_3, x_4, x_5) + b_3f_3(x_3, x_4), \qquad (31)$$

где функции f_1, f_2 нелинейны по параметрам и по переменным, все коэффициенты статистически значимы при уровне значимости 0.05, а коэффициент детерминации равен 98%.

6.5. Уравнение (21) показывает взаимосвязь *в течение года* радиационной и адвективной составляющей сезонных изменений характеристик суточных колебаний. Для определения этого уравнения в качестве зависимой переменной были приняты значения $\varepsilon_\sigma(j)$. Остальные переменные: $x_1 = P_a(j)$, $x_2 = P_b(j)$, $x_3 = P_\sigma(j)$, $x_4 = E_a(j)$, $x_5 = E_b(j)$, $x_6 = E_\sigma(j)$, $x_7 = \varepsilon_a(j)$, $x_8 = \varepsilon_b(j)$ рассматривались как независимые. Применяя шаговый метод включения переменных, были получены уравнения за все годы. Все коэффициенты полученных моделей оказались статистически значимыми. Наименьшее значение коэффициента детерминации R^2 равно 81%, а наибольшее значение – 97%. Модель для случая R^2 =93% имеет вид

$$\hat{\varepsilon}_\sigma = B_0 + B_1x_1f_1(x_2, x_3, x_4, x_5, x_6, x_7, x_8) + B_2\, x_2f_2(x_3, x_4, x_6, x_7, x_8) +$$
$$B_3x_3x_4f_3(x_5, x_6) + B_4x_6f_4(x_4, x_5) + B_6x_8f_5(x_4, x_6), \qquad (32)$$

где f_1 – f_5 являются сложными функциями от своих переменных.

На рис. 1 показана связь значений ε_{σ} и $\tilde{\varepsilon}_{\sigma}$. Наиболее существенно дисперсия мезомасштабных флуктуаций $\tilde{\varepsilon}_{\sigma}$ зависит от радиационной составляющей $x_1 = P_a$ при взаимодействии с остальными переменными.

6.6. Модели авторегрессий (22) предназначены для прогнозирования параметров суточных колебаний в течение года. Применяя шаговый метод, были получены авторегрессионные модели за все годы для всех трех параметров $\varepsilon_a(j)$, $\varepsilon_b(j)$, $\varepsilon_s(j)$. В среднем порядок авторегрессий для всех параметров равен 3, минимальный порядок оказался равным 2, а максимальный – 5. *Наименьшие* значение коэффициенты детерминации для всех параметров оказались достаточно высокими: 81% - для ε_a, 90% - для ε_b и 91% - для ε_{σ}.

Если известны значения параметров c_a, d_a, c_b, d_b, c_{σ}, d_{σ}, то используя полученные модели авторегрессии и функции $P_a(j)$, $P_b(j)$, $P_{\sigma}(j)$, $\bar{u}(j,k)$, можно с некоторой погрешностью прогнозировать среднесуточные значения, амплитуду и мезомасштабные флуктуации. Такой прогноз было бы крайне полезно использовать в численных моделях краткосрочных и долгосрочных прогнозов погоды. Следует подчеркнуть, что для такого прогноза суточных колебаний <u>на каждый день</u> необходимо знать параметры <u>климатических колебаний</u>, для чего необходимо найти полную систему уравнений (15) – (25).

6.7. Модели авторегрессий (23) – (25) предназначены для прогнозирования межгодовых колебаний параметров сезонных изменений характеристик суточного хода. Межгодовые колебания, как известно [2], испытывают колебания разных временных масштабов. Поэтому для их прогнозирования необходимо найти столько уравнений авторегрессии, сколько масштабов доступно для анализа. Для этого вначале надо, во-первых, осуществить декомпозицию климатических колебаний на разные временные масштабы. Во-вторых, для описания колебаний с характерным масштабом несколько лет дискретность определения параметров уравнений (12) - (14) должна быть значительно меньше года. Решить эти две задачи, используя период наблюдений в 18 лет не представляется возможным.

Если, тем не менее, проигнорировать эти два обстоятельства и применить шаговый метод включения переменных, то получим только одну модель авторегрессии для переменной $\tilde{S}_{\sigma}$ со статистически значимыми параметрами при

5%-ом уровне значимости, однако коэффициент детерминации этой модели не велик – 43%.

