

DOI: <https://doi.org/10.37162/2618-9631-2025-2-28-43>

УДК 551.509.33

## Сравнение двух систем внутрисезонного ансамблевого прогнозирования на основе модели ПЛАВ

*И.А. Куликова, Е.Н. Круглова, В.М. Хан,  
Р.Б. Зарипов, В.Г. Мизяк, М.А. Толстых*

*Гидрометеорологический научно-исследовательский центр  
Российской Федерации, г. Москва, Россия  
kulikova@mecom.ru; kruglova@mecom.ru; khan@mecom.ru*

Проводится сравнение двух систем ансамблевого прогнозирования с использованием модели ПЛАВ072L96 – S1 и S2, рассматриваются прогнозы с заблаговременностью до 6 недель. Система S1 использует ансамбль из 61 члена, для генерации ансамбля начальных состояний используется метод бридинга. В новой системе S2 используется 41 член, ансамбль начальных состояний подготавливается с использованием усвоения данных на базе локального ансамблевого фильтра Калмана. Показано, что, несмотря на меньшее количество членов ансамбля, система S2 не только не уступает системе S1, но и обладает некоторыми преимуществами, особенно на длительных интервалах интегрирования (до 46 дней). При этом обе системы характеризуются заниженным разбросом ансамбля и асимметрией распределения прогностических значений, что требует дальнейшей корректировки. Демонстрируется предпочтительность использования системы S2 и ее дальнейшего развития.

*Ключевые слова:* ансамбли прогнозов, метод бридинга, фильтр Калмана, верификация

## Comparison of two ensemble forecasting systems using the SLAV072L96 model

*I.A. Kulikova, E.N. Kruglova, V.M. Khan,  
R.B. Zaripov, V.G. Mizyak, M.A. Tolstykh*

*Hydrometeorological Research Center of Russian Federation, Moscow, Russia  
kulikova@mecom.ru; kruglova@mecom.ru; khan@mecom.ru*

The paper presents a comparison of two ensemble forecasting systems using the SLAV072L96 model: S1 and S2, considering forecasts with a lead time up to 6 weeks. The S1 system uses an ensemble of 61 members, the breeding method is applied to generate the ensemble of initial states. The newer S2 system uses 41 members, the ensemble of initial states is prepared using data assimilation based on the local ensemble Kalman filter. It is shown that despite the smaller number of ensemble members, the S2 system is not only not inferior to the S1 system, but also has some advantages, especially at long integration intervals (up to 46 days). At the same time, both systems are characterized by insufficient ensemble spread and asymmetry in the distribution of forecast values, which requires further adjustment and improvement. It is concluded that the use and further development of the S2 system is preferable.

*Keywords:* forecast ensembles, breeding method, Kalman filter, verification

## Введение

Существенный вклад в развитие и совершенствование метеорологического прогнозирования внесли ансамблевые прогностические системы. Прогнозирование с использованием ансамблей, инициированное в декабре 1992 года Национальным центром прогнозирования окружающей среды США и Европейским центром среднесрочных прогнозов погоды (ECMWF) в качестве дополнения к системе детерминистских среднесрочных прогнозов погоды, получило широкое распространение в мировой метеорологической практике [15, 18]. Важную роль в развитии данного подхода сыграли теоретические и экспериментальные работы Эпштейна [9], Флеминга [10, 11] и Лейта [14], посвященные проблемам неопределенности.

В настоящее время ансамблевый подход широко используется на практике не только для предсказания погоды, но и климатической изменчивости (на временных масштабах от сезонов до десятилетий). Целесообразность использования ансамблей связана не только с недостаточной точностью представления начальных данных, которые используются в моделях, и с несовершенством самих моделей, но и с хаотическим поведением атмосферы и океана. Для описания физических процессов, происходящих в атмосфере и океане, с учетом неопределенности необходимо использовать функции распределения вероятностей характеристик состояния в соответствующем фазовом пространстве. Математическим решением данной задачи может служить уравнение Лиувилля [7, 8]. Однако стохастический подход требует больших вычислительных ресурсов, поэтому на практике изменения функции распределения вероятностей во времени представляются путем использования ансамблевых прогнозов. Современные вычислительные системы позволяют глобальным метеорологическим центрам оперировать ансамблями размерностью порядка нескольких десятков реализаций прогнозов, на основе которых можно строить различные распределения метеорологических параметров, включая характеристики климатических экстремумов.

Основной вопрос, связанный с использованием ансамблей, заключается в выборе процедуры конструирования ансамбля. В Гидрометцентре России разработаны и используются на практике две основные системы ансамблевого прогнозирования, основанные на модели ПЛАВ072L96. В обеих системах схемы параметризаций идентичны. В первом случае при подготовке ансамбля начальных состояний используется метод бридинга [17] (система S1). «Выращивание» начальных возмущений при использовании бридинга проводится 6-часовыми циклами в течение 72 часов. Каждые 6 часов среднее по ансамблю приравнивается к данным анализа, а разброс остается неизменным (если он небольшой) или несколько уменьшается. Разброс членов ансамбля оценивается по высоте изобарической поверхности 500 гПа и полю ветра на ней, в отличие от полной энергии возмущений [1]. После анализа ко всем членам ансамбля добавляются

небольшие случайные возмущения. С подобных полей считается ансамблевый прогноз на следующие 6 часов. Для прогнозов по текущим данным используется ансамбль из одного невозмущенного члена и 60 возмущенных членов ансамбля, для исторических прогнозов число возмущенных членов ансамбля составляет 10 [5].

В [3] представлена новая система ансамблевого среднесрочного прогнозирования, которая базируется на ассимиляции метеорологических данных, включая спутниковые, с использованием локального ансамблевого фильтра Калмана с преобразованием ансамбля (Local Ensemble Transform Kalman Filter, LETKF) (система S2). В процессе ассимиляции метеорологических данных решение задачи ищется с учетом наблюдений в момент анализа, прогноза на интервале между предыдущим и текущим моментом анализа, а также ковариаций ошибок полей первого приближения. Уравнения прогностической модели применяются не только для получения прогноза среднего состояния (первого приближения), но и для прогнозирования ковариаций ошибок анализа, которые учитывают ковариации ошибок самой прогностической модели. В сравнении с методом 4D-VAR, данный подход позволяет явно учитывать ошибки прогностической модели. Кроме того, использование фильтров Калмана не ограничивает временной интервал окна ассимиляции.

