

Р. М. Вильфанд, В. А. Тищенко, В. М. Хан

МУЛЬТИМОДЕЛЬНЫЙ ПОДХОД ПРИ ПРОГНОЗИРОВАНИИ ПРИЗЕМНОЙ ТЕМПЕРАТУРЫ ВОЗДУХА НА СЕЗОН

Введение

В настоящее время ряд мировых прогностических центров ведет активную работу по выпуску ансамблевых гидродинамических прогнозов сроком на сезон и более [9, 11]. Не взирая на непрерывную деятельность по усовершенствованию гидродинамического прогнозирования на долгие сроки, в процессе интегрирования совместных гидродинамических моделей возникают систематические ошибки, которые неизбежно снижают уровень оправдываемости прогнозов. Это связано со следующим:

- 1) не совсем удовлетворительны необходимые для задания начальных условий моделей качество и разрешение по пространству и времени наблюдательной гидрометеорологической сети в глобальном масштабе;
- 2) недостаточное понимание физических процессов, происходящих в окружающей системе, вынуждает описывать их в моделях весьма в упрощенной форме.

Кроме того, в связи с ограниченными возможностями использования вычислительных ресурсов прогностические поля при долгосрочном прогнозировании имеют недостаточное разрешение.

Статистическая постпроцессинговая процедура — один из способов повышения предсказуемости прогнозов, она позволяет провести необходимую корректировку гидродинамических прогнозов к конкретному региону с учетом его физико-географических особенностей.

Основная цель данной работы — разработать схему прогноза приземной температуры воздуха на сезон на территории СНГ с использованием результатов прогнозирования совместных моделей океан—атмосфера из проекта DEMETER¹.

Проект DEMETER предложен, а затем поддержан в рамках 5-й рамочной программы европейского союза по исследованию окружающей среды [10]. Основная цель проекта — усовершенствовать концепцию

¹ Development of a European Multi-model Ensemble System for Seasonal to Interannual Prediction.

мультимодельного ансамблевого прогнозирования. Для реализации поставленной задачи на едином суперкомпьютере был установлен ряд современных глобальных совместных моделей общей циркуляции под руководством Европейского центра среднесрочных прогнозов погоды. Технологическая линия выпуска набора мультимодельных ансамблевых прогнозов сроком до 6 месяцев, сопровождаемая рутинными процедурами архивации и диагностики, была налажена.

Система DEMETER включает в себя семь совместных моделей океан—атмосфера, разработанных в различных исследовательских организациях Европы:

- 1) CERFACS (Европейский вычислительный и исследовательский центр, Франция);
- 2) ECMWF (Международный Европейский центр среднесрочных прогнозов погоды);
- 3) INGV (Национальный институт геофизики и вулканологии, Италия);
- 4) LODYS (Лаборатория динамической океанографии и климатологии, Франция);
- 5) Météo-France (Национальный центр метеорологических исследований, Франция);
- 6) Met Office (Метеорологический офис, Объединенное Королевство);
- 7) MPI (Метеорологический институт Макса Планка, Германия).

Для каждой модели, за исключением MPI, неопределенности начального состояния выражены ансамблем из девяти различных начальных состояний океана.

В последние годы появились публикации в зарубежной периодике, посвященные статистической интерпретации выходных данных одновременно нескольких гидродинамических моделей [9, 11]. Результаты этих работ продемонстрировали перспективность использования мультимодельного подхода в дальнейших постпроцессинговых процедурах.

Существуют разные подходы для создания мультиансамблевого прогноза. Так например, автор работы [9] использовал усовершенствованную регрессионную схему для создания суперансамбля. Суперансамблевый прогноз в работе [11] построен на базе техники декомпозиции сингулярного вектора.

В России в этом направлении активную работу ведут два центра:

- 1) ГГО им. А. И. Воейкова (ГГО) [7, 8];
- 2) Гидрометцентр России [5, 6].

