

DOI: <https://doi.org/10.37162/2618-9631-2021-4-30-49>

УДК 551.509.313+551.509.324.2+551.508.85

Обзор методов пространственной верификации и их применение для ансамблевых прогнозов

А.Ю. Бундель¹, А.В. Муравьев¹, Е.Д. Ольховая²

*¹Гидрометеорологический научно-исследовательский центр
Российской Федерации, г. Москва, Россия;*

*²МИРЭА – Российский технологический университет, г. Москва, Россия
a.bundel@gmail.com, muravev@mecom.ru*

Высокое разрешение современных моделей численного прогноза погоды позволяет детально воспроизводить мезомасштабные системы с высокой амплитудой метеовеличин и большими градиентами полей. Однако увеличение разрешения приводит к росту ошибок в пространстве и времени, причем количественная оценка таких ошибок наталкивается на известную "проблему двойного штрафа". Для решения этой проблемы в последние два десятилетия разработаны такие методы "пространственной верификации", в которых игнорируются умеренные ошибки (в первую очередь в положении), но сохраняется способность оценивать полезность детализированной информации.

В статье приводится классификация методов пространственной верификации с их кратким описанием и обзором международных проектов по сравнению свойств этих методов. Особое внимание уделяется публикациям по применению данных методов в ансамблевом прогнозировании. Перечисляются соответствующие программные средства и прилагается список наиболее употребимых в данной области англоязычных терминов с вариантами перевода на русский язык.

Ключевые слова: модели высокого разрешения, верификация, двойной штраф, пространственные методы, ансамблевый прогноз, объектно-ориентированные методы

Overview of spatial verification methods and their application to ensemble forecasting

A.Yu. Bundel¹, A.V. Muraviev¹, E.D. Olkhovaya²

¹Hydrometeorological Research Center of Russian Federation, Moscow, Russia;

*²MIREA – Russian Technological University, Moscow, Russia
a.bundel@gmail.com, muravev@mecom.ru*

State-of-the-art high-resolution NWP models simulate mesoscale systems with a high degree of detail, with large amplitudes and high gradients of fields of weather variables. Higher resolution leads to the spatial and temporal error growth and to a well-known double penalty problem. To solve this problem, the spatial verification methods have been developed over the last two decades, which ignore moderate errors (especially in the position), but can still evaluate the useful skill of a high-resolution model.

The paper refers to the updated classification of spatial verification methods, briefly describes the main methods, and gives an overview of the international projects for inter-comparison of the methods. Special attention is given to the application of the spatial

approach to ensemble forecasting. Popular software packages are considered. The Russian translation is proposed for the relevant English terms.

Keywords: high-resolution models, verification, double penalty, spatial methods, ensemble forecasting, object-based methods

Введение

По мере увеличения компьютерных ресурсов все шире используются модели численного прогноза погоды (ЧПП) высокого пространственного разрешения (2 км и выше). В целом, увеличение разрешения модели негативно сказывается на результате оценок для некоторых метеоэлементов при оценке стандартными методами. Модель высокого разрешения воспроизводит более точно очерченные мезоструктуры, большую амплитуду изменения метеоэлементов при более точной локализации и большие градиенты, что неизбежно ведет к большим ошибкам в пространстве и времени. Возникает проблема "двойного штрафа": при стандартном оценивании "точка с точкой", когда прогноз явления незначительно сдвинут по сравнению с наблюдаемым явлением, он штрафуются дважды: в точке, где явление прогнозировалось, но не случилось в реальности, и там, где оно было, но не спрогнозировано. Тем не менее такая модель дает больше полезной информации по сравнению с версией с более грубым разрешением, которая выдает более сглаженные поля. Возникает потребность в методах верификации, которые обладают некоторой терпимостью к умеренным ошибкам в пространстве и времени, сохраняя при этом способность оценивать истинное качество информации, которую дает модель высокого разрешения.

Кроме того, размерность задачи верификации пространственных полей метеоэлементов огромна, поскольку учитываются все точки сетки. Использование скалярных (одно число) метрик облегчает задачу, но не решает всех проблем. Традиционные оценки, сравнивающие прогноз и наблюдение в точке, имеют недостатки при оценке прогнозов высокого разрешения. Например, если мы оцениваем явление "Наличие осадков", все прогнозы с одинаковым количеством точек с явлением и отсутствием пересечения областей осадков в прогнозе и наблюдениях будут иметь одинаковые показатели, независимо от расстояния между областями осадков в поле прогноза и наблюдений, интенсивности осадков, формы областей и других свойств. Проблема двойного штрафа проиллюстрирована на рис. 1. Все прогнозы на рис. 1, кроме (е), дадут, например, оценку угроз $TS = 0$, хотя (е) субъективно не является лучшим прогнозом. Во многих ситуациях (а) является лучшим прогнозом, хотя будет иметь такие же плохие стандартные оценки, как и откровенно неправильный прогноз.

На протяжении последних 20 лет был разработан ряд методов, позволяющих уменьшить влияние двойного штрафа. Такие методы традиционно называются пространственными (spatial methods) [19]. Однако

некоторые из них позволяют учесть также сдвиг по времени. Они не наказывают незначительные сдвиги прогноза относительно наблюдения в пространстве и времени, т. е. позволяют избежать двойного штрафа. Также пространственные методы помогают снизить размерность задачи верификации, главным образом за счет введения понятия объекта с набором свойств вместо всего множества точек поля.

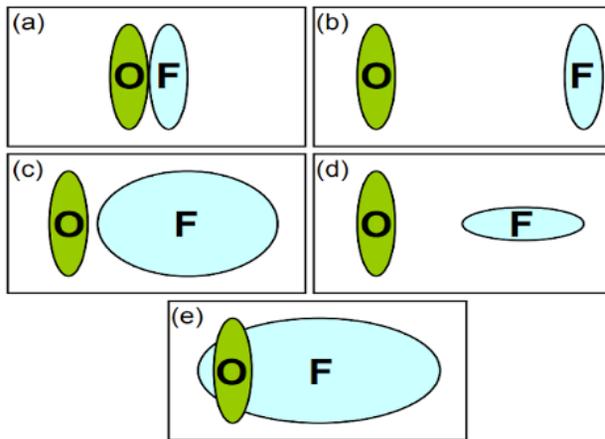


Рис. 1. Гипотетическая область прогноза (F) и наблюдения (O) явления [8]. При стандартном оценивании «точка с точкой», когда прогноз явления незначительно сдвинут по сравнению с наблюдаемым явлением, он штрафуются дважды: в точке, где явление прогнозировалось, но не случилось в реальности, и там, где оно было, но не спрогнозировано.

Fig. 1. Idealized areas of forecast and observation events [8]. The slight shift of the forecast event compared to the observation is penalized twice by the point-based scores: where the event is forecast but not occurred, and where it occurred but not forecast.

Традиционную оценку мезомасштабных прогнозов рекомендуется дополнять методами и приемами пространственной верификации, особенно это касается разрывных полей, таких как поля количества осадков и облачности.