В данной работе с использованием статистических методов и общепринятых на практике инструментов верификации прогнозов проводится сравнение двух систем ансамблевого прогнозирования на основе модели ПЛАВ072L96: S1 и S2. Показано, что система S2, несмотря на меньшее количество членов ансамбля, не только не уступает системе S1, но и обладает некоторыми преимуществами, наиболее четко выраженными на длительных интервалах интегрирования (до 46 дней). И в том, и в другом случае наблюдения слишком часто выходят за пределы ансамбля, что указывает на недостаточный разброс ансамбля. Отсюда следует, что одним из потенциалов повышения качества прогнозов на базе модели ПЛАВ является дальнейшее усовершенствование и развитие ансамблевой прогностической системы. С учетом требуемых вычислительных ресурсов система S2 является более предпочтительной по сравнению с S1 и может с успехом заменить ее в оперативной практике.

## 1. Исходные данные и методы исследования

Для сравнения двух систем ансамблевого прогнозирования S1 и S2 использовались детализированные по времени прогнозы, полученные с помощью модели ПЛАВ072L96. Данные охватывали период с 20 июня 2024 года по 14 ноября 2024 года и включали 22 случая прогнозов, выполненных в оперативном режиме. Анализировались среднесуточные числовые поля, представленные на сетке с разрешением  $2.5^{\circ} \times 2.5^{\circ}$  для пяти метеорологических параметров: геопотенциала на уровне 500 гПа (H500), атмосферного давления на уровне моря (MSLP), температуры воздуха на

уровне 850 гПа (T850), температуры у поверхности Земли (TRSF) и интенсивности атмосферных осадков (PREC). Чтобы минимизировать систематические ошибки, связанные с долгосрочным моделированием, прогнозы были представлены в виде аномалий относительно модельного климата за период 1991–2015 годов.

В анализе учитывались не только усредненные аномалии по ансамблю, но и аномалии, рассчитанные для каждого отдельного члена ансамбля. Также оценивались вероятности трех категорий аномалий – терцилей (ниже нормы, норма, выше нормы) для пяти параметров (H500, MSLP, T850, TRSF, PREC) в рамках шести временных интервалов: четырех недельных (1–7, 8–14, 15–21, 22–28 сутки) и двух месячных (1–30 и 16–46 сутки). При построении диаграмм Талаграна, основанных на суточных аномалиях для каждого члена ансамбля, временные интервалы варьировались, включая как более короткие, так и более длинные периоды по сравнению с указанными. Кроме терцилей, для определения границ категорий аномалий использовались 10-й и 90-й процентиля климатического распределения, что позволило выделить экстремальные значения метеорологических параметров.

Эталонными архивами служили данные реанализа ЕЦСПП (ERA5) полей аномалий H500, MSLP, T850, TRSF, PREC за 1991–2015 гг., а также за рассматриваемый прогностический период с 20 июня по 31 декабря 2024 года.

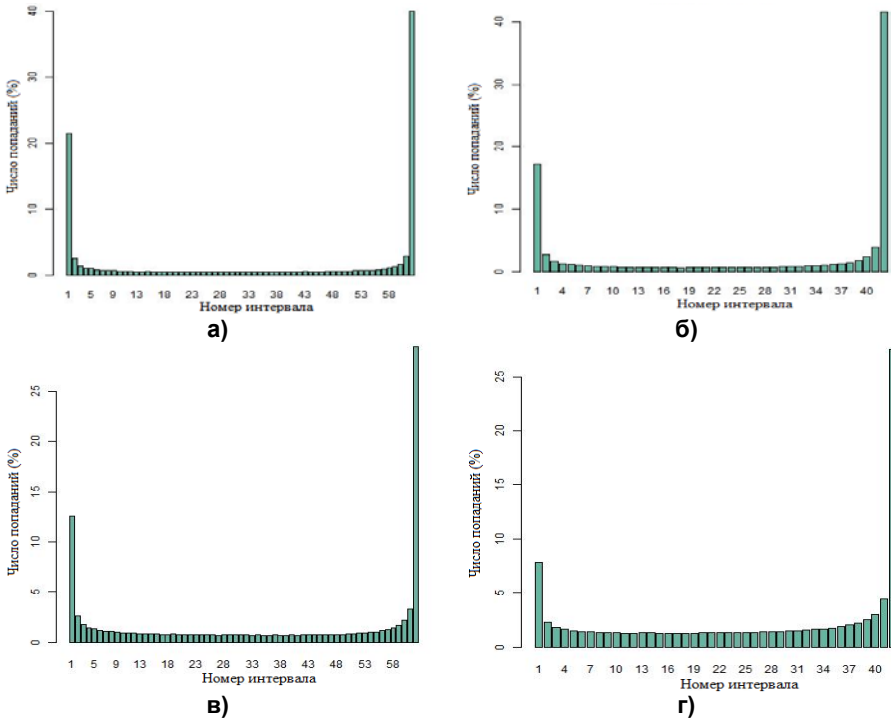
В состав критериев качества прогнозов по величине аномалий включены рекомендованные руководством [16] параметры: средняя квадратическая ошибка (RMSE), пространственный коэффициент корреляции аномалий (ACC) и оценка мастерства прогнозов по отношению к климату (MSSS). Оценка успешности вероятностных прогнозов трех градаций осуществлялась посредством расчетов площади под кривой ROC (AROC), построенной на сопоставлении относительной доли удачных прогнозов (HR) и ложных тревог (FAR) при различных уровнях пороговой вероятности. Формулы для расчетов и более подробную информацию по данным критериям можно найти в [4]. Перечисленные критерии рассчитывались по сеточным данным в пределах трех регионов: внетропические широты Северного полушария (20–90° с. ш.), тропики (20° ю. ш. – 20° с. ш.), Северная Евразия (30–85° с. ш., 0–140° в. д.)