Гидродинамико-статистический прогноз сроком до 1 месяца, разработанный в ГГО, основан на статистической интерпретации результатов интегрирования уравнений глобальной спектральной модели об-

щей циркуляции атмосферы (ОЦА) T30L14. Результаты испытаний продемонстрировали, что качество месячных прогнозов выше, чем инерционных. В 75—80 % случаев знак ожидаемой аномалии прогнозируемого параметра предсказан правильно.

В Гидрометцентре России схема динамико-статистического долгосрочного прогноза разработана на базе модели класса T41L15. Результаты проведенных исследований приведены в работе [4].

Используемые данные и методология

В работе задействованы следующие данные:

1) прогностические поля семи совместных моделей из DEMETER за 1981—2000 гг. Ретропрогнозы рассчитываются по девяти наборам начальных условий на два сезона вперед и имеют отсчет с 1 февраля, 1 мая, 1 августа и 1 ноября каждого года. Разрешение сеточных полей составляет $2,5^\circ$ по широте и долготе;

2) фактические данные приземной температуры воздуха по 70 станциям, расположенным на территории СНГ. Ежедневные данные о температуре воздуха переведены в формат месячных средних;

3) данные сеточных полей NCEP/NCAR реанализа за 1952—2002 гг. с шагом $2,5^\circ$.

Работа по созданию прогностической схемы прогноза температуры состоит из двух этапов.

Первый этап связан с отысканием оптимальной статистической комбинации различных гидродинамических моделей, имеющих достаточно большую выборку ансамблевых долгосрочных прогнозов.

Конкретно, мультиансамблевый прогноз в узловой точке для отдельного метеоэлемента определялся формулой вида

$$S = O + \sum_{i=1}^n A_i (F_i - F_{m_i}), \quad (1)$$

где S — мультиансамблевый прогноз, A_i — весовые коэффициенты моделей, F_i — прогноз отдельно взятой модели, F_{m_i} — осредненный прогноз по отдельно взятой модели.

Весовые коэффициенты в каждой точке регулярной сетки определяются по методу отыскания минимального значения остатка:

$$G = \sum_{i=1}^n (S_i - O_i)^2, \quad (2)$$

где O_i — фактическое значение элемента погоды.

Второй этап работы связан с задействованием полученных прогностических значений суперансамбля в аналого-статистической схеме для расчета приземной температуры воздуха по территории СНГ.

При этом применялось два различных подхода.

1. Статистическая интерпретация, когда прогноз температуры давался на основе предвычисленных на архивном материале показателей взаимосвязи между метеорологическими полями в тропосфере и температурой воздуха у земли.

Статистическая интерпретация включает в себя:

- классификацию полей на архиве;
- расчет корреляций температуры в пунктах и сеточных значений геопотенциала (температуры на изобарических поверхностях) для каждого класса, поиск наилучших районов в пространстве (дальних связей) для прогноза в каждом пункте;
- построение уравнений множественной регрессии для прогноза температуры;
- проверку прогноза на результатах мультиансамблевого гидродинамического прогноза.

2. Использование в качестве предиктора «ведущего потока», когда значения температуры в пунктах прогноза вычислялись по классам (типам) средней (для данного месяца и региона) изолинии прогностических полей геопотенциала и температуры на изобарических поверхностях. Метод не является прогностическим в прямом смысле, это один из способов, позволяющих проводить априорную оценку прогнозируемости среднемесячной температуры воздуха в течение сезона.

Прогноз по «ведущему потоку» включает в себя:

- вычисление среднего ведущего потока (или средней изолинии температуры) для каждого прогностического месяца (по данным DEMETER);
- классификацию полученных результатов;
- сравнение термического режима на исследуемой территории и основных потоков по гидродинамическим прогнозам.

В результате статистических экспериментов по исследованию зависимостей между предиктантом и предикторами в качестве информативных параметров выбраны абсолютный геопотенциал H_{500} , H_{1000} и температура на поверхности 850 гПа (T_{850}), которая хорошо коррелирует с приземной температурой воздуха T_0 .