1. Классификация пространственных методов

Новизна методов пространственной верификации в сопоставлении с традиционными обусловлена в первую очередь богатством и многозначностью понятий "объекта" и "сходства объектов" в сопоставлении с предельной простотой понятий "точки" и "разности двух чисел".

Обилие предлагаемых пространственных методов и не всегда понятные свойства порождаемой ими информации послужили стимулом к их целенаправленному анализу в рамках международных проектов, среди которых наиболее значимыми представляются проект ICP по сравнению

методов пространственной верификации (spatial forecast verification Inter-Comparison Project, ICP, время выполнения 2007–2009 гг.) и развивающий его результаты проект MesoVICT по сравнению методов мезомасштабной верификации над сложным рельефом (Mesoscale Verification Inter-Comparison over Complex Terrain, время выполнения 2014–2019 гг.) [15, 16]. Подчеркнем, что речь идет о сравнении самих методов верификации и их свойств, т. е. о "метаверификации" – науке о методах и о системах верификации [24], а не о сравнении прогностических систем. Результаты этих проектов уже находят практическое применение в исследованиях и в оперативной деятельности метеорологических центров и служб [9]. В России пространственный аспект верификации разрабатывался в работах [1–6]. При этом многие методологические вопросы пространственной верификации еще нуждаются в уточнении и далеки от решения.

В проекте ICP были сформулированы следующие задачи:

- 1) каким образом методы информируют о качестве на разных масштабах;
- 2) каким образом методы информируют о локальных ошибках;
- 3) поставляют ли методы информацию об ошибках в интенсивности и в распределении;
- 4) поставляют ли методы информацию об ошибках в структуре;
- 5) поставляют ли методы информацию о попаданиях, промахах и ложных тревогах;
- 6) есть ли в методах что-либо, противоречащее интуиции;
- 7) имеются ли в методах какие-либо параметры и насколько результаты чувствительны к их выбору;
- 8) насколько просто производить агрегирование результатов по множеству случаев;
- 9) способны ли методы идентифицировать ошибки в определении моментов времени;
- 10) могут ли быть быстро рассчитаны доверительные интервалы и критерии гипотез.

В сравнении участвовало 18 методов, разработанных в период с 2000 по 2009 год. В самом начале выполнения проекта была проведена классификация почти всех разработанных до этого времени методов на две большие категории – на методы фильтрации (filtering methods) и методы смещения (displacement methods). В свою очередь, методы фильтрации были подразделены на две группы – окрестные и масштабно-разделительные (neighborhood and scale separation), а методы смещения – на объектно-ориентированные и деформационные (features based and field deformation) [20].

В статье [7] описаны использованные участниками проекта наборы идеализированных геометрических и возмущенных прогнозов с заданными ошибками, а также набор реальных прогнозов в центральной части США конвективных осадков на сетке 4 км. Геометрические случаи

состояли в представлении областей осадков в виде эллиптических форм со сдвинутым ядром высокой интенсивности и при типичных, схематизированных ошибках прогнозирования этих областей (сдвиг по пространству, различие площадей, занятых осадками, разница в частоте прогнозирования явления по отношению к частоте явления по наблюдениям и др.).

В качестве "традиционных оценок" качества прогнозов использованы 13 показателей, в том числе: критический индекс успешности – Critical Success Index, CSI; показатель Пирса – Обухова – Hanssen – Kuipers (H-K); показатель мастерства Хайдке – Heidke skill score (HSS); предупрежденность явления, ложная предупрежденность явления, среднеквадратическая ошибка RMSE, коэффициент корреляции. Из методов пространственной верификации особое внимание уделено оценке по пространственным долям FSS (Fractions Skill Score), разложению по масштабам ошибки прогноза (Intensity Scale, IS), объектно-ориентированному методу SAL (Structure, Amplitude, Location) (некоторые из этих методов описаны ниже).

Было проведено сравнение методов, проанализированы результаты для геометрических случаев и представлены субъективные и традиционные оценки по реальным случаям. В результате, в частности, выяснилось, что все новые методы смогли обнаружить систематическую ошибку (bias), а объектно-ориентированные и деформационные методы оказались способными диагностировать ошибки смещения областей осадков.

При сравнении модельных прогнозов с реальными данными традиционные оценки чаще всего не совпадали с субъективными оценками прогнозов синоптиками. В [7] обсуждается важная проблема согласования "объективного" оценивания прогнозов одной мерой с субъективными оценками синоптика-прогнозиста. Притом что никакая единственная традиционная оценка не может дать содержательной информации, было бы желательно отыскать некоторое объединение традиционных оценок в одну такую, которая бы привела к согласию с субъективным оцениванием.

Основные итоги выполненного проекта ICP были сформулированы в статье [19]. Была составлена итоговая таблица соответствия четырех групп методов следующим пяти типам порождаемой ими информации с характеристиками: 1) пространственных масштабов мастерства прогнозов; 2) ошибок локализации; 3) ошибок в интенсивности; 4) ошибок в структуре; 5) повторяемости (попадания, промахи и ложные тревоги). По таблице, в частности, можно увидеть, что методы фильтрации (окрестные и масштабно-разделительные) способны описать "масштаб", на котором прогноз достигает определенного уровня мастерства, а методы смещения (объектно-ориентированные и деформационные) способны оценить ошибки в смещениях.

В проекте MesoVICT основное внимание было уделено новым вопросам [16]:

- Какова способность метода оценивать переменные помимо осадков?
- Может ли метод быть адаптирован для ансамблевого прогноза?
- Применим ли метод в условиях сложного рельефа?
- Насколько чувствителен метод к настройчным параметрам, выбору метода интерполяции?
- Адекватно ли метод сравнивает модели грубого и высокого разрешения?
- Возможен ли в методе учет неопределенности наблюдений?

В рамках MesoVICT была создана база данных модельных, в том числе ансамблевых прогнозов высокого разрешения, а также качественный архив наблюдений, включая ансамбль наблюдений VERA в регионе Альп. Также было предложено шесть основных тестовых случаев с характерными синоптическими условиями.

Классификация пространственных методов была дополнена и уточнена, отдельно была выделена группа мер дальности (distance metrics) (рис. 2). Пересечения на рисунке означают взаимопроникновение отдельных групп. Так, например, сглаживание, характерное для методов фильтрации, часто применяется в качестве начального шага в объектно-ориентированных методах.

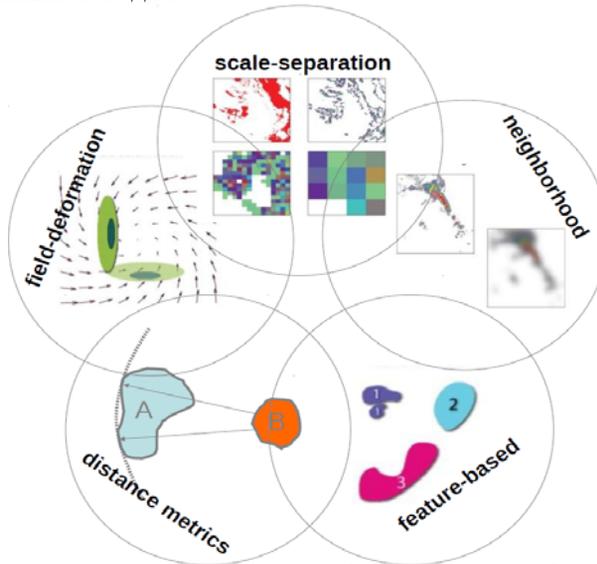


Рис. 2. Классификация пространственных методов верификации [16] на окрестные и масштабно-разделительные (neighborhood and scale-separation), объектно-ориентированные и деформационные (feature-based and field-deformation), а также меры дальности (distance metrics).