## 2. Анализ качества ансамблевых систем

Одним из наиболее распространенных на практике инструментов оценки качества ансамблевой прогностической системы являются диаграммы Талаграна (или ранговые гистограммы) [12]. Для их построения применяется следующая процедура. Ансамбль из  $N$  прогнозов  $f_1, \dots, f_N$  упорядочивается по возрастанию, и весь диапазон прогнозов разбивается на  $N+1$  интервалов (первый интервал включает значения меньше минимального значения  $f$ ,  $(N+1)$ -й интервал включает значения больше

максимального значения  $f$ ). Затем для каждого интервала определяется относительная частота наблюдений. При построении диаграмм используется метод агрегации или объединения, при котором все точки, расположенные в регионе, например, в Северной Евразии (1311 узлов сетки), объединяются в один массив, на основе которого рассчитываются относительные частоты наблюдений. С учетом 22 начальных дат прогноза, объем выборки для территории Северной Евразии составляет 28842 случая.

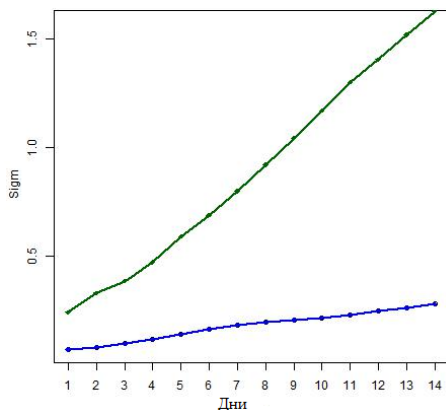
На рис. 1 представлены диаграммы Талаграна для двух ансамблевых систем S1 и S2, построенные для приземной температуры воздуха на недельном и месячном интервалах интегрирования. U-образный вид диаграмм свидетельствует о недостаточности разброса в ансамбле и в том, и в другом случаях, а асимметрия распределения прогностических значений указывают на наличие систематической ошибки в прогнозах.



**Рис. 1.** Диаграммы Талаграна, построенные для двух систем: S1 (левая панель) и S2 (правая панель) модели ПЛАВ072L96 на недельном (верхняя панель) и месячном (нижняя панель) интервалах интегрирования на территории Северной Евразии для приземной температуры воздуха (28842 случая).  
**Fig. 1.** Talagrand diagrams constructed for two schemes S1 (left panel) and S2 (right panel) of the SLAV072L96 model on weekly (upper panel) and monthly (lower panel) integration intervals over the territory of Northern Eurasia for surface air temperature (28842 cases).

На недельном интервале интегрирования для систем S1 (рис. 1а) и S2 (рис. 1б) примерно 61 % и 58 % данных наблюдений (анализа), соответственно, оказываются за пределами ансамбля. При этом небольшое преимущество системы S2 отмечается в области прогнозов отрицательных аномалий температуры воздуха. На месячном интервале интегрирования преимущества системы S2 (рис. 1г) по сравнению с системой S1 (рис. 1в) становятся более заметными, – за пределами ансамбля прогнозов оказываются соответственно 34 % и 43 % данных анализа.

Для дальнейшего сравнения двух схем ансамблевого прогноза рассчитывалась величина разброса прогнозируемых системой величин, обычно определяемая как среднеквадратическое отклонение прогностических значений во всех реализациях ансамбля от среднего по ансамблю прогноза. На рис. 2 приводятся графики изменения среднеквадратического отклонения со временем для двух систем S1 и S2 для приземной температуры воздуха на территории Северной Евразии. Как видно, на всех интервалах интегрирования величина разброса для системы S2 больше, чем для системы S1. С увеличением заблаговременности прогноза различия становятся более существенными. Аналогичные выводы получены для Северного полушария и тропиков, а также для других метеорологических параметров. Поскольку в идеальном случае разброс в ансамбле должен включать фактическое значение метеорологического элемента, можно предположить, что система S2 имеет определенные преимущества, особенно на долговременных интервалах интегрирования, т. е. в контексте долгосрочных прогнозов.



**Рис. 2.** Временной ход среднеквадратического отклонения (sigma) прогностических значений аномалий температуры воздуха во всех реализациях ансамбля от среднего по ансамблю прогноза для двух схем прогноза S1 (синяя кривая) и S2 (зеленая кривая) на территории Северной Евразии.

**Fig. 2.** Time course of the standard deviation (sigma) of the forecast values of air temperature anomalies in all ensemble realizations from the ensemble average forecast for two forecast schemes S1 (blue curve) and S2 (green curve) in the territory of Northern Eurasia.

### 3. Оценки качества прогнозов на внутрисезонных интервалах времени

В таблице приводятся осредненные за весь рассматриваемый период значения коэффициента корреляции аномалий ACC, рассчитанные для внутритропических широт Северного полушария. Согласно таблице, система S2 имеет небольшие преимущества по сравнению с S1, наиболее четко выраженные на длительных интервалах времени. Более заметны различия для атмосферного давления (MSLP), являющегося, вследствие большой изменчивости, одним из наиболее трудных элементов для прогнозирования. В таблице для средних значений ACC представлены доверительные интервалы для тех случаев, когда различия между системами S1 и S2 являются статистически значимыми.

**Таблица.** Средние значения коэффициента корреляции аномалий ACC среднего по ансамблю прогнозов, полученные на базе двух систем ансамблевого прогноза S1 и S2 для Северного полушария (22 начальные даты прогноза: с 20 июня по 14 ноября 2024 г., в скобках представлены доверительные интервалы)

**Table.** Averaged anomaly correlation coefficient (ACC) values of ensemble mean forecast obtained with two ensemble prediction systems S1 and S2 for Northern hemisphere (forecasts for 22 initial dates from 20 of June to 14th of November 2024; confidence intervals are given in the parentheses)

Система/ интервал	Неделя 1	Неделя 2	Неделя 3	Неделя 4	Месяц 1	Месяц 2
<b>TRSF</b>						
S1	0.68	0.36	0.23	0.21	0.45	0.46
S2	0.68	0.39	0.25	0.23	0.47	0.49
<b>PREC</b>						
S1	0.36	0.14	0.10	0.09	0.26	0.19
S2	0.37	0.16	0.12	0.12	0.28	0.21
<b>H500</b>						
S1	0.86	0.40	0.25	0.12	0.49	0.18
S2	0.87	0.45	0.28	0.17	0.55	0.27
<b>MSLP</b>						
S1	0.84	0.39	0.19	0.05	0.44 (0.380-0.500)	0.10 (0.015-0.185)
S2	0.86	0.44	0.27	0.12	0.56 (0.503-0.620)	0.26 (0.186-0.339)
<b>T850</b>						
S1	0.75	0.37	0.22	0.12	0.47	0.19
S2	0.76	0.40	0.24	0.17	0.49	0.23