При составлении прогнозов на достаточно пролонгированные сроки ведущая роль уделяется крупномасштабным составляющим метеорологических полей, поскольку они обладают достаточной инерцией по пространству и времени. По этой причине поля H_{500} , H_{1000} , T_{850} пред-

варительно разлагались по естественным ортогональным составляющим (ЕОС)

$$X_{ih} = \sum_{j=1}^k Y_{ij} V_{jh}, \quad (3)$$

где X_{ih} — среднемесячные значения H_{500} , H_{1000} и T_{850} ; V_{ij} — собственные векторы ковариационной матрицы временного хода указанных величин; Y_{jh} — коэффициенты разложения по времени; i — порядковые номера полей; h — узлы географической сетки.

При всех дальнейших расчетах учтены первые три главных компонента, которые в большинстве случаев описывали более 70 % изменчивости полей.

С целью выявить закономерности влияния комбинаций прогностических полей на изменчивость приземной температуры воздуха в интересующем регионе проводилась автоматическая классификация [1, 2] среднемесячных полей H_{500} для каждого сезона года для территории первого естественного синоптического района (ЕСР), сочетающая черты алгоритма Ланда и метода k -средних. Классификация представляет собой один из алгоритмов «самообучения», когда ни число классов (типов), ни сами типы заранее не известны и определяются только свойствами рассматриваемых полей.

К каждой новой реализации первого предиктора (среднедекадному значению или трем коэффициентам при главных компонентах H_{500}) подбирались группа аналогов, составляющая около 50 % исходной выборки. Мерой аналогичности служили евклидовы расстояния между исходным и всеми остальными векторами-предикторами. Остальные предикторы выстраивались в порядке возрастания евклидова расстояния между исключенным и архивным предикторами. Евклидово расстояние $D_{\alpha, \beta}$ вычислялось по формуле

$$D_{\alpha, \beta} = \left\{ \sum_i P_i [X_i(\alpha) - X_i(\beta)]^2 \right\}^{1/2}, \quad (4)$$

где P_i — весовой коэффициент i -го признака; $X_i(\alpha)$ и $X_i(\beta)$ — i -й признак объектов α и β .

Если признаками являлись среднедекадные значения предикторов, то данное выражение вырождалось в сумму по одному индексу, а весовой коэффициент равнялся нулю. Если в качестве признаков рассматривались коэффициенты разложения по ЕОС, то P_i представляли собой величины, обратные собственным числам λ_i .

Далее на аналоговой выборке решалось уравнение множественной линейной регрессии и проводилась проверка на независимом материале.

Анализ взаимосвязей между предиктором и предиктантами

На первом этапе рассчитывались матрицы коэффициентов корреляции между среднемесячной температурой в 70 пунктах СНГ и полями предикторов за тот же месяц, представленными в виде коэффициентов разложения временного хода эмпирических ортогональных функций в узлах регулярной географической сетки на территории 1-го ЕСР (рис. 1).

В работах [3, 4] показано, что наилучшие показатели оправдываемости прогнозов температуры на месяц отмечаются при определенных типах атмосферной циркуляции. Поэтому предварительно были классифицированы архивные поля геопотенциала на территории 1-го ЕСР по методу, изложенному в предыдущем разделе. Классификация проведена для нормированных отклонений H_{500} (для сезонов зима—осень):

$$dH = (H_i - \bar{H}) / [\sigma(H)]. \quad (5)$$

Оптимальным с точки зрения отношения внутриклассовой дисперсии к общей дисперсии и по объему выборки оказалось использование шести классов полей. Классификация проводилась по 1-му ЕСР. На рис. 1 представлены поля эталонов классов dH_{500} для весны (март—май).

Примем следующие названия для периодов, на которые даются прогнозы:

- период 1 — февраль—июль,
- период 2 — май—октябрь,
- период 3 — август—январь,
- период 4 — ноябрь—апрель.