Fig. 2. Classification of spatial verification methods [16]: neighborhood and scale-separation methods, feature-based and field-deformation methods, and distance metrics.

В результате MesoVICT появилось много исследований по применению пространственных методов для нетрадиционных переменных, сложного рельефа, ансамблей (часть из них цитируется в данной работе), однако результаты проекта еще находятся в стадии обобщения.

Приведем краткое описание основных групп методов.

2. Окрестные (neighborhood) методы

Раньше эти методы чаще назывались Fuzzy (нечеткие). Основное их свойство – доверие к прогнозу в окрестности [18]. Оценка соответствия прогноза явления и наблюдения в точке замещается оценкой прогноза и/или наблюдения в окрестности точки, которая предполагает какую-либо обработку данных внутри окна (окрестности), например:

- Расчет среднего значения (апскейлинг – upscaling)
- Наличие явления где-нибудь внутри окна
- Частота явления в окне – аналог вероятности
- Распределение величин в окне

Если в окне рассматривается также наблюдение – это позволяет косвенно учесть неопределенность наблюдений.

Возможно, самый популярный окрестный метод – это оценка по пространственным долям, Fractions Skill Score, FSS [25]. Рассматривается доля ячеек в окрестности, занятых явлением, P , в поле прогноза и поле наблюдений (рис. 3).

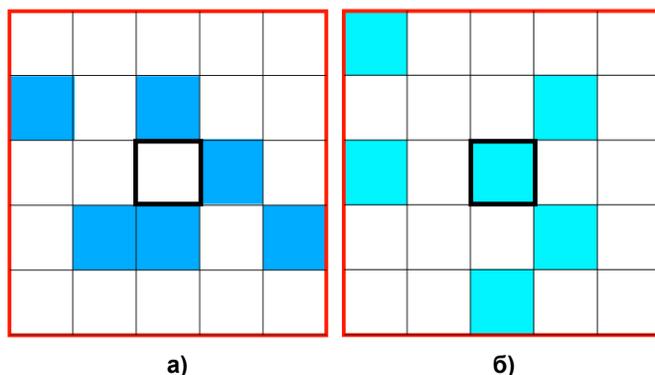


Рис. 3. Доля ячеек, занятых явлением (закрашенные ячейки) в окрестности 5×5 точек одинакова для прогноза (а) и наблюдения (б) и равна $6/25=0.24$, в то время как оценка только в центральной точке окрестности (выделена жирной черной рамкой) дала бы плохие оценки (по прогнозу явления нет, а по наблюдению есть).

Fig. 3. Fraction of grid cells with an event in a 5×5 neighborhood is equal for the forecast (a) and observation (b) ($6/25=0.24$). The score at the central point only would be low.

Далее рассчитывается среднеквадратическая разность P прогноза и наблюдений по всем точкам сетки N , в знаменателе FSS фигурирует наилучший возможный прогноз:

$$FSS = 1 - \frac{\frac{1}{N} \sum_{i=1}^N (P_{fcst} - P_{obs})^2}{\frac{1}{N} \sum_{i=1}^N P_{fcst}^2 + \frac{1}{N} \sum_{i=1}^N P_{obs}^2}$$

Величина FSS полезного прогноза задается больше $FSS_{uniform} = 0.5 + f_0/2$, где f_0 доля ячеек с явлением по наблюдениям во всей области прогноза ($FSS_{uniform}$ соответствует FSS в центральной точке окрестности, которую имел бы прогноз со значением f_0 в каждой точке области). Оценку FSS следует рассчитывать для размеров окрестности возрастающей величины: в центральной точке, окрестности 3×3 точки, 5×5 , 7×7 и т. д.

С помощью FSS мы хотим узнать:

- Как изменяется мастерство прогноза с ростом окрестности?
- Какой наименьший размер окрестности дает достаточно точные прогнозы?
- Дает ли модель более высокого разрешения более точные прогнозы?

В [28] показан способ использования FSS для оценки смещения прогностической области явления в пространстве относительно наблюдений. Метод основан на соотношении между FSS, размером окрестности и смещением по пространству. Показано, что метод работает только при отсутствии большого сдвига частоты (FBI, frequency bias) в прогнозе, конкретно при значениях FBI от 0,75 до 1,5. Однако это ограничение можно преодолеть заданием порогового значения на основе перцентиля исследуемой переменной. Далее Г. Скок [29] ввел модификацию FSS, окрестная мера мастерства – Neighborhood Skill Score, NSS. Оценка смещения NSS не ограничена значениями FBI, что делает ее более удобной в использовании.

В [14] FSS была использована для расчета соотношения пространственного разброса по ансамблю и ошибки ансамбля. Для этого были рассчитаны среднее значение и стандартное отклонение FSS, рассчитанной для каждой пары членов ансамбля (dispersive FSS, dFSS), а также среднее и стандартное отклонение FSS, рассчитанной для каждого члена ансамбля и наблюдения (eFSS). Высокие значения среднего dFSS означают, что все члены ансамбля предсказывают область явления близко друг от друга, соответственно, пространственный разброс по ансамблю мал. Стандартное отклонение dFSS показывает интервал значений dFSS, высокое стандартное отклонение dFSS означает, что имеются выбросы с особенно низким или высоким dFSS. Для eFSS делаются соответствующие заключения, но не для разброса внутри ансамбля, а для ошибки каждого

из членов ансамбля по отношению к наблюдениям. Затем проводится анализ соотношения пространственного разброса и ошибки ансамбля аналогично стандартному анализу соотношения разброс-ошибка ансамбля (spread-error relationship). Данный анализ широко применяется в настоящее время, например в [10].

При применении FSS и ее модификаций следует иметь в виду, что эти оценки дают неточные результаты для объектов, близких к границе области.

Примеры пространственных оценок с использованием апскейлинга, а также их сравнение с традиционными оценками для экспериментальной системы сверхкраткосрочного прогноза погоды, подготовленной в ФГБУ «ЦАО» с использованием модели атмосферы WRF ARW, системы трехмерного вариационного усвоения данных WRF 3D-Var и технологий даунскейлинга даны в [1].

3. Масштабно-разделительные методы (scale-separation methods)

Масштабно-разделительные методы оценивают ошибку прогноза в зависимости от масштаба. Наиболее известный метод – разложение по масштабам ошибки прогноза – метод Intensity-scale verification approach (IS) [11]. Бинарная ошибка прогностического поля относительно поля наблюдений раскладывается на сумму компонентов разных масштабов, как правило, с помощью вейвлет-фильтра (например, вейвлет 2D Haar). Для компонентов каждого из масштабов рассчитывается ряд оценок (среднеквадратическая мера мастерства, оценка Брайера, и т. п.).