Задача статистического оценивания распределения среднего решалась путем многократного повторного случайного отбора с заменой из исходной выборки. Данный подход, известный как «бутстреп» [6],

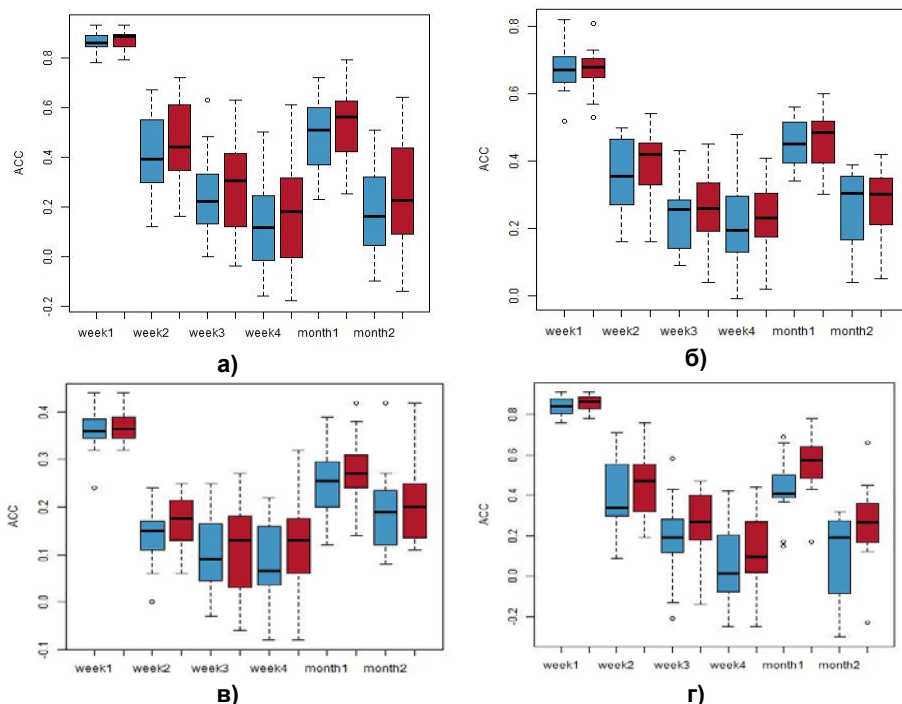
особенно полезен, когда теоретическое распределение статистики неизвестно или сложно вычислить. Повторные бутстреп-выборки были сгенерированы с использованием генератора случайных чисел в количестве равном 500 для каждого временного ряда отдельно. Доверительные интервалы, найденные для выборочных статистик, рассматривались как интервальные оценки соответствующих параметров генеральной совокупности. Средние значения АСС считались значимо отличающимися друг от друга (в пределах 10 % уровня значимости) в том случае, если соответствующие доверительные интервалы не перекрывались.

Более наглядные примеры, демонстрирующие сходство и различия двух систем S1 и S2, могут быть получены с помощью квартильного анализа, позволяющего разделить данные на четыре равные части (квартили), каждая из которых содержит 25 % наблюдений [19]. Квартили используются для построения диаграммы размаха (так называемого «ящика с усами»), которая наглядно отображает распределение данных, медиану и выбросы. Квартильный размах (разница между третьим и первым квартилем) дает представление о разбросе данных вокруг медианы и их изменчивости. Квартильный анализ относится к непараметрическим методам, т. е. не требует предположений о нормальности распределения данных. Среди основных преимуществ квартильного анализа нельзя не отметить также устойчивость к выбросам и возможность сравнения данных различной размерности.

На рис. 3 приводятся диаграммы размаха для коэффициента корреляции аномалий АСС. Полученные графики свидетельствуют о небольшом преимуществе системы S2 по сравнению с S1, как с точки зрения медианы распределения, так и нижнего (верхнего) квартилей. Значения коэффициента корреляции аномалий АСС, равные 0.6, обычно, в частности в ЕЦСПП, используются в качестве порогового значения, ограничивающего интервал практической предсказуемости [13]. Как видно из графиков, для обеих систем ансамблевого прогноза интервал практической предсказуемости, за исключением полей осадков, ограничивается первой прогностической неделей. Для осадков значения коэффициента корреляции аномалий АСС даже для первой недели прогноза не превышают 0.5. На месячном интервале интегрирования для системы S2 примерно в 50 % случаях, за исключением прогнозов осадков, значения коэффициента корреляции аномалий превышают 0.5–0.6.

В ряде исследований [например, 2] показаны преимущества использования вероятностных формулировок прогнозов. На рис. 4 приводятся результаты квартильного анализа агрегированного показателя качества вероятностных прогнозов, а именно площади под кривой ROC (AROC) [4, 16] для градации «выше нормы». Когда доли оправдавшихся прогнозов и ложных тревог равны (отсутствие полезного сигнала),  $AROC = 0.5$  (уровень случайного прогноза). Критерий ROC используется для характеристики разрешающей способности прогностической системы (resolution)