Корреляционный анализ на архивном материале проведен для каждого из этих периодов с учетом классификации по сезонам.

Исследования, проведенные для случаев, попавших в соответствующие классы, показали наличие следующих особенностей взаимосвязей:

- в течение года районы, где отмечаются максимальные коэффициенты корреляции, перемещаются в пространстве;
- в теплое полугодие максимальные коэффициенты корреляции наблюдаются вблизи исследуемых пунктов, зимой они преобладают в районах, близких к климатическим центрам действия или фронтальным зонам. При этом средние коэффициенты корреляции для всех сезонов меньше, чем для каждого класса внутри сезонов, так как в первом случае не учитываются «дальние связи» для каждой формы циркуляции.

В табл. 1 приведена часть расчетной матрицы максимальных коэффициентов корреляции (между среднемесячной температурой и H_{500}) и соответствующих номеров узлов сетки для всех периодов и классов полей геопотенциала H_{500} . Подобным образом рассчитаны коэффици-

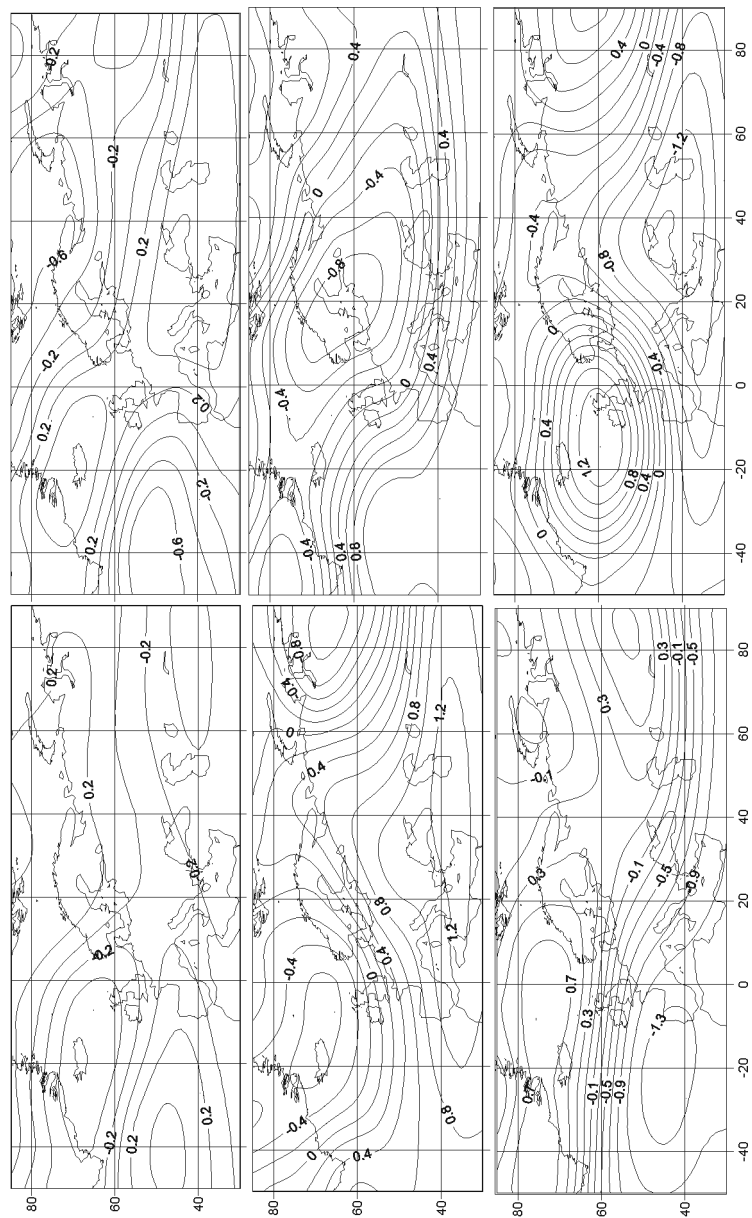


Рис. 1. Классы полей dH_{500} (сезон — весна).