Отметим, что различие окрестных и масштабных методов состоит в том, что окрестные методы отфильтровывают ошибку на мелких масштабах, а масштабные методы выделяют ошибку на каждом из масштабов.

4. Объектно-ориентированные методы

Объектно-ориентированная верификация (object-based или features-based) предназначена для оценки способности прогностической системы воспроизводить структурные элементы полей вместе с их перемещением и эволюцией [12, 13, 17, 23]. Основные и первые примеры применения пространственной верификации относятся к осадкам в силу их высокой пространственно-временной изменчивости, которая затрудняет идентификацию и прогнозирование областей осадков и делает осадки своего рода "метрикой качества гидродинамических моделей".

Под "объектами" здесь понимаются пространственные, в основном односвязные области полей на земной поверхности, выделяемые некоторым пороговым значением исследуемой величины, например – суммарное количество осадков больше некоторого порога. Таким образом, идентификация объекта производится, как правило, выбором произвольной

изолинии, очерчивающей замкнутую границу объекта. Часто предварительно проводят процедуру сглаживания исходного поля, например, с помощью фильтра свертки, т. е. осреднения внутри некоторого радиуса вокруг каждой точки поля.

Более сложен вопрос связности области, выделенной замкнутой линией – учитывать ли многосвязность или устранять её? Собственно говоря, требуется выбрать такой радиус осреднения, при котором и граница станет более гладкой, и область более односвязной, но при этом сохранятся важные для анализа области малого масштаба. Все эти проблемы возникают каждый раз при анализе и прогнозе областей мезомасштабных осадков, и удовлетворить все запросы не представляется возможным.

После пространственного осреднения полей, выделения объектов подходящей гладкости границ и близких к односвязности возникает очередная проблема статистического описания объектов. Простейшие характеристики – это центр тяжести, учитывающий интенсивность осадков, и размер, оцениваемый в количестве точек или в площади. Несмотря на сглаживание, границы зон осадков могут принимать самые разнообразные формы и для характеристик объекта, отличных от его объема (как, например, в [13]), сопоставление полос в наблюдениях и кругов в прогнозах становится затруднительно.

В России объектно-ориентированный подход к верификации применен в [3–6].

CRA (Contiguous Rain Area), Сопряженная область дождя – объектно-ориентированный метод верификации прогнозов, предложенный в [17]. Вкратце он состоит в следующем: вначале в полях наблюдений и прогноза выделяются объекты (чаще всего осадков) на основе порогового значения. Затем с помощью той или иной функции (в которой используется расстояние между объектами) выделяются пары объектов в поле наблюдений и прогнозов. Затем определяется оптимальный сдвиг прогностического объекта в пространстве к наблюдаемому объекту путем минимизации ошибки. Здесь также могут применяться различные критерии, например коэффициент корреляции или среднеквадратичная ошибка MSE. Находится разность исходной MSE (представляемой как $MSE_{total} = MSE_{displacement} + MSE_{volume} + MSE_{pattern}$) и MSE после сдвига, это ошибка из-за сдвига в пространстве $MSE_{displacement} = MSE_{total} - MSE_{shifted}$. Оставшаяся MSE состоит из квадрата разности среднего количества осадков в прогностическом и наблюдаемом объекте $MSE_{volume} = (F - X)^2$ и мелкомасштабных различий $MSE_{pattern} = MSE_{shifted} - MSE_{volume}$. Пример применения CRA для оценки наукастинга ФГБУ «Гидрометцентр России» можно найти в [6]

MODE, Method for Object-Based Diagnostic Evaluation – Метод объектно-ориентированной диагностической верификации [12, 13] предназначен для сравнения сеточных прогнозов и сеточных наблюдений.

Метод MODE может использоваться для решения общих задач сравнения любых двух полей, содержащих однозначно определенные объекты. Метод MODE был разработан, чтобы приблизить объективный расчет качества прогноза объектов к субъективному суждению синоптика.

Расчеты оценок с помощью MODE производятся в следующей последовательности.

1. Определение объектов в полях прогноза и наблюдения с помощью пользовательских параметров (выбор фильтра для сглаживания полей, выбор пороговых значений).

2. Выбор и расчет атрибутов для каждого из этих объектов: например, площади, центроиды ("центры масс"), углы наклона оси, интенсивности, сложность контура и др.

3. Расчет разностей между атрибутами для каждой пары объектов в поле прогноза и наблюдения: например, отношение площадей, отношение площади перекрытия, расстояние между центроидами, разность углов наклона, отношение интенсивностей и др.

4. Выбор функций нечеткой логики F для расчета суммарной важности (total interest) для каждой пары объектов из соответствующих полей прогноза и наблюдения (на рис. 4 приведен пример таких функций). Смысл их в том, что хорошему соответствию атрибута (например, малому расстоянию между центроидами) придается значение 1, которое уменьшается (как правило, линейно) до нуля вплоть до значения разности атрибутов, бессмысленного для пользователя (например, расстояние между центроидами слишком большое). Таким образом, все значения разностей атрибутов переводятся с помощью функций нечеткой логики в значения от 0 до 1.

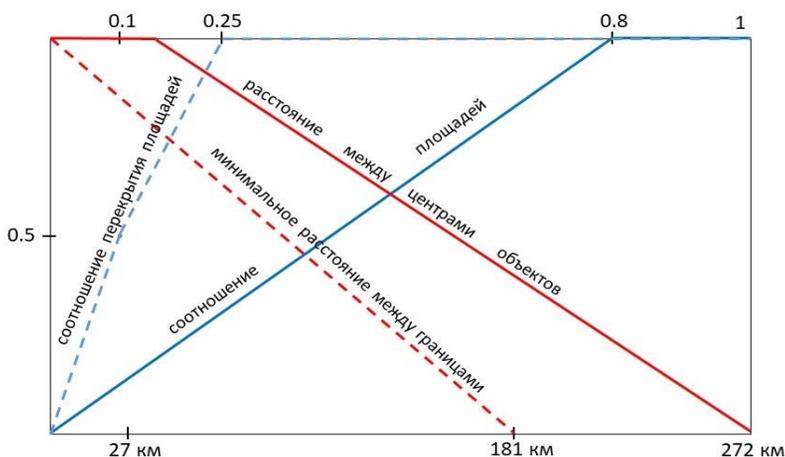


Рис. 4. Пример функций нечеткой логики для перевода разностей атрибутов в значения от 0 до 1.

Fig. 4. Example of fuzzy logic functions to pass from attribute differences to values from 0 to 1.

5. Выбор весов w для каждого из атрибутов, а также коэффициента надежности c (отражающего, насколько мы уверены в точности определения значения атрибута. Например, при малой разности углов наклона надежность меньше; на практике обычно берут 1).

6. Вычисление суммарной важности (Total Interest, TI) для всех пар объектов в поле прогноза и наблюдений:

$$TI_j = \frac{\sum_i I_{ij}}{\sum_i w_i c_i}$$

$$I_{ij} = F_{ij} w_i c_i$$

где i – номер атрибута; j – номер пары объектов.