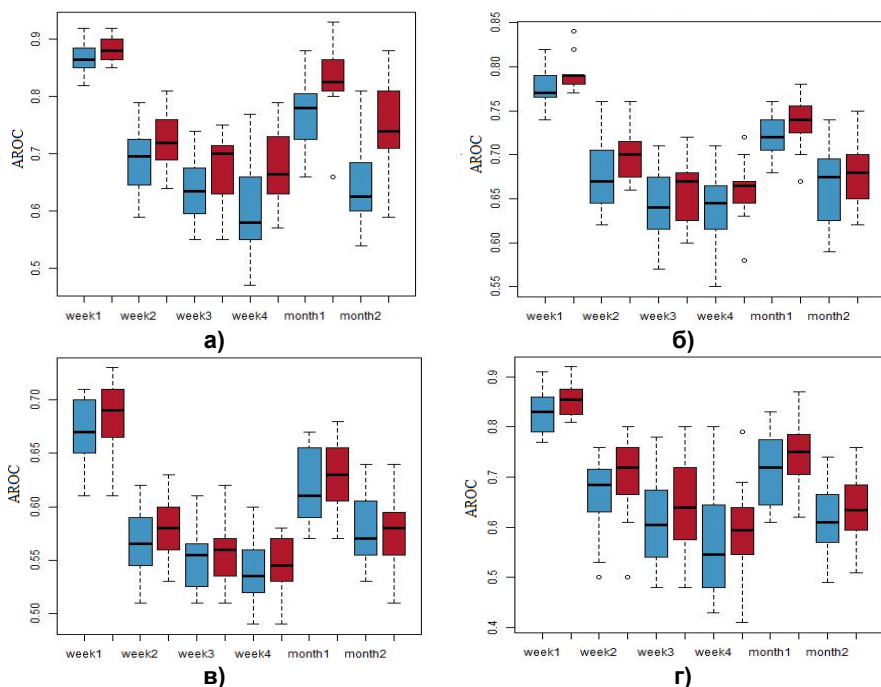
или, иными словами, «умения» распределять наборы событий по подмножествам в зависимости от распределения наблюдаемых частот. Разрешающая способность модели при использовании системы S2 повышается, хотя и очень незначительно. Важно отметить, что улучшения затрагивают месячные интервалы интегрирования, а также прогнозы такого элемента, как осадки.



**Рис. 3.** Квартильный анализ оценок качества прогнозов (коэффициент корреляции аномалий ACC) для полей: H500 (а); приземной температуры воздуха (б); осадков (в); атмосферного давления (г) для Северного полушария (синий цвет соответствует системе S1, красный – S2).

**Fig. 3.** Quartile analysis of quality assessments of forecasts (anomaly correlation coefficient ACC) for the fields: H500 (а); surface air temperature (б); precipitation (в); atmospheric pressure (г) for the Northern Hemisphere (blue color corresponds to the S1 scheme, red – S2).

Полученные выше оценки позволяют сделать вывод, что использование вероятностных формулировок расширяет временные рамки «полезности» прогнозов по сравнению с детерминистским подходом с одной недели до одного месяца, как для системы S2, так и для системы S1. В то же время на временных интервалах в пределах второй, третьей и четвертой недель предсказуемость метеорологических величин ограничивается как фундаментальными свойствами атмосферы, в частности неустойчивостью, так и современным уровнем развития гидродинамического моделирования.



**Рис. 4.** Квартильный анализ оценок качества вероятностных прогнозов (AROC) для полей: H500 (а); приземной температуры воздуха(б); осадков (в); атмосферного давления (г) для градации выше нормы (AROC) для Северного полушария (20 июня – 14 ноября 2024 г.) (синий цвет соответствует схеме S1, красный – S2).

**Fig. 4.** Quartile analysis of the quality estimates of probabilistic forecasts (AROC) for the fields: H500 (a); surface air temperature (б); precipitation (в); atmospheric pressure (г) for the gradation above normal (AROC) for the Northern Hemisphere (June 20 – November 14, 2024) (blue color corresponds to the S1 scheme, red – S2).

#### 4. Экстремальные явления

Ансамблевый подход, позволяя определять не только вероятности метеорологических событий, но и выбранные статистические параметры прогнозируемого распределения вероятностей вектора состояния, является важным инструментом прогнозирования экстремальных метеорологических явлений на длительных интервалах времени. Выбор схемы (S1 или S2), обладающей определенными преимуществами в этом контексте, может помочь снизить риски, связанные с экстремальными метеорологическими явлениями, оказывающие все возрастающее негативное влияние на различные отрасли экономики и социальные сферы жизни общества.

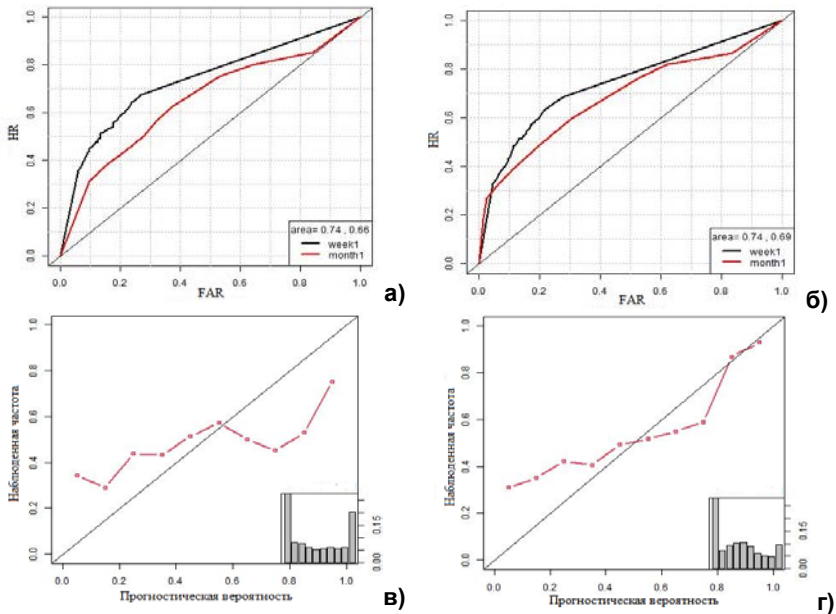
Оценка успешности вероятностных прогнозов экстремальных метеорологических явлений требует комплексного подхода, учитывающего как статистические метрики, так и практическую значимость

прогнозов. Весьма полезной в этом контексте может оказаться кривая ROC (Receiver Operating Characteristic curve), оценивающая разрешающую способность модели различать экстремальные события и не-события. Площадь под кривой (AROC) используется как мера качества прогноза. Для получения более полного представления о качестве прогнозов строятся диаграммы надежности (reliability) и гистограммы частот, характеризующие соответственно надежность, показывающую насколько хорошо прогнозируемые вероятности соответствуют наблюдаемым частотам событий, и избирательность (sharpness) системы прогнозов [16].