Таблица 1

**Часть расчетной матрицы максимальных коэффициентов корреляций
(между среднемесячной температурой и H_{500})
и соответствующих номеров узлов сетки для нескольких пунктов**

Номер класса	Мурманск	Кандалакша	Кемь	Архангельск
1 - й период				
1	$\frac{0,46}{271}$	$\frac{0,43}{271}$	$\frac{0,49}{125}$	$\frac{0,48}{271}$
2	$\frac{0,52}{72}$	$\frac{0,55}{60}$	$\frac{0,57}{72}$	$\frac{0,57}{72}$
3	$\frac{0,58}{201}$	$\frac{0,40}{201}$	$\frac{0,44}{201}$	$\frac{0,48}{177}$
4	$\frac{0,42}{139}$	$\frac{0,43}{247}$	$\frac{0,46}{259}$	$\frac{0,51}{247}$
5	$\frac{0,40}{6}$	$\frac{0,48}{6}$	$\frac{0,47}{41}$	$\frac{0,49}{42}$
6	$\frac{0,52}{297}$	$\frac{0,51}{297}$	$\frac{0,71}{298}$	$\frac{0,61}{298}$
2 - й период				
1	$\frac{0,49}{123}$	$\frac{0,47}{308}$	$\frac{0,48}{99}$	$\frac{0,49}{99}$
2	$\frac{0,41}{144}$	$\frac{0,46}{23}$	$\frac{0,51}{24}$	$\frac{0,44}{35}$
3	$\frac{0,40}{348}$	$\frac{0,40}{348}$	$\frac{0,49}{348}$	$\frac{0,56}{348}$
4	$\frac{0,49}{331}$	$\frac{0,45}{331}$	$\frac{0,50}{343}$	$\frac{0,51}{343}$
5	$\frac{0,58}{7}$	$\frac{0,55}{7}$	$\frac{0,57}{107}$	$\frac{0,57}{95}$
6	$\frac{0,49}{71}$	$\frac{0,47}{71}$	$\frac{0,55}{111}$	$\frac{0,55}{87}$

Примечание: числитель — коэффициент корреляции, знаменатель — номер узла сетки от крайнего северо-западного узла 1-го ЕСР. Например, узел 271 соответствует 55° с. ш., 60° в. д., узел 72 соответствует 30° с. ш., 25° з. д.

енты корреляции между среднемесячными значениями температуры в 70 пунктах и значениями метеорологических полей в узлах регулярной географической сетки для всех сезонов. На основе полученных взаимосвязей решалось уравнение множественной линейной регрессии и проводилась проверка на независимом материале.

Для метода априорного предсказания качества прогноза «по ведущему потоку» рассчитывались координаты изогипс (изотерм), соответ-

ствующих ведущему потоку для данного месяца. Предварительно проводилась билинейная интерполяция прогностических полей геопотенциала (температуры) с целью минимизировать интервал между узлами регулярной сетки до 0,50. Это позволило проводить поиск нужной изолинии по минимуму разности

$$\Delta x = x_i - \bar{x},$$

где x_i — значение метеорологической величины в i -м узле сетки вдоль меридиана, \bar{x} — значение основной изолинии для данного месяца в исследуемом регионе.

Такой механизм поиска возможен только для очень гладких полей, какими являются среднемесячные поля метеорологических величин в свободной атмосфере.

Получены векторы координат «ведущего потока»:

$$\xi = \lambda\{\varphi[\min(\Delta x)]\},$$

где φ — широта, λ — долгота.

На рис. 2 приведены четыре типа ξ для весенних месяцев. Для расчета использованы только прогностические величины совместных моделей из DEMETER.