7. Выбор порогового значения TI для сопоставления (спаривания) объектов в парах. Если TI больше порогового значения в какой-либо паре объектов прогноз-наблюдение, данный объект считается правильно спрогнозированным. Имеется также ряд алгоритмов для объединения (слияния) объектов внутри одного и того же поля – получение кластеров объектов.

8. Вывод обобщенных статистических характеристик отдельных объектов, пар объектов и кластерных объектов, полученных сопоставлением и объединением простых объектов. Расчет медианы максимального интереса – основной интегральной характеристики качества прогноза MODE.

Достоинства MODE:

- Может применяться в режиме спаривания объектов и нет
- Рассчитывает единый показатель качества (ММИ)
- Дает информацию о пропусках и ложных тревогах

MODE – очень гибко настраиваемый метод за счет введения функций нечеткой логики, пороговых значений и весов в зависимости от цели и объекта верификации. Это в то же время является его недостатком, поскольку MODE содержит много настроечных параметров. В отличие от CRA, где общая ошибка сразу раскладывается на составные части, характеризующие разные аспекты качества прогноза объектов, суммарная важность TI в MODE объединяет несколько аспектов качества прогнозов. Это, однако, может быть преодолено расчетом TI для различных атрибутов по отдельности.

5. Применение пространственных методов для ансамблевого прогноза

В последние несколько лет применение пространственных методов для ансамблей является особенно актуальной задачей, поскольку с ростом вычислительных мощностей стал возможным ансамблевый прогноз высокого разрешения. Как и в случае детерминированного прогноза, наиболее прочно вошла в практику верификации комбинация окрестного

и объектно-ориентированных подходов. Применение FSS для ансамбля было описано в п. 3. Остановимся на еще двух методах, доказавших свою перспективность.

Метод ансамблевой вероятности в окрестности – Neighborhood ensemble probability, NEP, и **метод максимальной вероятности в окрестности** – Neighborhood maximum ensemble probability, NMEP [26, 27]. В [27] метод создания вероятностного прогноза из детерминистского на основе количества точек с явлением в окрестности [30] был перенесен на ансамблевый прогноз. В [26] были введены названия Neighborhood ensemble probability, NEP, и метод максимальной вероятности в окрестности – Neighborhood maximum ensemble probability, NMEP. В NEP рассчитывается вероятность явления в каждой точке выбранной окрестности, а затем находится средняя в окрестности вероятность. В NMEP для каждой точки рассчитывается доля членов ансамбля, для которых явление предсказывается где-либо в окрестности точки. Как и обычно, явление задается с помощью порогового значения.

На рис. 5 приведен пример поля NMEP. Далее полученное поле оценивается с помощью обычных вероятностных метрик качества. Такие продукты полезны для синоптика, однако часто они оказываются слишком сглаженными и возникает потребность дополнять их объектно-ориентированными методами [22].

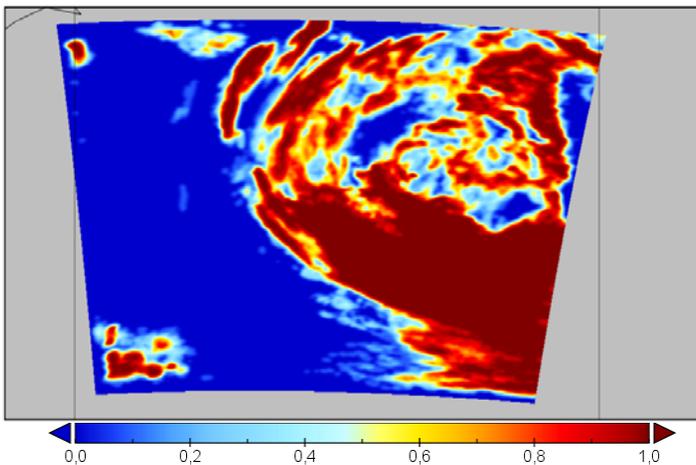


Рис. 5. Ансамблевая максимальная вероятность количества осадков > 0.1 мм/ч в окрестности (Neighborhood maximum ensemble probability, NMEP), Ансамбль ICON-Ru-EPS, 2.2 км, 11 членов, окрестность – 9 точек, старт – 00 ч ВСВ, дата – 01.07.2021.

Fig. 5. Neighborhood maximum ensemble probability, NMEP, precipitation accumulation > 0.1 mm/h, ICON-Ru-EPS ensemble with 2.2 km grid mesh, 11 members, neighborhood of 9 points, run from 2021.07.01, 00 UTC.

Метод псевдочлена ансамбля [22] основан на применении MODE для ансамбля. Джонсон с коллегами из университета Оклахомы на протяжении десятилетия развивают свой подход к оцениванию конвективно-разрешающего ансамбля. Вначале они использовали объекты контрольного члена ансамбля как эталон и находили соответствующие объекты в полях остальных членов ансамбля. На основе количества объектов, соответствующих каждому из объектов контрольного члена, рассчитывалась вероятность каждого из объектов контрольного члена [21]. Ограничением этого метода является выбор контрольного члена в качестве основного. Затем метод был усовершенствован. В новом методе создается так называемый «псевдочлен» ансамбля из наиболее репрезентативных объектов.

На рис. 6 приведен пример создания псевдочлена.

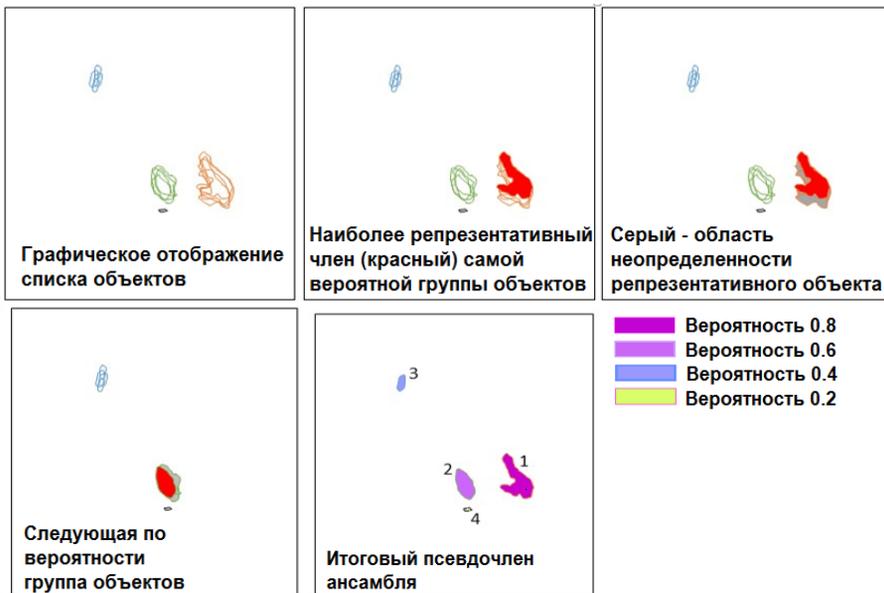


Рис. 6. Схематический пример создания псевдочлена ансамбля.

