

А. Ю. Бундель

**ОЦЕНКА КАЧЕСТВА СРЕДНЕСРОЧНЫХ ПРОГНОЗОВ
КОЛИЧЕСТВА ОСАДКОВ
ПО ДАННЫМ АНСАМБЛЕВОГО ИНТЕГРИРОВАНИЯ
ПРИ ПОМОЩИ СПЕКТРАЛЬНОЙ МОДЕЛИ ГИДРОМЕТЦЕНТРА РОССИИ**

Введение

В лаборатории гидродинамических среднесрочных прогнозов Гидрометцентра России разработана методика ансамблевого прогнозирования при помощи бридинг-метода создания начальных полей возмущений [2]. Ансамблевый подход потребовал ввести вероятностные оценки в систему верификации. В настоящее время Всемирная Метеорологическая Организация (ВМО) рекомендует следующие стандартные вероятностные характеристики качества прогностической системы [8]:

- ROC-диаграмму и площадь под ней;
- ранговую гистограмму (называемую также диаграммой Талаграна);
- оценку Брайера;
- диаграммы надежности и разрешения;
- неопределенность, характеризующую синоптическую ситуацию.

Очень важна для практического использования прогнозов — задача принятия решения. Она зависит от потребностей конкретного пользователя. В настоящее время хорошо разработана оценка потенциальной экономической эффективности, учитывающая затраты на меры предохранения от неблагоприятного погодного явления и потери в результате этого явления при отсутствии предохранительных мер. Далее обсуждены эти оценки в применении к среднесрочным ансамблевым прогнозам осадков при помощи спектральной модели Гидрометцентра России.

Технологические аспекты

В настоящее время используется версия модели T85L31 с 7—11 членами ансамбля, однако настройки разрешения модели и числа членов ансамбля легко можно изменить.

Расчеты ведутся от начальной даты для 28 заблаговременностей (до 10 сут): 0, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66, 72, 78, 84, 96, 108, 120, 132, 144, 156, 168, 180, 192, 204, 216, 228, 240 ч.