Проведено сравнение термического режима на исследуемой территории и основных потоков по гидродинамическим прогнозам.

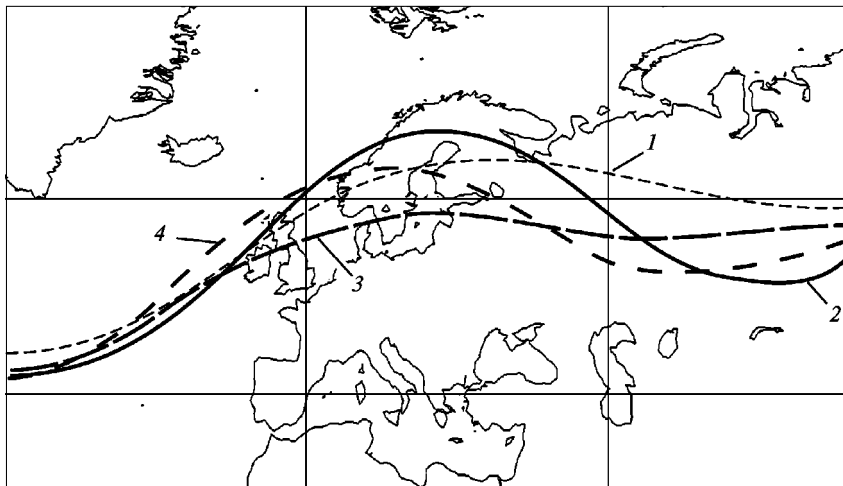


Рис. 2. Типы ξ для весенних месяцев.

Обсуждение полученных результатов

Для оценки качества сезонных прогнозов использованы следующие параметры:

— систематические и абсолютные ошибки, относительные ошибки, вычисляемые как отношение квадрата абсолютных ошибок прогноза к дисперсии фактической температуры относительно климата;

— обеспеченность прогноза H_{500} , T_{850} , T_0 (в процентах) при заданных пределах абсолютных ошибок, оценки ρ и коэффициенты корреляции (r) между прогностическими и фактическими значениями температуры.

Предварительно проверено качество выходной продукции моделей DEMETER и мультиансамблевого прогноза. Вычислены абсолютные ошибки (AM) и обеспеченность (Δ) прогнозов полей геопотенциала и температуры воздуха. Обеспеченность в данном случае представляет собой коэффициент, отражающий число случаев, когда значение величины попадает в заданный интервал (для геопотенциала H_{500} — 8 дам, для температуры T_{850} — 5 °С). Вычислены коэффициенты корреляции между фактическими и прогностическими полями. Выявлено, что практически во всех случаях прогностические значения геопотенциала занижены (при хорошей оценке по коэффициенту корреляции — средний коэффициент корреляции 0,94, минимальный 0,86). В табл. 2 приведены средние абсолютные ошибки и обеспеченность прогнозов по четырем прогностическим периодам.

Статистическая интерпретация отдельных ансамблевых прогнозов, а также мультимодельных прогнозов в целом за прогностический период не позволяет сделать вывод об удовлетворительном предсказании среднемесячных значений температуры воздуха у земли на исследуемой территории. Средние значения коэффициентов корреляции ниже статистически значимых, средние абсолютные ошибки и обеспеченность

Таблица 2

Средние абсолютные ошибки и обеспеченность прогнозов полей геопотенциала H_{500} и температуры воздуха T_{850}

Характеристика	Период 1	Период 2	Период 3	Период 4
AM(H_{500}), дам	–9	–4	–11	–8
AM(T_{850}), °С	0	1	3	1
Δ (H_{500})*	0,79	0,79	0,69	0,72
Δ (T_{850})*	0,91	0,93	0,81	0,85

* Обеспеченность вычислена с поправкой на среднюю абсолютную ошибку.