Fig. 6. Schematic picture of pseudomember creation.

Алгоритм состоит из следующих шагов:

- Создать список всех объектов в ансамбле
- Рассчитать суммарную важность (Total Interest, TI, см. п. 5 MODE) для всех пар объектов из всех членов ансамбля
- Определить соответствующие друг другу пары объектов (matching) на основе порогового значения суммарной важности TI
- Рассчитать вероятность появления для каждого объекта на основе количества членов ансамбля с объектом, соответствующим данному (включая его самого), по отношению к общему количеству членов ансамбля

- Выбрать группу соответствующих друг другу объектов с наибольшей вероятностью
- Внутри этой группы выбрать репрезентативный объект с наибольшим значением средней суммарной важности по отношению к другим объектам группы, добавить этот объект в псевдо-член. Убрать этот объект и всю группу соответствующих ему объектов из списка
- Выбрать группу соответствующих объектов со следующим значением вероятности
- Повторить данный алгоритм, пока в списке не останется ни одного объекта

Таким образом, на выходе мы получаем список объектов псевдо-члена с разными вероятностями, а также область неопределенности каждого объекта псевдо-члена, определяемую группой влияющих (соответствующих ему) объектов. Далее можно установить пороговые значения вероятности и рассчитать обычные оценки MODE для полученных объектов.

6. Наиболее популярные пакеты программного обеспечения для реализации пространственных методов

CRA. Исторически первым стал пакет CRA, написанный на языке IDL и реализованный для нужд австралийской службы погоды. Пакет опробован в нескольких странах, получил высокую оценку и стал применяться в экспериментальном режиме в некоторых метеослужбах. Однако, с одной стороны, узкая направленность методики с расчетом около десятка характеристик, а с другой стороны, использование лицензионного языка программирования не способствовали широкому распространению пакета, хотя он продолжает модифицироваться и до сих пор служит средством оценок и удобным пособием для университетского обучения в некоторых странах [17].

MET. Система верификации прогностических моделей MET (Model Evaluation Tools), первая версия которого выпущена в январе 2008 г., была разработана в США Исследовательским Центром для Тестирования моделей DTC (Developmental Testbed Center) [<https://dtcenter.org/community-code/metplus>]. Пакет MET включает возможность расчета оценок на основе окрестного подхода, метод MODE, а также метод разложения по масштабам ошибки прогноза, Intensity-scale (IS).

SpatialVx (R). Комплекс SpatialVx ("библиотека" в языке R) является, возможно, наиболее современным, полным и подробно документированным программным средством, реализующим большинство известных методов пространственной верификации. Комплекс написан Э. Гиллеландом в 2012 г. по плану научно-исследовательской деятельности лаборатории

RAL и по результатам выполнения проекта ICP. Автор регулярно производит корректировку программного обеспечения, включая исправление обнаруженных пользователями ошибок, и расширяет комплекс новыми возможностями [<https://cran.r-project.org/web/packages/SpatialVx/index.html>]. Приведем список методов пространственной верификации, заключенный в комплексе версии 0.7 от 31 августа 2020 г.

Автор сохраняет классификацию методов на две большие категории методов фильтрации и методов перемещения, введя группу мер близости [16] во вторую категорию. Таким образом, категория фильтрации (Filter Methods) включает в себя окрестные и масштабо-разделительные методы (Neighborhood Methods, Scale Separation Methods), а категория перемещения (Displacement Methods) состоит из суммарных мер близости и пространственного выравнивания (Distance-based and Spatial-Alignment Summary Measures), деформации полей (Field deformation), объектно-ориентированных методов (Features-based methods) и мер геометрической характеристики (Geometrical characterization measures).

Несомненным достоинством комплекса является наличие функций для оценки статистической значимости результатов верификации прогноза полей (field significance) с помощью процедур перевыборок с повтором (блочный бутстреп) и полу-параметрический монте-карловский метод Ливзи – Чена. Однако эти функции требуют организации полей во временной последовательности. Помимо статистических характеристик пространственной верификации имеется возможность расчета некоторых стандартных показатели качества, например, градиентной оценки S1 и коэффициента корреляции аномалий.

Работа выполнена в рамках исследований по плану ЦНТП Росгидромета НИТР 1.1.4 и 1.1.5.

Заключение

Традиционные методы верификации современных моделей ЧПП высокого разрешения рекомендуется дополнять новыми пространственными методами. В статье приведена классификация пространственных методов и описание основных из них. Наиболее часто в мировой практике применяются окрестные методы в комбинации с объектно-ориентированными, в том числе для оценки ансамблевого прогноза.

Список литературы

1. *Зарипов Р.Б., Павлюков Ю.Б., Шумилин А.А., Травов А.В.* Использование радиолокационной информации при оценке прогноза погоды с высоким разрешением // Гидрометеорологические исследования и прогнозы. 2018. № 2 (368). С. 60-86.
2. *Калинин Н.А., Шихов А.Н., Быков А.В., Тарасов А.В.* Анализ результатов численного прогноза ливневых осадков по модели WRF с применением различных параметризаций конвекции (на примере территории Пермского края) // Гидрометеорологические исследования и прогнозы. 2019. № 3 (373). С. 43-60.

3. Кисельникова В.З. Объектно-ориентированная оценка качества прогноза осадков WRF // Метеорология и гидрология. 2013. № 4. С. 5-10.

4. Кисельникова В.З. Результаты объектно-ориентированной оценки качества прогнозов осадков по мезомасштабной модели Гидрометцентра России // Труды Гидрометцентра России. 2014. Вып. 351. С. 34-42.

5. Кисельникова В.З. Объектно-ориентированная оценка качества прогноза осадков для теплого периода (май–сентябрь) 2016–2020 гг. по модели COSMO-Ru2 // Гидрометеорологические исследования и прогнозы. 2021. № 2 (380). С. 43-51.

6. Муравьев А.В., Киктёв Д.Б., Смирнов А.В. Оперативная технология наукастинга осадков на основе радарных данных и результаты верификации для теплого периода года (май–сентябрь 2017 г.) // Гидрометеорологические исследования и прогнозы. 2018. № 1 (367). С. 6-38.

7. Ahijevych, D., Gilleland E., Brown B.G., Ebert E.E. Application of spatial verification methods to idealized and NWP-gridded precipitation forecasts // Wea. Forecasting. 2009. Vol. 24. P. 1485-1497:

8. Brown B.G., Bullock R., David C.A., Gotway J.H., Chapman M.B., Takacs A., Gilleland E., Manning K., Mahoney J. New verification approaches for convective weather forecasts // 11th Conf. Aviation, Range, and Aerospace Meteorology, Amer. Met. Soc., 4-8 Oct 2004, Hyannis, MA.