С помощью данных критериев решалась задача оценки успешности вероятностных прогнозов экстремальных значений температуры воздуха (больше 90 и меньше 10-го перцентилей), полученных на основе прогностических ансамблей S1 и S2 для каждого вышеуказанного региона и временного интервала в отдельности. На рис. 5а и рис. 5б приводятся кривые ROC, построенные для прогнозов экстремальных значений температуры воздуха (выше 90-го перцентиля) по результатам интегрирования модели ПЛАВ072L96 на недельном и месячном интервалах для систем S1 и S2 на территории Северной Евразии (начальная дата прогноза 15 августа 2024 г.). Как видно, на недельном интервале интегрирования различия практически отсутствуют, на месячном интервале небольшое преимущество имеет система S2. Диаграммы надежности подтверждают указанные различия, предсказанные системой S2 вероятности (рис. 5г) располагаются ближе к диагональной линии, чем в случае системы S1 (рис. 5в).

Избирательность (sharpness) относится к способности модели предсказывать экстремальные или редкие события с высокой точностью и уверенностью. Эта метрика характеризует, насколько хорошо выражены прогнозы модели, т. е. насколько модель склонна выдавать предсказания с высокой уверенностью (например, с вероятностью близкой к 0 % или 100 %), а не оставаться в области неопределенности (50 % вероятности). В этом контексте нельзя не отметить небольшое преимущество системы S1 по сравнению с S2 (рис. 5в и рис. 5г соответственно). По мнению экспертов [16], кривая ROC более подходит для оценки качества долгосрочных прогнозов, чем другие критерии. Важным преимуществом кривой ROC является ее способность оценивать качество модели при всех возможных порогах классификации прогнозов, что позволяет анализировать её производительность независимо от выбора конкретного порога. Кривые надежности (включая гистограммы частот для оценки избирательности, sharpness) рекомендуется строить по большим выборкам вероятностных прогнозов с агрегированием по большим по площади территориям.

Для оценки практической полезности систем S1 и S2 использовался метод анализа затрат и потерь (cost-loss analysis), который позволяет определить, насколько выгодно использовать прогнозы для минимизации убытков, связанных с экстремальными метеорологическими явлениями.



**Рис. 5.** Кривые ROC (а и б), диаграммы надежности (в и г) и гистограммы частот (в и г справа в нижнем углу), построенные для прогнозов экстремальных значений приземной температуры воздуха (выше 90-го перцентиля) на недельном и месячном интервалах интегрирования для систем S1 (левая панель) и S2 (правая панель) модели ПЛАВ072L96 на территории Северной Евразии (начальная дата прогноза 15 августа 2024 г.). HR – доля попаданий, FAR – доля ложных тревог. Диаграммы надежности и гистограммы частот построены для месячного интервала интегрирования.

**Fig. 5.** ROC curves (a and б), reliability diagrams (в and г) and frequency histograms (c and d in the lower right corner) constructed for forecasts of extreme surface air temperature values (above the 90th percentile) on weekly and monthly integration intervals for the S1 (left panel) and S2 (right panel) systems of the SLAV072L96 model in Northern Eurasia (the initial forecast date is August 15, 2024). HR – Hit Rate, FAR – False Alarm Ratio. Reliability diagrams and frequency histograms are constructed for a monthly integration interval.

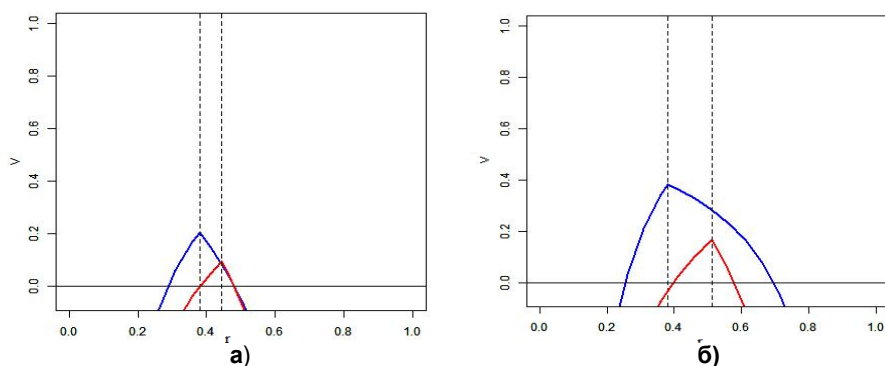
Потенциальная экономическая эффективность прогнозов оценивалась по формуле [20]:

$$V = \frac{\min[o, r] - (HR + FAR)r - m}{\min[o, r] - or},$$

где  $V$  – потенциальная экономическая эффективность прогностической системы, с одной стороны, зависящая от трех основных параметров: HR, FAR и  $m$ , характеризующих, соответственно, долю «попаданий», долю ложных «тревог» и пропуск цели (т. е. качество прогнозов). С другой стороны, величина  $V$  определяется климатической повторяемостью явления  $o$  и соотношением между затратами и потерями потребителя  $r$ .

Положительные значения  $V$  означают, что использование прогнозов позволяет снизить ожидаемые убытки. Прогнозы экономически выгодны. Если  $V < 0$ , то прогностическая система обходится пользователю дороже, чем климатическая информация. Наряду с параметром  $V$  для оценки экономической полезности прогнозов использовался так называемый показатель Клейтона (Clayton skill score (CSS) [12]), характеризующий диапазон соотношений затрат/потерь, при которых пользователи могут получить выгоду от прогнозов.

На рис. 6 приводятся графики зависимости экономической эффективности вероятностных прогнозов экстремальных значений температуры воздуха (выше 90-го перцентиля)  $V$  от соотношения между затратами и потерями потребителя  $r$  для двух ансамблевых систем S1 и S2 модели ПЛАВ072L96 на недельном и месячном интервалах интегрирования для территории Северной Евразии.



**Рис. 6.** «Экономическая эффективность» вероятностных прогнозов экстремальных значений температуры воздуха (выше 90-го перцентиля) для двух ансамблевых систем модели ПЛАВ072L96: S1 (а) и S2 (б) на недельном (синяя кривая) и месячном (красная кривая) интервалах интегрирования на территории Северной Евразии (начальная дата прогноза 15 августа 2024 г.).  
**Fig. 6.** “Economic efficiency” of probabilistic forecasts of extreme air temperature values (above the 90th percentile) for two ensemble systems of the SLAV072L96 model: S1 (a) and S2 (b) on weekly (blue curve) and monthly (red curve) integration intervals in the territory of Northern Eurasia (initial forecast date is August 15, 2024).