Таблица 3

Среднее качество прогнозов среднемесячной температуры по периодам					
Характеристика	Предиктор	Период 1	Период 2	Период 3	Период 4
AM(T_0)	H_{500}	-0,46	-0,10	-0,27	-0,63
	T_{850}	-0,44	-0,07	-0,28	-0,62
	$H_{500} + T_{850}$	-0,45	-0,08	-0,27	-0,63
$\Delta(T_0)$	H_{500}	0,73	0,74	0,74	0,72
	T_{850}	0,73	0,74	0,74	0,72
	$H_{500} + T_{850}$	0,76	0,77	0,75	0,73

прогнозов приведены в табл. 3. При этом в отдельные месяцы (в начале прогностического периода) качество прогнозов удовлетворительное. Это связано в первую очередь с выходными данными моделей, близкими к фактическим полям метеорологических величин.

Наилучшие показатели оправдываемости прогнозов температуры в пунктах получились для 1—3-го месяцев прогностических периодов. При этом для 1-го и 2-го периодов максимальные коэффициенты корреляции и обеспеченность получены в 1—2-м месяцах, для 3-го и 4-го периодов — в 3-м месяце. В оставшиеся 3 месяца периодов прогноз температуры хуже, что сказывается на средних оценках.

Абсолютные ошибки, обеспеченность (Δ_1 °С) и коэффициенты корреляции $r(T_0)$ прогнозов среднемесячной температуры на первые 3 месяца прогностических периодов приведены в табл. 4. Абсолютные ошибки могут различаться, при этом $\Delta(T_0)$ и $r(T_0)$ заметно больше, чем для других месяцев.

Таблица 4

Характеристики качества прогнозов среднемесячной температуры за 1—3-й месяцы прогностических периодов					
Характеристика	Месяц	1-й период	2-й период	3-й период	4-й период
AM(T_0)	1-й	-1,4	0,3	-0,4	-1,7
	2-й	0,6	0,9	0,9	0,6
	3-й	-1,4	0,3	-0,4	-1,7
$\Delta(T_0)$	1-й	0,78	0,77	0,65	0,67
	2-й	0,80	0,81	0,62	0,70
	3-й	0,70	0,66	0,82	0,80
$r(T_0)$	1-й	0,24	0,23	0,08	0,15
	2-й	0,19	0,16	0,12	0,09
	3-й	0,12	0,09	0,23	0,27

Примечание: выделены максимальные значения характеристик.

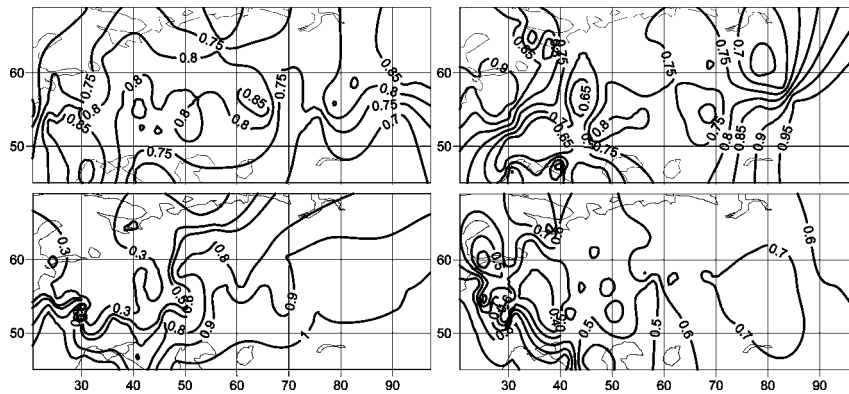
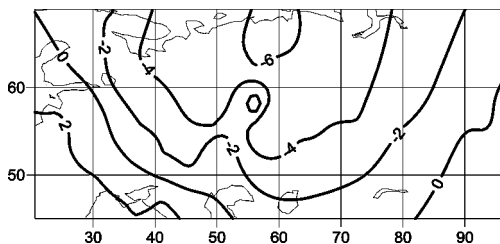


Рис. 5. Среднее поле температуры воздуха при циркуляции, которая соответствует классу 3 «ведущего потока».



