

DOI: <https://doi.org/10.37162/2618-9631-2025-4-7-32>

УДК 551.509.313+551.509.324.2+551.508.85

Наукастинг порывов ветра на основе численных прогнозов, радиолокации и машинного обучения: реализация, возможности и ограничения

А.В. Смирнов, Д.Б. Киктев, А.В. Муравьев

*Гидрометеорологический научно-исследовательский центр
Российской Федерации, г. Москва, Россия
muravev@mecom.ru*

Изложены результаты испытаний системы наукастинга порывов приземного ветра, проведенных в мае–сентябре 2024 года в Гидрометцентре России с применением алгоритма машинного обучения "Случайный лес" к выходной продукции статистической модели наукастинга ruSTEPS и системы численного прогноза погоды COSMO-Ru2.2. Оценена значимость наблюдений автоматических метеостанций Центрального федерального округа в качестве контрольных данных для наукастинга порывов ветра. Выявлены некоторые систематические особенности численных прогнозов порывов в период испытаний. Обсуждаются проблемы синхронизации данных от различных источников, детально анализируется погодная ситуация с заметными порывами на территории Московской области, приводятся оценки качества с помощью показателя FSS. В разработанной версии системы наукастинга, на накопленной информационной базе и по значениям показателя FSS горизонт полезного прогноза порывов ветра ограничен 30 минутами.

Ключевые слова: наукастинг порывов ветра, численный прогноз погоды, радиолокационные наблюдения, машинное обучение

Wind gust nowcasting using numerical forecasts, radar technologies and machine learning: implementation, capabilities and limitations

A.V. Smirnov, D.B. Kiktev, A.V. Muravev

*Hydrometeorological Research Center of Russian Federation,
Moscow, Russia
muravev@mecom.ru*

The paper presents the results of testing a surface wind gust nowcasting system conducted in May–September 2024 at the Hydrometcentre of Russia using the machine learning Random Forest algorithm applied to the output products of the pySTEPS statistical nowcasting model and the COSMO-Ru2.2 numerical weather prediction system. The significance of observations from automatic weather stations in the Central Federal District as control data for wind gust nowcasting is assessed. Some systematic features of numerical gust forecasts during the testing period are revealed. The problems of synchronizing data from various sources are discussed, the weather conditions with significant gusts over

a large area are analyzed in detail, and the quality assessment is provided using the FSS metric. In the developed version of the nowcasting system, on the basis of accumulated information and according to the FSS skill score, the useful forecast horizon for wind gusts is limited to 30 minutes.

Keywords: wind gust nowcasting, numerical weather prediction, radar observations, machine learning

Введение

В информационном сообщении [2] приведены основные результаты испытаний системы наукастинга *порывов приземного ветра* с применением методов машинного обучения к выходной продукции статистической модели наукастинга ruSTEPS и системы численного прогноза погоды COSMO-Ru2.2 (ЧПП). Испытания проведены в мае–сентябре 2024 года с использованием алгоритма машинного обучения "Случайный лес". В связи с разработкой системы наукастинга порывов составлен краткий обзор передовых тенденций в данной области [6]. В настоящей статье более подробно обсуждаются особенности накопленных для испытаний архивов, ограничения и возможности наукастинга порывов ветра в отечественных условиях с учетом физико-статистических особенностей этой величины и имеющейся наблюдательной базы.

Опишем состав архива данных по периоду 1 мая – 23 сентября (последняя неделя сентября не включена по объективным причинам). Количественные характеристики округляются до десятков. Для краткости будет называть наборы данных условно результирующими, обучающими и контрольными.

1. *Результирующие* прогностические поля порывов (наукасты) в области обзоров радиолокаторов ДМРЛ-С сроком до 150 мин в 10-минутных интервалах (144 прогноза ежесуточно). Шаг расчетной ячейки ~1 км, размер расчетных матриц данных 987×987. Общее количество полей для одной заблаговременности около 20600.

2. *Обучающие* численные прогнозы порывов (мезопргнозы) сроком на 720 мин через 3 часа (8 прогнозов ежесуточно в 10-минутной дискретности). Шаг расчетной ячейки 2.2 км, размер матриц данных 601×601. Общее количество полей для одной заблаговременности около 1150.

3. *Контрольные* наблюдения около 180 автоматических метеорологических станций (АМС) в 10-минутных интервалах (144 телеграммы ежесуточно). Общее количество телеграмм от сети АМС около 21000.