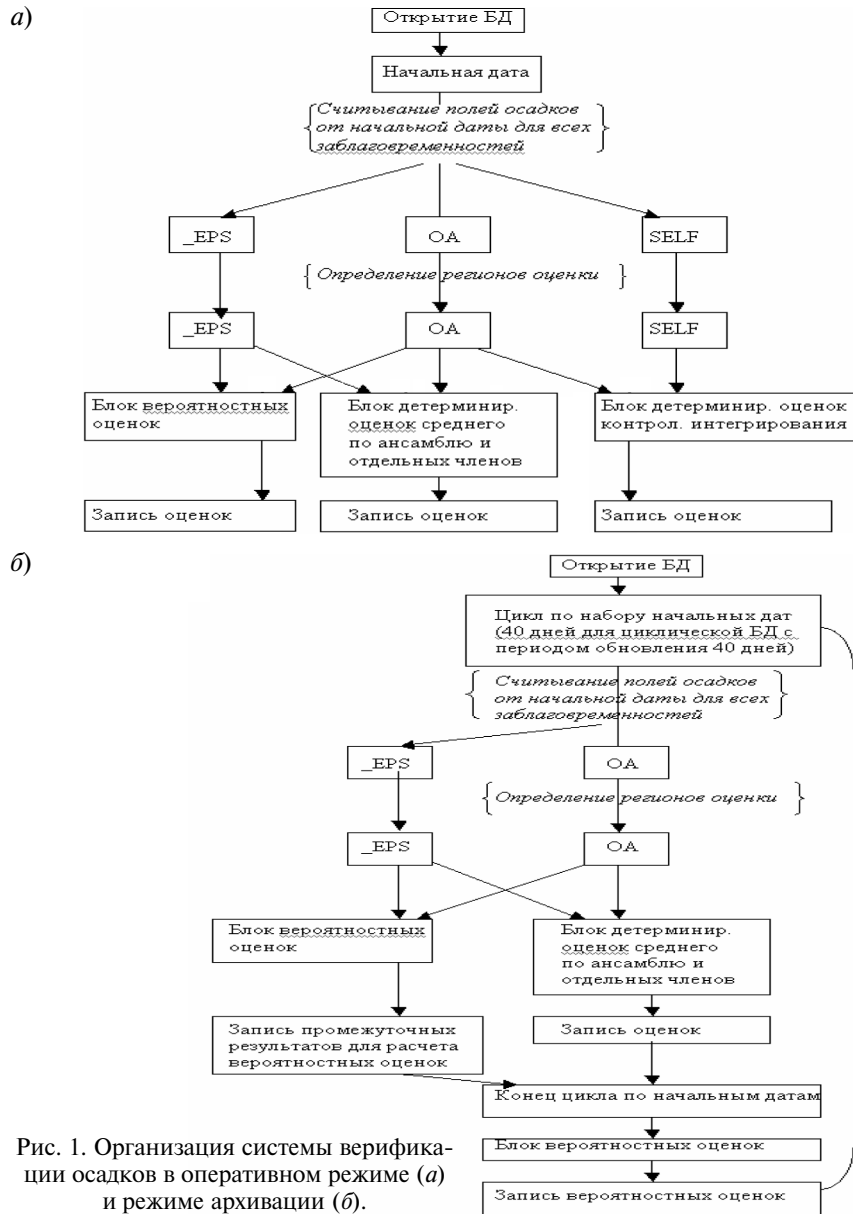


Рис. 1. Организация системы верификации осадков в оперативном режиме (а) и режиме архивации (б).

В качестве данных наблюдений для верификации осадков используются сеточные поля осадков, получаемые в оперативном режиме в схеме восстановления уточненных значений шестичасовых сумм осадков по данным сетевых метеорологических измерений [1]. Далее они названы объективным анализом (ОА) осадков.

Система реализована в двух режимах: 1) оперативном и 2) архивации данных (рис. 1).

Архивация данных необходима для получения статистически значимых вероятностных оценок. Поэтому потребовалось создать специальную циклическую базу данных (БД) с периодом 40 сут для накопления результатов ансамблевого прогноза. ОА осадков сохраняется в отделе информационного обеспечения Гидрометцентра России в циклической базе данных с периодом 60 сут. Архивация данных позволит также оценить качество моделирования в конкретных синоптических ситуациях, например, при интенсивных циклонах.

В настоящее время проводится накопление результатов для расчета статистически обеспеченных оценок. Приведенные далее оценки рассчитаны по данным отдельного случая.

Оценки

Метод ROC¹ представляет собой расширение методики таблиц сопряженности для вероятностных прогнозов. При вероятностном прогнозе вместо категорического утверждения о том, будет явление или нет, вычисляется вероятность явления.

Имеются методы учета неопределенности при оценке прогностической вероятности по конечному числу членов ансамбля (например, байесовский подход), однако в данной работе применен стандартный способ расчета прогностической вероятности как отношения членов ансамбля, предсказывающих явление, к общему числу членов ансамбля.

Весь диапазон вероятностей от 0 до 100 % можно разбить на интервалы пороговыми значениями вероятностей. В нашем случае пороговые значения выбирались в зависимости от числа членов ансамбля, т. е., использовалось $N + 1$ категорий от $0/N$ до N/N для ансамбля из N членов. Использование пороговых вероятностей как бы переводит вероятностный прогноз в детерминированный:

— прогноз явления с вероятностью, превышающей пороговое значение, рассматривается как прогноз наличия явления;

¹ Relative Operating Characteristic — относительная рабочая характеристика.

— прогноз явления с вероятностью, которая меньше порогового значения, рассматривается как прогноз отсутствия явления.

Таким образом, для каждой пороговой вероятности строится обычная таблица сопряженности, а на графике попаданий—ложных тревог строится набор точек по числу интервалов вероятностей. Этот график и определяет кривую, называемую ROC-кривой. ROC-кривая по определению проходит через точки с координатами (0,0) и (1,1) соответственно для событий, прогнозируемых только со значениями вероятности более 100 % (чего никогда не бывает), и для всех значений вероятности не менее 0 % (эта градация включает в себя весь набор возможных реализаций, так как всегда либо ни один член ансамбля не прогнозирует событие, либо сколько-то членов ансамбля прогнозируют событие).

Мерой качества прогноза является площадь под этой кривой, которая лежит в диапазоне от 1 (идеальная вероятностная прогностическая система) до 0. Площадь, равная 0,5, соответствует диагональному графику, при котором доли точных прогнозов равны долям ложных тревог. Отрицательные значения площади ROC (график под диагональной кривой) соответствуют системе, в которой доли ложных тревог больше долей точных прогнозов, т. е. отсутствует полезный сигнал. В мировой практике считается, что площадь ROC = 80 соответствует хорошей прогностической системе, а площадь ROC = 70 — полезной [4].