определенных типах циркуляции (обычно меридиональном положении изолиний) предсказуемость приземной температуры заметно выше, чем при других типах «ведущего потока». В частности, при циркуляции, соответствующей классу 2 (см. рис. 2), поля среднемесячной температуры мало отличаются от среднего поля, приведенного на рис. 5.

Выводы

На основе проведенных исследований можно сделать следующие выводы.

1. Анализ среднемесячных метеорологических полей в свободной тропосфере (с учетом их типизации) и приземной температуры воздуха показал возможность использования полей H_{500} и T_{850} в качестве предикторов при статистической интерпретации гидродинамических долгосрочных прогнозов.

2. Качество выходной продукции отдельных ансамблевых прогнозов, а также мультиансамблевого прогноза (DEMETER) на 6 месяцев при аналого-статистической интерпретации не позволяет на данном этапе улучшить прогнозируемость термического режима на 1–3-й месяц. В то же время отдельные прогнозы полей среднемесячной температуры с нулевой и 1–2-месячной заблаговременностью соответствуют реальным полям.

СПИСОК ЛИТЕРАТУРЫ

1. Батырева О. В., Васильев А. А., Вильфанд Р. М., Волобуев Н. М. и др. Автоматическая классификация ежедневных полей геопотенциала // Тр. ГМЦ РФ. — 1999. — Вып. 330. — С. 84–89.

2. Боннер Р. Е. Некоторые виды классификации. // Автоматический анализ сложных изображений. — М.: 1969. — С. 205–235.

3. Вильфанд Р. М., Тищенко В. А., Хан В. М. Результаты исследований по прогнозированию сглаженного хода температуры в течение месяца // Тр. ГМЦ РФ. — 2002. — Вып. 337. — С. 16—32.
4. Вильфанд Р. М., Тищенко В. А., Хан В. М. Статистический прогноз температуры воздуха на месяц с использованием выходных данных гидродинамических моделей // Метеорология и гидрология. — 2007. — №3. — С.5—13.
5. Муравьев А. В., Куликова И. А., Круглова Е. Н., Казначеева В. Д. Использование ансамблей в прогнозе метеорологических полей // Метеорология и гидрология. — 2005. — № 3. — С. 5—23.
6. Муравьев А. В., Куликова И. А. Ансамбли прогнозов: методы, проблемы и перспективы // Метеорология и гидрология. — 2005. — № 7. — С. 5—17.
7. Мирвис В. М., Мелешко В. П., Гаврилина В. М., Матюгин В. А. и др. Прогноз метеорологических величин на предстоящий месяц гидродинамико-статистическим методом ГГО. I. Категорические прогнозы — результаты 26-летних серий испытаний // Метеорология и гидрология. — 2006. — № 1. — С. 5—18.
8. Мирвис В. М., Мелешко В. П., Гаврилина В. М., Матюгин В. А. и др. Прогноз метеорологических величин на предстоящий месяц гидродинамико-статистическим методом ГГО. II. Вероятностный прогноз: анализ и интерпретация распределения ансамбля, методика и качество прогнозов // Метеорология и гидрология. — 2006. — № 2. — С. 5—16.
9. Krishnamurti T. N., Kishtawal C. M., Shin D. W., Williford C. E. Multimodel superensemble forecasts for weather and seasonal climate // J. Climate. — 2000b. — 13. — P. 4196—4216.
10. Palmer T. N., Alessandri A., Andersen U., Cantelaube P. et al. Development of a European multi-model ensemble system for seasonal to inter-annual prediction (DEMETER) // Bull/ of the American Meteorol. Soc. — 2004. — 85. — P. 853—872.
11. Yun W. -T., Krishnamurti T. N. Improvement of the superensemble technique for seasonal forecasts // J. Climate. — 2003. — 16. — P. 3834—3840.