Как и следовало ожидать, наибольшие положительные значения эффективности  $V$  наблюдаются на недельном интервале интегрирования. На месячном интервале интегрирования с нулевой заблаговременностью положительный эффект хотя и уменьшается, но сохраняется. С точки зрения параметра  $V$  система S2 имеет определенные преимущества как на недельном, так и на месячных интервалах интегрирования. Параметр CSS служит дополнительным подтверждением, – диапазон значений  $r$  для системы S2 по сравнению с системой S1 увеличивается (круг пользователей расширяется). К преимуществам вероятностного прогнозирования можно отнести и возможность максимизировать экономическую ценность

для любого конкретного пользователя (соотношения затрат/потерь) за счет выбора конкретного порогового значения вероятности.

Таким образом, существующие небольшие различия в оценках качества прогнозов ансамблевых прогностических систем свидетельствуют в пользу системы S2 по сравнению с системой S1. Существенным преимуществом системы S2 являются меньшие вычислительные затраты на подготовку ансамблевого прогноза. При расчете прогнозов на 46 дней требуется около 40 % вычислительных ресурсов, необходимых для расчета с использованием системы S1. Экономия ресурсов достигается не только за счет сокращения числа членов ансамбля, но и благодаря использованию начальных данных, подготовленных для ансамблевого среднесрочного прогноза погоды (с использованием модели ПЛАВ072L96).

### Заключение

Существенный вклад в развитие и совершенствование метеорологического прогнозирования вносят ансамблевые прогностические системы, позволяющие учесть неопределенности, связанные, с одной стороны, с недостаточной точностью представления исходных метеорологических и океанографических полей и несовершенством моделей, с другой, – с хаотическим поведением самой атмосферы. В Гидрометцентре России разработаны две системы ансамблевого прогноза S1 и S2 модели ПЛАВ072L96. Обе системы идентичны, за исключением способа подготовки ансамбля начальных состояний (и самого размера ансамбля – 61 и 41 член соответственно). В первом случае при подготовке ансамбля прогнозов используется метод бридинга, во втором – алгоритм усвоения данных на основе локального ансамблевого фильтра Калмана.

Результаты верификации двух систем ансамблевого прогноза с использованием диаграмм Талаграна и среднеквадратического отклонения прогностических значений во всех реализациях ансамбля от среднего по ансамблю прогноза показывают, что система S2 обеспечивает соизмеримое или несколько лучшее качество прогнозов, чем система S1. К сожалению, обе системы характеризуются недостаточностью разброса ансамбля, а асимметрия распределения прогностических значений указывает на наличие систематической ошибки в прогнозах. Небольшие преимущества (статистически значимые в отдельных случаях) системы S2 по сравнению с системой S1 отмечаются также при оценке качества ансамблевых прогнозов, в том числе прогнозов экстремальных величин.

Можно сделать вывод о том, что использование системы S2 является предпочтительным по сравнению с S1, поскольку может обеспечить не только небольшое улучшение качества прогнозов, но и существенно сэкономить вычислительные ресурсы (за счет уменьшения числа членов ансамбля). Ввиду этого с начала 2025 года оперативные долгосрочные прогнозы Гидрометцентра России, подготавливаемые с использованием модели ПЛАВ072L96, используют систему S2. Метод бридинга

используется при подготовке модельного климата (т. е. при счете ансамблевых прогнозов от данных 1991–2015 годов).

Работа выполнена при поддержке Российского научного фонда проект РНФ № 22-17-00247-П.

### Список литературы

1. Астахова Е.Д. Ансамблевый среднесрочный прогноз погоды: реализация технологии на современной компьютерной базе // Труды Гидрометцентра России. 2011. Вып. 346. С. 38-52.
2. Куликова И.А., Вильфанд П.М., Хан В.М., Круглова Е.Н., Тищенко В.А., Емелина С.В., Каверина Е.С., Набокова Е.В., Субботин А.В., Сумерова К.А., Толстых М.А. Климатические прогнозы. Часть II. Вероятностные подходы // Метеорология и гидрология. 2024. № 8. С. 5-19.
3. Мизяк В.Г., Алипова К.А., Толстых М.А., Розутов В.С. Результаты оперативных испытаний глобальной системы ансамблевого среднесрочного прогноза погоды на основе модели ПЛАВ // Результаты испытания новых и усовершенствованных технологий, моделей и методов гидрометеорологических прогнозов. 2023. Информационный сборник № 50. С. 44-65.
4. Муравьев А.В., Вильфанд П.М. О стандартизации оценок качества прогнозов на средние и долгие сроки // Метеорология и гидрология. 2000. № 12. С. 24-34.
5. Толстых М.А., Фадеев Р.Ю., Шашкин В.В., Зарипов Р.Б. и др. Модель долгосрочного метеорологического прогноза ПЛАВ072L96 // Метеорология и гидрология. 2024. № 7. С. 25-39.
6. Эфрон Б. Нетрадиционные методы многомерного статистического анализа. М.: Финансы и статистика, 1988. 263 с.
7. Ehrendorfer M. The Liouville equation and its potential usefulness for the prediction of forecast skill. Part I: Theory // Mon. Wea. Rev. 1994. Vol. 122. P. 703-713.
8. Ehrendorfer M. The Liouville equation and its potential usefulness for the prediction of forecast skill. Part II: Applications // Mon. Wea. Rev. 1994. Vol. 122. P. 714-728.
9. Epstein E.S. Stochastic dynamic predictions // Tellus. 1969. Vol. 21. P. 739-759.
10. Fleming R.J. On stochastic dynamic prediction. Part I: The energetics of uncertainty and the question of closure // Mon. Wea. Rev. 1971. Vol. 99. P. 851-872.
11. Fleming R.J. On stochastic dynamic prediction. Part II: Predictability and utility // Mon. Wea. Rev. 1971. Vol. 99. P. 927-938.
12. Forecast Verification in Atmospheric Science. A Practitioner's Guide: Second Ed., Eds. I. Jolliffe and D. Stephenson. John Wiley & Sons Ltd, 2012. 274 p.
13. Hollingsworth A., Arpe K., Tiedtke M. et al. The Performance of a Medium Range Forecast Model in Winter-Impact Of Physical Parameterization // Mon. Wea. Rev. 1980. Vol. 108. P. 1736-1773.
14. Leith C.E. Theoretical skill of Monte Carlo forecasts // Mon. Wea. Rev. 1974. Vol. 102. P. 409-418.
15. Palmer T.N., Molteni F., Mureau R., Buizza R., Chapelet P., Tribbia J. Ensemble prediction // Proceedings of the ECMWF Seminar on Validation of Models over Europe. 1993. Vol. 1. P. 21-66.
16. Standardized Verification System (SVS) for Long-Range Forecasts (LRF). New Attachment II-9 to the Manual on the GDPS (WMO-No. 485), 2002. Vol. I.
17. Toth Z., Kalnay E. Ensemble forecasting at NCEP and the breeding method // Mon. Wea. Rev. 1997. Vol. 125. P. 3297-3319.
18. Tracton M.S., Kalnay E. Operational ensemble prediction at the National Meteorological Center: Practical aspects // Wea. Forecasting. 1993. Vol. 8. P. 379-398.
19. Wilks D.S. Statistical methods in the atmospheric sciences: 3d ed. London, Academic Press, 2011. 676 p.
20. Zhu Y., Toth Z., Wobus R., Richardson D., Mylne K. The economic value of ensemble-based weather forecasts // BAMS. 2002. Vol. 83. P. 73-82.