Географическая область наукастинга порывов представляет собой сферический прямоугольник в зоне обзоров радиолокаторов ДМРЛ-С, развернутых на территории ЦФО России (рис. 1). Довольно частое отсутствие данных от каких-либо локаторов маркируется в расчетной матрице ("композитном поле") константами отсутствия в соответствующих точках выходного наукаста.

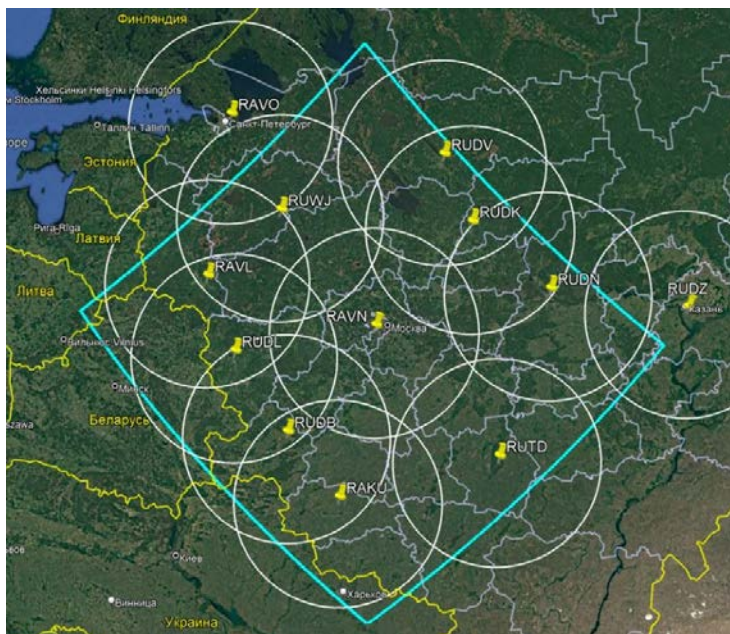


Рис. 1. Зоны обзора радиолокаторов и расчетная область наукастинга скорости порывов ветра.

Fig. 1. Radar coverage areas and computation area for wind gust speed nowcasting.

1. Реализация технологии наукастинга порывов приземного ветра

Сложность прогноза шквалистых усилений ветра начинается со сложности их диагноза по имеющимся наблюдениям. Эти явления часто не удается зарегистрировать надлежащим образом средствами контактных измерений – расстояние между станциями велико в сравнении с характерными размерами такого рода явлений.

Для радарных наблюдений нижний пятидесятиметровый слой над земной поверхностью, где происходят максимальные шквалистые усиления ветра, обычно находится ниже зоны прямой видимости радара и для косвенной оценки приземного ветра требуется привлекать гипотезы о его связи с непосредственно измеряемыми характеристиками.

Анализ ситуаций по 10-минутным измерениям автоматических метеорологических станций (АМС) и радарным данным показывает, что фиксируемые на станциях порывистые скорости ветра могут сильно различаться при близких значениях в их окрестности таких влияющих параметров, как высота верхней границы облачности (ВГО), отражаемость, влагосодержание и др.

Такие ситуации возникают, когда активная конвекция, идентифицируемая максимальной отражаемостью выше 50 dBZ и высотой ВГО более

12 км, не сопровождается заметными порывами ветра на ближайших АМС. Среди возможных объяснений такого несовпадения могут быть:

– особенности структуры конвективного облака, где области повышенного радиоэха и области максимальных шквалистых усилений скорости ветра пространственно разнесены (максимальные усиления ветра чаще возникают на фронтах порывистости перед конвективным облаком, которые при этом могут опережать передний край зоны ливневых осадков на 10–15 км);

– недостаточный учет в диагностических алгоритмах тонких эффектов, связанных со структурой конвективного облака и его окружения, как в атмосфере, так и на подстилающей поверхности (например, увлажненность, альbedo и др. могут оказывать существенное влияние на энергетику конвекции);

– погрешности пространственно-временной привязки радарных данных к станционным (при сканировании каждые 10 минут и при скорости потока около 30 км/ч перенос за 5-минутный интервал составляет 2–3 км, вследствие чего фиксируемые на станциях явления могут оказаться разнесенными с радарными наблюдениями по соседним 10-минутным срокам).

Существенные пробелы в фиксации шквалистых усилений ветра наземными измерениями затрудняют настройку диагностических алгоритмов по данным дистанционного зондирования, что сказывается на качестве верификации прогнозов этих явлений, способствуя в первую очередь обилию ложных тревог в прогностической продукции.