Начальное приближение параметров суточных и сезонных колебаний получены в предположении постоянства адвективной составляющей. Для определения последующих приближений необходимо учесть изменение ее во времени и параметры уравнения необходимо оценивать в неявном виде.

Выводы. Возможность конструирования системы эмпирических моделей обеспечивается исключительно благодаря тому, что для этой цели используется функция с разделенными переменными. Для прогнозирования процессов разных временных масштабов можно использовать эмпирические модели авторегрессии. На примере анализа одного временного ряда 8-ми срочных наблюдений приземной температуры воздуха за период 18 лет получены эмпирические модели с высокими статистическими характеристиками. При увеличении объема данных (количества временных рядов) погрешности эмпирических моделей уменьшаются. Это уменьшение обусловлено свойством функции с разделенными переменными и свойством метода мягкого сглаживания.

Из приведенной системы уравнений видна неразрывная связь колебаний суточного и климатического временного масштаба.

Заключение. Метод легко адаптировать для описания глобальных процессов в приземном слое и свободной атмосфере. Систему эмпирических моделей можно использовать в практической работе Гидрометеорологического центра для прогнозирования температуры воздуха и упругости водяного пара в приземном слое атмосферы. Согласно теореме Такенса, не высокий порядок моделей авторегрессий показывает, что для описания сезонных и межгодовых колебаний можно найти систему обыкновенных дифференциальных уравнений невысокого порядка. Возможно, что эмпирическая система уравнений может стать основой для разработки физической теории климата, альтернативной современному подходу, который требует выполнения гигантской работы по конструированию сотен механизмов взаимодействия процессов различной физической природы [11].

Изложенный подход можно использовать для разработки модели эмпирических данных государственного Гидрометеорологического фонда. Можно показать, что это могло бы существенно повысить качество гидрометеорологического обеспечения субъектов народного хозяйства.

Литература

1. Боровиков В.П., Ивченко Г.И. Прогнозирование в системе STATISTICA в среде WINDOWS. Основы теории и интенсивная практика на компьютере:Учеб.пособие.-М.:Финансы и статистика, 2000.-384 с.

2. Груза Г.В., Ранькова Э.Я. Эмпирико-статистический анализ структуры и изменений наблюдаемого климата. //Труды ВНИИГМИ-МЦД, 1980, вып. 68, с.3-22.

3. Дрейпер Н., Смит Г. Прикладной регрессионный анализ. Кн.1, 2. – М.: Финансы и статистика, 1986. – 366 с.

4. Малинецкий Г.Г., Потапов А.Б. Современные проблемы нелинейной динамики. Эдиториал УРСС. Москва. 2000.

5. Пуголовкин В.В., Степаненко С.Р., Шаймарданов М.З. О выявлении систематических ошибок в архивных данных. – Метеорология и гидрология, 1996, № 1, с. 58-67.

6. Степаненко С.Р. Об одном способе эмпирического разложения гидрометеорологических процессов и полей. – Обнинск, ВНИИГМИ-МЦД, 1991, деп. Рукопись № 1093-ГМ91.

7. Степаненко С.Р. Климатические изменения температуры воздуха и высоты поверхности H_{500} Северного полушария. – Метеорология и гидрология, 1995, № 8, с. 14-22.

8. Степаненко С.Р. К вопросу оценки погрешности климатических характеристик. Труды ВНИИГМИ-МЦД. –1996. Вып. 162. – с. 58-65.

9. Степаненко С.Р. Метод мягкого сглаживания эмпирических зависимостей. Труды ВНИИГМИ-МЦД. –1996. Вып. 162. – с. 51-57.

10. Степаненко С.Р., Яхрюшин В.Н. Применение модели нестационарных случайных процессов для мониторинга и прогнозирования сложных систем. //Проблемы безопасности и чрезвычайных ситуаций. № 1, 2006, с. 91-104

11. Физические основы теории климата и его моделирования. Труды Международной научной конференции.–Л.,Гидрометеоиздат, 1977, 271 с.

12. Hasselmann K. Klimamodelle. – Annalen Meteorologi (Neu Folge). – 1980, Nr/15, s. 81-82.