В реализованной системе оценивают события: «количество суммарных осадков больше некоторого порогового значения». Используют следующие пороговые значения: «осадки больше 0,1 мм/6 ч, 1 мм/6 ч, 2,5 мм/6 ч, 5 мм/6 ч». Данные пороговые значения соответствуют функции распределения осадков и известны из литературы и опыта отдела среднесрочных прогнозов Гидрометцентра России. Значение 5 мм/6 ч соответствует интенсивным осадкам.

Рассмотрим подробно случай ансамблевого прогноза для начальной даты 22 ноября 2005 г. (рис. 2).

Для осадков больше 0,1 мм/6 ч в четырех случаях из шести площадь ROC = 0,65...0,67, для Западной Сибири площадь ROC = -0,46. Интенсивные осадки (больше 5 мм/ч) в некоторых регионах предсказываются с очень хорошим качеством:

- площадь ROC = 0,79 для Западной Европы с заблаговременностью 12 ч;
- площадь ROC = 0,80 для европейской территории России (ЕТР) с заблаговременностью 60 ч.

Однако для Западной Сибири отсутствует полезный сигнал.

Все перечисленные данные являются результатами отдельного случая. Статистически значимые оценки будут получены после завершения периода архивирования результатов в базе данных.

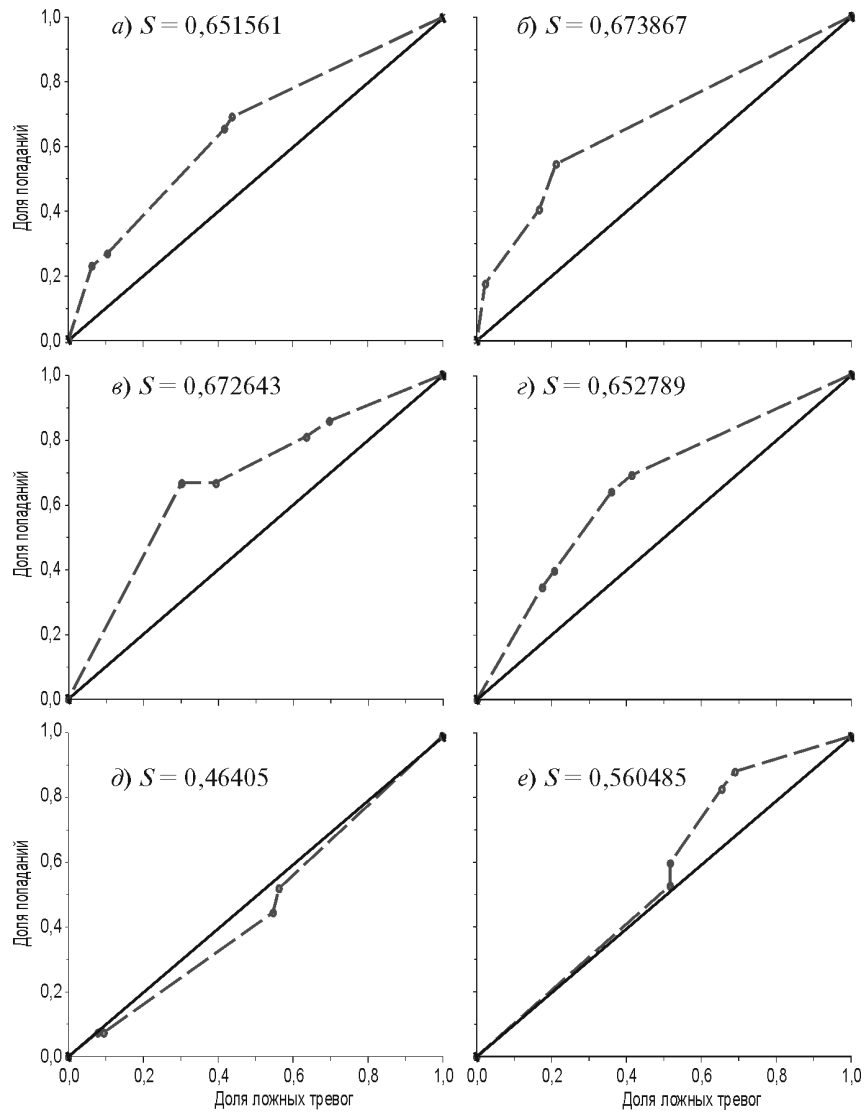
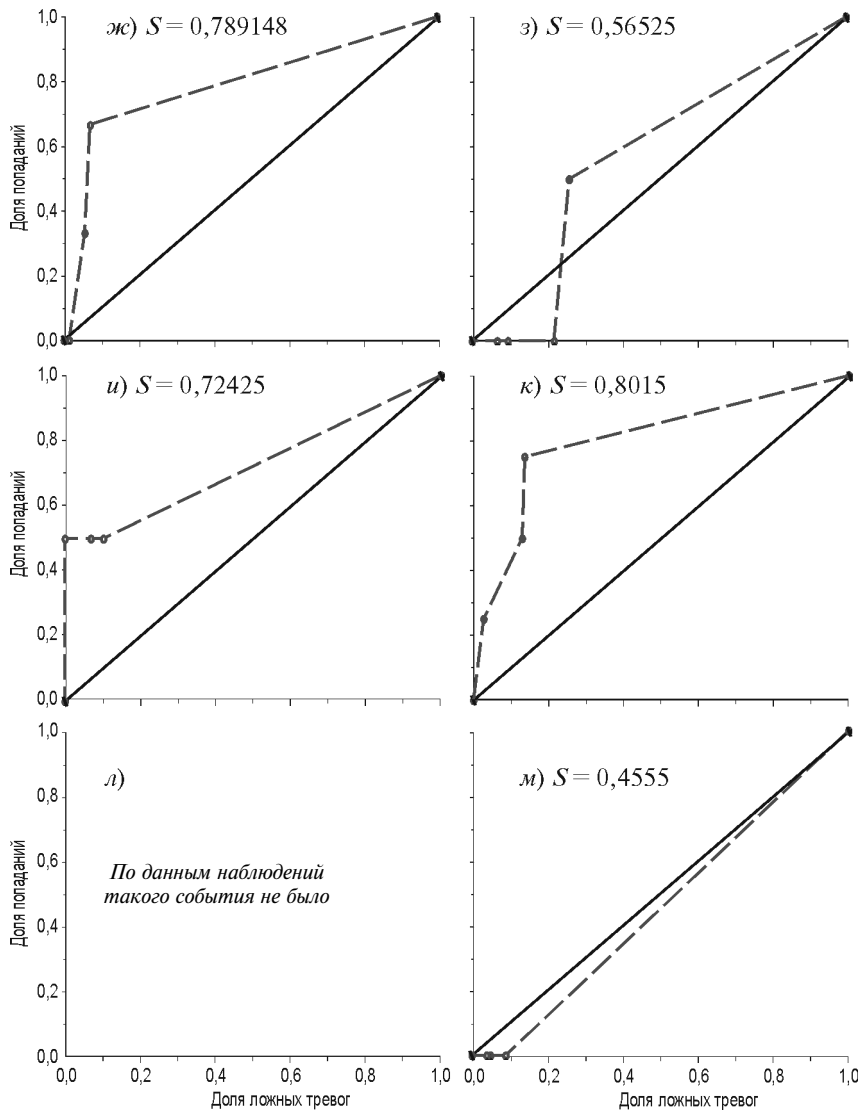


Рис. 2. ROC-кривые суммарных осадков больше 0,1 мм/6 ч (а–е)
 а, в, д, ж, и, л — заблаговременность 12 ч,
 а, б, в, ж, з — Западная Европа (40–60° с. ш., 10° з. д.–20° в. д.); в, г, и, к — ЕТР (45–65 с. ш.,



и больше 5,0 мм/6 ч (ж—м). Начальная дата: 22 ноября 2005 г., срок 12 ч.

б, г, е, з, к, м — заблаговременность 60 ч;

20° в. д.—60° в. д.); д, е, л, м — Западная Сибирь (50—70° с. ш., 60—90° в. д.)