9. Bundel A., Gofa F. et al. COSMO Priority Project INSPECT. Final Report. January 2019. 67 p. DOI: 10.5676/DWD pub/nwv/cosmo-tr37

10. Cafaro C., Woodhams B.J., Stein T.H.M., Birch C.E., Webste, S., Bain C.L., Hartley A., Clarke S., Ferrett S., Hill P. Do Convection-Permitting Ensembles Lead to More Skillful Short-Range Probabilistic Rainfall Forecasts over Tropical East Africa? // Wea. Forecast. 2021. Vol. 36, is. 2. P. 697-716. DOI: 10.1175/WAF-D-20-0172.1

11. Casati B., Ross G., Stephenson D.B. A new intensity-scale verification approach for the verification of spatial precipitation forecasts // Meteorol Appl. 2004. Vol. 11. P. 141-154.

12. Davis C.A., Brown B.G., Bullock R., Halley-Gotway J. The Method for Object-Based Diagnostic Evaluation (MODE) Applied to Numerical Forecasts from the 2005 NSSL/SPC Spring Program DOI // Wea. Forecast. 2009. Vol. 24, is. 5. <https://doi.org/10.1175/2009WAF2222241.1>

13. Davis C., Brown B., Bullock R. Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. // Mon. Wea. Rev. 2006. Vol. 134. P. 1772-1784.

14. Dey S.R., Leoncini G., Roberts N.M., Plant R.S., Migliorini S. A spatial view of ensemble spread in convection permitting ensembles // Mon. Wea. Rev. 2014. Vol. 142. P. 4091-4107, <https://doi.org/10.1175/MWR-D-14-00172.1>.

15. Dorninger M., Mittermaier M.P., Gilleland E., Ebert E.E., Brown B.G., Wilson L.J. MesoVICT: Mesoscale Verification Inter-Comparison over Complex Terrain // NCAR Technical Note NCAR/TN-505+STR, 2013. 23 p. DOI:10.5065/D6416V21. <http://nldr.library.ucar.edu/repository/collections/TECH-NOTE-000-000-000-874>

16. Dorninger M., Gilleland E., Casati B., Mittermaier M.P., Ebert E.E., Brown B.G., Wilson L.J. The setup of the MesoVICT project // Bull. Amer. Meteor. Soc. 2018. Vol. 99. P. 1887-1906.

17. Ebert E., McBride J. Verification of precipitation in weather systems: Determination of systematic errors // J. Hydrol. 2000. Vol. 239. P. 179-202.

18. Ebert E. Fuzzy verification of high resolution gridded forecasts: A review and proposed framework // Meteor. Appl. 2008. Vol. 15. P. 51-64. DOI:10.1002/met.25
19. Gilleland E., Ahijevych D.A., Brown B.G., Ebert E.E. Verifying Forecasts Spatially // Bull. Amer. Meteor. Soc. 2010. Vol. 91. P. 1365-1373. DOI: <http://dx.doi.org/10.1175/2010BAMS2819.1>
20. Gilleland E., Ahijevych D., Brown B.G., Casati B., Ebert E.E. Intercomparison of spatial forecast verification methods // Wea. Forecasting. 2009. Vol. 24. P. 1416-1430.
21. Johnson and Wang. Verification and Calibration of Neighborhood and Object-Based Probabilistic Precipitation Forecasts from a Multimodel Convection-Allowing Ensemble // Mon. Wea. Rev. 2012. Vol. 140. P. 3054-3077.
22. Johnson A., Wang X., Wang Y., Reinhart A., Clark A.J., Jirak I.L. Neighborhood- and Object-Based Probabilistic Verification of the OU MAP Ensemble Forecasts during 2017 and 2018 Hazardous Weather Testbeds // Wea. Forecast. Vol. 2020. Vol. 35. P. 169-191. <https://doi.org/10.1175/WAF-D-19-0060.1>
23. Jolliffe Ian T., Stephenson D.B. Forecast Verification: A Practitioner's Guide in Atmospheric Science: 2nd Edition. John Wiley & Sons Ltd, 2012, 292 p.
24. Murphy A.H. The Finley affair: a signal event in the history of forecast verification // Wea. Forecast. 1996. Vol. 11. P. 3-20.
25. Roberts N.M., Lean H.W. Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events // Mon. Wea. Rev. 2008. Vol. 136. P. 78-97. <https://doi.org/10.1175/2007MWR2123.1>
26. Schwartz C.S., Sobash R.A. Generating Probabilistic Forecasts from Convection-Allowing Ensembles Using Neighborhood Approaches: A Review and Recommendations // Mon. Wea. Rev. 2017. Vol. 135. P. 3397-3418. DOI: <https://doi.org/10.1175/MWR-D-16-0400.1>
27. Schwartz C.S., Kain J.S., Weiss S.J., Xue M., Bright D.R., Kong F., Thomas K.W., Levit J.J., Coniglio M.C., Wandishin M.S. Toward Improved Convection-Allowing Ensembles: Model Physics Sensitivities and Optimizing Probabilistic Guidance with Small Ensemble Membership // Wea. Forecast. 2010. Vol. 25. P. 263-280. <https://doi.org/10.1175/2009WAF2222267.1>
28. Skok G., Roberts N. Estimating the displacement in precipitation forecasts using the Fractions Skill Score // Quarterly Journal of the Royal Meteorological Society. 2018. Vol. 144. P. 414-425. <https://doi.org/https://doi.org/10.1002/qj.3212>
29. Skok G. A new spatial displacement measure for continuous fields, IVMW2020, 2020. https://www.univie.ac.at/img-wien/jwgfvr/SpatialMethods/2020111615/Skok_2020IVMWO_SpatialMethods_16Nov15UTC.pdf,
30. Theis S. E., Hense A., Damrath U. Probabilistic precipitation forecasts from a deterministic model: A pragmatic approach // Meteor. Appl. 2005. Vol. 12. P. 257-268. <https://doi.org/10.1017/S1350482705001763>

References

1. Zaripov R.B., Pavlyukov Yu.B., Shumilin A.A., Travov A.V. Application of radar information for verification of the high-resolution numerical weather prediction. *Gidrometeorologicheskie issledovaniya i prognozy* [Hydrometeorological Research and Forecasting], 2018, vol. 368, no. 2, pp. 60-86 [in Russ.].
2. Kalinin N.A., Shikhov A.N., Bykov A.V., Tarasov A.V. Analysis of numerical forecasts of heavy precipitation with the WRF model using various convection schemes

(a case study for the Perm Krai). *Gidrometeorologicheskie issledovaniya i prognozy* [Hydrometeorological Research and Forecasting], 2019, vol. 373, no. 3, pp. 43-60 [in Russ.].

3. Kisel'nikova V.Z. Object-based evaluation of precipitation forecast quality. *Russ. Meteorol. Hydrol.*, 2013, vol. 38, no. 4, pp. 217-221. DOI: 10.3103/S1068373913040018.

4. Kisel'nikova V.Z. The results of object-oriented estimation of precipitation forecast quality using the mesoscale model of the Hydrometcentre of Russia. *Trudy Gidromettsentra Rossii* [Proceedings of the Hydrometcentre of Russia], 2014, vol. 351, pp. 34-42 [in Russ.].

5. Kisel'nikova V.Z. Object-oriented assessment of COSMO-Ru2 precipitation forecast quality for the warm season (May-September) in 2016-2020. *Gidrometeorologicheskie issledovaniya i prognozy* [Hydrometeorological Research and Forecasting], 2021, vol. 380, no. 2, pp. 43-51 [in Russ.].