## References

1. Astakhova E.D. Ensemble medium-range weather prediction: the technology implementation on a modern computer base. *Trudy Gidromettsentra Rossii [Proceedings of the Hydrometcentre of Russia]*, 2011, vol. 346, pp. 38-52 [in Russ.].
2. Kulikova I.A., Vilfand R.M., Khan V.M., Kруглова E.N., Tishchenko V.A., Emelina S.V., Kaverina E.S., Nabokova E.V., Subbotin A.V., Sumerova K.A., Tolstykh M.A. Climate Forecasts. Part II: Probabilistic Approaches. *Russ. Meteorol. Hydrol.*, 2024, vol. 49, pp. 659-668. DOI: 10.3103/S1068373924080016.
3. Mizyak V.G., Alipova K.A., Tolstykh M.A., Rogutov V.S. Rezul'taty operativnykh ispytaniy global'noy sistemy ansamblevogo srednesrochnogo prognoza pogody na osnove modeli PLAV. *Rezul'taty ispytaniya novykh i usovershenstvovannykh tekhnologiy, modeley i metodov gidrometeorologicheskikh prognozov*. 2023, Informacionnyy sbornik № 50, pp. 44-65 [in Russ.].
4. Murav'ev A.V., Vilfand R.M. Standardization of Skill Scores for Medium- and Long-Term Weather Forecasts. *Meteorologiya i Gidrologiya [Russ. Meteorol. Hydrol.]*, 2000, no. 12, pp. 24-34 [in Russ.].
5. Tolstykh M.A., Fadeev R.Y., Shashkin V.V., Zaripov R.B., Travova S.V., Goyman G.S., Alipova K.A., Mizyak V.G., Tishchenko V.A., Kруглова E.N. The SLAV072L96 Model for Long-range Meteorological Forecasts. *Russ. Meteorol. Hydrol.*, 2024, vol. 49, no. 7, pp. 576-86. DOI: 10.3103/S1068373924070021.
6. Efron B. *Netradicionnye metody mnogomernogo statisticheskogo analiza*. M.: Finansy i statistika publ., 1988, 263 p. [in Russ.].
7. Ehrendorfer M. The Liouville equation and its potential usefulness for the prediction of forecast skill. Part I: Theory. *Mon. Wea. Rev.*, 1994, vol. 122, pp. 703-713.
8. Ehrendorfer M. The Liouville equation and its potential usefulness for the prediction of forecast skill. Part II: Applications. *Mon. Wea. Rev.*, 1994, vol. 122, pp. 714-728.
9. Epstein E.S. Stochastic dynamic predictions. *Tellus*, 1969, vol. 21, pp. 739-759.
10. Fleming R.J. On stochastic dynamic prediction. Part I: The energetics of uncertainty and the question of closure. *Mon. Wea. Rev.*, 1971, vol. 99, pp. 851-872.
11. Fleming R.J. On stochastic dynamic prediction. Part II: Predictability and utility. *Mon. Wea. Rev.*, 1971, vol. 99, pp. 927-938.
12. Forecast Verification in Atmospheric Science. A Practitioner's Guide: Second Ed., Eds. I. Jolliffe and D. Stephenson. John Wiley & Sons Ltd, 2012, 274 p.
13. Hollingsworth A., Arpe K., Tiedtke M. et al. The Performance of a Medium Range Forecast Model in Winter-Impact Of Physical Parameterization. *Mon. Wea. Rev.*, 1980, vol. 108, pp. 1736-1773.
14. Leith C.E. Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, 1974, vol. 102, pp. 409-418.
15. Palmer T.N., Molteni F., Mureau R., Buizza R., Chapelet P., Tribbia J. Ensemble prediction. Proceedings of the ECMWF Seminar on Validation of Models over Europe, 1993, vol. 1, pp. 21-66.
16. Standardized Verification System (SVS) for Long-Range Forecasts (LRF). New Attachment II-9 to the Manual on the GDPS. *WMO-No. 485*, 2002, vol. I.
17. Toth Z., Kalnay E. Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, 1997, vol. 125, pp. 3297-3319.
18. Tracton M.S., Kalnay E. Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Wea. Forecasting*, 1993, vol. 8, pp. 379-398.
19. Wilks D.S. *Statistical methods in the atmospheric sciences*: 3d ed. London, Academic Press publ., 2011, 676 p.
20. Zhu Y., Toth Z., Wobus R., Richardson D., Mylne K. The economic value of ensemble-based weather forecasts. *BAMS*, 2002, vol. 83, pp. 73-82.

Поступила 26.02.2025; одобрена после рецензирования 04.06.2025;  
принята в печать 18.06.2025.

Submitted 26.02.2025; approved after reviewing 04.06.2025;  
accepted for publication 18.06.2025.