Ранговая гистограмма (диаграммой Талаграна)

В этом методе учитывается попадание фактической величины F в пределы одной из $N + 1$ градаций ранжированного ряда из N членов ансамбля в предположении независимости членов ансамбля друг от друга [5, 9]. В каждом прогностическом эпизоде имеется ансамбль значений прогностической величины:

$$x_1, \dots, x_N.$$

Из этого ряда строится ранжированный ряд величин

$$x_{(1)} \dots x_{(N)}.$$

Оговариваются условия присваивания номеров градаций при равенстве членов ранжированного ряда и при попадании фактического значения на границу градации. Значение гистограммы в каждой градации определяется из условия

$$r_1 = P(F < x_{(1)}), \quad r_j = P(x_{(j-1)} \leq F < x_{(j)}), \quad r_{N+1} = P(x_{(N)} \leq F),$$

где P — частота попадания F в каждую из градаций $j = 1, \dots, N + 1$, т. е. число случаев попадания факта (F) в пределы градации, деленное на общее число случаев.

Теоретически функция распределения прогностического ансамбля, построенного на основе идеальной модели, совпадает с функцией распределения фактической величины. Это выражается в строго равномерном (горизонтальном) расположении результирующей кривой гистограммы. Отличие кривой гистограммы от равномерной — четкий признак несовершенства прогностической системы. Более или менее равномерное расположение кривой свидетельствует об удовлетворительном качестве построенной ансамблевой схемы.

Оценка Брайера, диаграммы надежности и разрешения

Оценка Брайера BS (Brier score) — это аналог среднеквадратической ошибки $RMSE$ (root-mean-square error) для вероятностных прогнозов и определяется по формуле

$$BS = \overline{(p - o)^2}.$$

При помощи BS оценивается разность между прогностической вероятностью события P и его частотой o , равной 0 или 1 в зависимости от того, было событие или нет. Чем меньше значение BS — тем лучше оценка прогноза.

Оценка BSS определяется как мастерство системы вероятностного прогноза по отношению к какому-либо другому вероятностному методу:

$$BSS = (BS_{ref} - BS) / BS_{ref}$$

где BS_{ref} — оценка Брайера метода, с которым сравнивается оцениваемый метод.

BS можно разложить на три составляющие, характеризующие различные аспекты системы:

$$BS = \underbrace{(p_k - c)^2}_{\text{Надежность}} - \underbrace{(c_k - c)^2}_{\text{Разрешение}} + \underbrace{(1 - c)c}_{\text{Неопределенность}}$$

где вся выборка вероятностных прогнозов поделена на $N + 1$ категорий ($k = 1, 2, \dots, N + 1$), в каждой из которых содержится n_k прогнозов с вероятностью p_k ; c_k — частота наблюдаемого события в данной категории; c — частота наблюдаемого события по всей выборке.

Надежность — среднеквадратическая разность между прогностической вероятностью и наблюдаемой частотой в разных вероятностных категориях. Она означает способность системы прогнозировать точные вероятности. Например, явление наблюдалось в 30 % случаев прогнозов с вероятностью 30 %. Надежность направлена отрицательно, т. е. чем меньше значение этого члена — тем лучше.

Разрешение — это среднеквадратичная разность между наблюдаемой частотой в каждой категории и средней наблюдаемой частотой явления во всей выборке. Этот член означает способность прогностической системы разделять всю выборку прогнозов на подвыборки с различающимися наблюдаемыми частотами явления.

Неопределенность зависит только от изменчивости наблюдений и характеризует трудность прогностической ситуации. Она не зависит от качества прогностической системы.

Предварительные результаты расчетов ранговых гистограмм и всех компонентов оценки Брайера по расчетам от одной начальной даты показали, что для статистической обеспеченности оценок требуется большее число прогностических случаев.

Оценка экономической эффективности

Очень важен для практического использования прогнозов процесс принятия решения. Он зависит от конкретных задач пользователя. В настоящее время хорошо разработана оценка экономической эффективности, учитывающая затраты на предохранительные меры на случай

неблагоприятного погодного явления и потери в результате этого явления. Методика подробно описана в работах [3, 6, 7, 10].

Приведем окончательную формулу для расчета экономической эффективности:

$$V(p_k) = \frac{\min[r, 0] - FAR(p_k)(1 - o)r + HR(p_k)(1 - r)o - o}{\min[r, 0] - or},$$

где $r = C/L$ — соотношение затрат (C) и потерь (L) конкретного потребителя, o — наблюдаемая частота явления по всей выборке, HR и FAR^* — доли попаданий и ложных тревог соответственно, p_k — пороговая вероятность, k — номер вероятностной категории (в нашем случае $k = 1, \dots, N + 1$ для ансамбля из N членов).