6. Muravev A.V., Kiktev D.B., Smirnov A.V. Operational precipitation nowcasting system based on radar data and verification results for the warm period of the year (May-September 2017). *Gidrometeorologicheskie issledovaniya i prognozy* [Hydrometeorological Research and Forecasting], 2018, vol. 367, no. 1, pp. 6-38 [in Russ.].

7. Ahijevych, D., Gilleland E., Brown B.G., Ebert E.E. Application of spatial verification methods to idealized and NWP-gridded precipitation forecasts. *Wea. Forecast.* 2009, vol. 24, pp. 1485-1497. DOI: 10.1175/2009WAF2222298.1.

8. Brown B.G., Bullock R., David C.A., Gotway J.H., Chapman M.B., Takacs A., Gilleland E., Manning K., Mahoney J. New verification approaches for convective weather forecasts. 11th Conf. Aviation, Range, and Aerospace Meteorology, Amer. Met. Soc., 4-8 Oct 2004, Hyannis, MA.

9. Bundel A., Gofa F. et al. COSMO Priority Project INSPECT. Final Report. 2019, 67 p. DOI: 10.5676/DWD pub/nwv/cosmo-tr37.

10. Cafaro C., Woodhams B.J., Stein T.H.M., Birch C.E., Webste, S., Bain C.L., Hartley A., Clarke S., Ferrett S., Hill P. Do Convection-Permitting Ensembles Lead to More Skillful Short-Range Probabilistic Rainfall Forecasts over Tropical East Africa? *Wea Forecast.* 2021, vol. 36, is. 2, pp. 697-716. DOI: 10.1175/WAF-D-20-0172.1

11. Casati B., Ross G., Stephenson D.B. A new intensity-scale verification approach for the verification of spatial precipitation forecasts. *Meteorol Appl.*, 2004, vol. 11, pp. 141-154.

12. Davis C.A., Brown B.G., Bullock R., Halley-Gotway J. The Method for Object-Based Diagnostic Evaluation (MODE) Applied to Numerical Forecasts from the 2005 NSSL/SPC Spring Program. DOI: *Wea. Forecast.*, 2009, vol. 24, no. 5. DOI: 10.1175/2009WAF2222241.1.

13. Davis C., Brown B., Bullock R. Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, 2006, vol. 134, pp. 1772-1784.

14. Dey S.R., Leoncini G., Roberts N.M., Plant R.S., Migliorini S. A spatial view of ensemble spread in convection permitting ensembles. *Mon. Wea. Rev.*, 2014, vol. 142, p. 4091-4107, <https://doi.org/10.1175/MWR-D-14-00172.1>.

15. Dorninger M., Mittermaier M.P., Gilleland E., Ebert E.E., Brown B.G., Wilson L.J. MesoVICT: Mesoscale Verification Inter-Comparison over Complex Terrain. *NCAR Technical Note NCAR/TN-505+STR*, 2013, 23 p. DOI: 10.5065/D6416V21.

16. Dorninger M., Gilleland E., Casati B., Mittermaier M.P., Ebert E.E., Brown B.G., Wilson L.J. The setup of the MesoVICT project. *Bull. Amer. Meteor. Soc.*, 2018, vol. 99, pp. 1887-1906.

17. Ebert E., McBride J. Verification of precipitation in weather systems: Determination of systematic errors. *J. Hydrol.*, 2000, vol. 239, pp. 179-202.
18. Ebert E. Fuzzy verification of high resolution gridded forecasts: A review and proposed framework. *Meteor. Appl.*, 2008, vol. 15, pp 51-64. DOI:10.1002/met.25.
19. Gilleland E., Ahijevych D.A., Brown B.G., Ebert E.E. Verifying Forecasts Spatially. *Bull. Amer. Meteor. Soc.*, 2010, vol. 91, pp. 1365-1373. DOI: 10.1175/2010BAMS2819.1.
20. Gilleland E., Ahijevych D., Brown B.G., Casati B., Ebert E.E. Intercomparison of spatial forecast verification methods. *Wea. Forecast.*, 2009, vol. 24, pp. 1416-1430.
21. Johnson and Wang. Verification and Calibration of Neighborhood and Object-Based Probabilistic Precipitation Forecasts from a Multimodel Convection-Allowing Ensemble. *Mon. Wea. Rev.*, 2012, vol. 140, pp. 3054-3077.
22. Johnson A., Wang X., Wang Y., Reinhart A., Clark A.J., Jirak I.L. Neighborhood- and Object-Based Probabilistic Verification of the OU MAP Ensemble Forecasts during 2017 and 2018 Hazardous Weather Testbeds. *Wea. Forecast.*, 2020, vol. 35, pp. 169-191. DOI: 10.1175/WAF-D-19-0060.1.
23. Jolliffe Ian T., Stephenson D.B. Forecast Verification: A Practitioner's Guide in Atmospheric Science: 2nd Edition. John Wiley & Sons Ltd, 2012, 292 p.
24. Murphy A.H. The Finley affair: a signal event in the history of forecast verification. *Wea. Forecast.*, 1996, vol. 11, pp. 3-20.
25. Roberts N.M., Lean H.W. Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Mon. Wea. Rev.*, 2008, vol. 136, pp. 78-97. DOI: 10.1175/2007MWR2123.1.
26. Schwartz C.S., Sobash R.A. Generating Probabilistic Forecasts from Convection-Allowing Ensembles Using Neighborhood Approaches: A Review and Recommendations. *Mon. Wea. Rev.*, 2017, vol. 135, pp. 3397-3418. DOI: 10.1175/MWR-D-16-0400.1.
27. Schwartz C.S., Kain J.S., Weiss S.J., Xue M., Bright D.R., Kong F., Thomas K.W., Levit J.J., Coniglio M.C., Wandishin M.S. Toward Improved Convection-Allowing Ensembles: Model Physics Sensitivities and Optimizing Probabilistic Guidance with Small Ensemble Membership. *Wea. Forecast.*, 2010, vol. 25, pp. 263-280. DOI: 10.1175/2009WAF2222267.1.
28. Skok G., Roberts N. Estimating the displacement in precipitation forecasts using the Fractions Skill Score. *Quarterly Journal of the Royal Meteorological Society*, 2018, vol. 144, p. 414-425. <https://doi.org/> <https://doi.org/10.1002/qj.3212>
29. Skok G. A new spatial displacement measure for continuous fields, *IVMW2020*, 2020. https://www.univie.ac.at/img-wien/jwgfvr/SpatialMethods/2020111615/Skok_2020IVMWO_SpatialMethods_16Nov15UTC.pdf
30. Theis S. E., Hense A., Damrath U. Probabilistic precipitation forecasts from a deterministic model: A pragmatic approach. *Meteor. Appl.*, 2005, vol. 12, p. 257-268. <https://doi.org/10.1017/S1350482705001763>

Поступила 15.11.2021; одобрена после рецензирования 30.11.2021;
принята в печать 13.12.2021.

Submitted 15.11.2021; approved after reviewing 30.11.2021;
accepted for publication 13.12.2021.