Достоинство данного метода — возможность определить оптимальную пороговую вероятность $p_{\text{макс}}$ (при прогнозе с вероятностью, превышающей $p_{\text{макс}}$, принимается решение о проведении защитных мер), при которой экономическая эффективность максимальна и равна

$$V_{\text{опт}} = \text{Max}_{p_k} [V(p_k)].$$

Таким образом, метод позволяет повысить эффективность использования ансамблевого прогноза. Данная характеристика требует знания наблюдаемой частоты явления на временном интервале оценки прогностической системы, поэтому она будет рассчитана после завершения накопления данных.

Заключение

Создана система верификации среднесрочных прогнозов атмосферных осадков по данным ансамблевого моделирования в отделе среднесрочных прогнозов погоды Гидрометцентра России. Система предназначена для работы в оперативном режиме и режиме архивации, однако в связи с нехваткой вычислительных ресурсов работа в оперативном режиме в настоящее время затруднительна. Проводятся работы по созданию базы данных для архивирования результатов ансамблевого моделирования, что позволит получить статистически обеспеченные оценки качества ансамблевого моделирования количества осадков на срок до 10 сут. Архивация позволит также оценить качество ансамблевых прогнозов в конкретных синоптических ситуациях, например, количе-

* HR и FAR — «Hit rate» и «False alarm rate» соответственно.

ство осадков в интенсивном циклоне. Система ориентирована на оценку осадков, однако некоторые ее блоки можно применить для верификации прогнозов других погодных характеристик при условии, что они сформулированы в виде бинарного события.

Автор выражает благодарность всем сотрудникам отдела среднесрочных прогнозов, создавшим систему среднесрочного ансамблевого прогноза, и конкретно Ю. В. Алферову, Е. Д. Астаховой и И. А. Розинкиной за помощь и сотрудничество при реализации данной работы.

Работа выполнена при поддержке гранта РФФИ 04-05-64530-а.

СПИСОК ЛИТЕРАТУРЫ

1. Алферов Ю. В. О восстановлении значений 6-часовых сумм осадков по данным синоптических наблюдений // Тр. ГМЦ РФ. — 2000. — Вып. 334. — С. 44—51.
2. Астахова Е. Д. Построение ансамблей начальных полей для системы кратко- и среднесрочного ансамблевого прогнозирования погоды // См. наст. сб. — С. 98—117.
3. Муравьев А. В., Куликова И. А. Оценка экономической эффективности ансамблей гидродинамических прогнозов на основе вероятностных оценок качества // Метеорология и гидрология. — 2004 — № 11. — С. 17—26.
4. Buizza R., Hollingsworth A., Lalaurette F., Ghelli A. Probabilistic Predictions of Precipitation Using the ECMWF Ensemble Prediction System // Weather and Forecasting. — 1999. — Vol. 14. — P. 168—189.
5. Hamill T. M., Colucci S. J. Evaluation of Eta-RSM ensemble probabilistic precipitation forecasts // Mon. Wea. Rev. — 1998. — Vol. 126. — P. 711—724.
6. Mylne K. R. Decision-making from probability forecasts based on forecast value // Meteor. Appl. — 2002. — Vol. 9. — P. 307—315.
7. Richardson D. R. Obtaining economic value from the EPS // ECMWF Newsletter. — 1998. — N 80. — P. 8—12.
8. Standardised Verification System (SVS) for Long-Range Forecasts (LRF) // New Attachment II-9 to the Manual on the GDPS. Vol. I. — 2002. — WMO-No. 485. — P. 1—22*.
9. Talagrand O. Statistical consistency of ensemble prediction systems // Discussion paper, SAC Working Group on EPS, ECMWF. — Reading, UK, 1997. — P. 1—25.
10. Zhu Y., Toth Z., Wobus R., Richardson D., Mylne K. The economic value of ensemble-based weather forecasts // Bull. Amer. Met. Soc. — 2002. — January. — P. 73—83.

* Доступно по адресу <http://www.wmo.ch/pages/prog/www/DPS/LRF/LRF-standardised-verif-sys-2002.doc>.