В решающих правилах программного обеспечения вторичной обработки информации ДМРЛ-С [7] шквалы диагностируются в том случае, когда отражаемость на уровне изотермы -15°C , максимальная отражаемость в столбе и высота ВГО относительно уровня тропопаузы превышают предписанные пороговые значения. Эти правила не являются универсальными для охвата многообразных и физически сложных ситуаций образования шквалов, что также отражается на качестве их детектирования. Разработчики "ГИМЕТ-2010" считают необходимым к единственному фактору интенсивности конвекции, учитываемому в текущей версии диагностики шквалов, добавить «динамический», т. е. адвективный фактор [3, 7].

2. Машинное обучение: алгоритм "Случайный лес"

Алгоритм машинного обучения "Случайный лес" (использован пакет randomForestSRC из библиотек языка R) применен к выходным данным ЧПП, которые имеются одновременно и в продукции радиолокационного зондирования (ВГО, интегральное влагосодержание облаков, отражаемость, скорость и направления ветра на различных уровнях), а также к высотам изотерм и некоторым другим потенциальным предикторам шквалистых усилений ветра. Данный алгоритм зарекомендовал себя как простой, надежный и гарантированный от переобучения метод, чем он выгодно отличается, в частности, от нейросетевых моделей, которые при довольно

сложной настройке требуют и значительных вычислительных ресурсов для поддержания режима частого обновления прогнозов.

В предположении наличия связи этих предикторов со скоростью порывов формируется обучающая выборка для диагноза шквалов по данным ЧПП и строятся решающие диагностические процедуры для последующего использования построенных связей в радарном диагнозе и наукастинге.

Перенос связей в область радарного диагноза и прогноза обусловлен тем, что мультипликативная каскадная модель наукастинга ruSTEPS способна воспроизводить структуру и временную эволюцию не только полей осадков, но и полей упомянутых выше потенциальных предикторов. Следует, однако, учитывать, что оперативная модель COSMO-Ru2.2 в целом достаточно реалистично прогнозирует скорость порывов ветра у земли, но довольно грубо воспроизводит сложные процессы глубокой конвекции, ввиду чего перенос построенных машинным обучением диагностических связей на данные радарного наукастинга сопряжен с определенным риском и должен внимательно контролироваться и перепроверяться. Как будет показано ниже, такие опасения вполне обоснованы.

Для построения решающих правил используется информация последнего доступного оперативного ЧПП модели COSMO-Ru2.2, представленного в 10-минутной временной дискретности внутри периода прогноза. Расчет ЧПП на срок до 12 часов проводится восемь раз в сутки и через 2.5–3 часа после стартового срока результаты расчетов становятся доступными для технологии наукастинга.

В качестве предикторов скорости приземных порывов ветра из данных ЧПП выбираются следующие параметры:

- максимальная отражаемость,
- высота верхней границы облачности,
- высота нулевой изотермы,
- средняя скорость ветра на высоте 1 км,
- средняя скорость ветра на высоте 3 км,
- средняя скорость ветра на 10 м.

В обучающую выборку включаются все попадающие на временной интервал наукастинга 10-минутные модельные поля скорости порывов и шести предикторов из *последнего доступного* численного прогноза погоды. Перед процедурой машинного обучения модельные сеточные значения предикторов и предиктанта преобразовываются в значения, агрегированные по окрестным узлам сетки. Размеры окна агрегации и функции преобразования в окне сеточных значений подобраны в результате предварительных экспериментов (табл. 1).

По алгоритму "Случайный лес" строится ансамбль множества решающих деревьев (рис. 2), а итоговый прогноз "обобщает" ответы множества деревьев. Тренировка деревьев происходит независимо друг от друга на разных подмножествах, что препятствует построению одинаковых деревьев на одной выборке и делает этот алгоритм удобным для распределённых вычислений.

Таблица 1. Перечень обучающих параметров, размеры используемых окон агрегации и способы преобразования окрестных сеточных значений

Table 1. List of training parameters, sizes of aggregation windows used and methods for transforming neighboring grid values

| Параметр | Единицы | Размер окна (узлы сетки) | Функция |
|---------------------------------------|---------|--------------------------|---------|
| Максимальная отражаемость | dBZ | 11×11 | max |
| Высота верхней границы облачности | м | 15×15 | max |
| Высота нулевой изотермы | м | 11×11 | mean |
| Средняя скорость ветра на высоте 1 км | м/с | 11×11 | mean |
| Средняя скорость ветра на высоте 3 км | м/с | 11×11 | mean |
| Средняя скорость ветра на высоте 10 м | м/с | 25×25 | mean |
| Скорость приземных порывов ветра | м/с | 5×5 | mean |

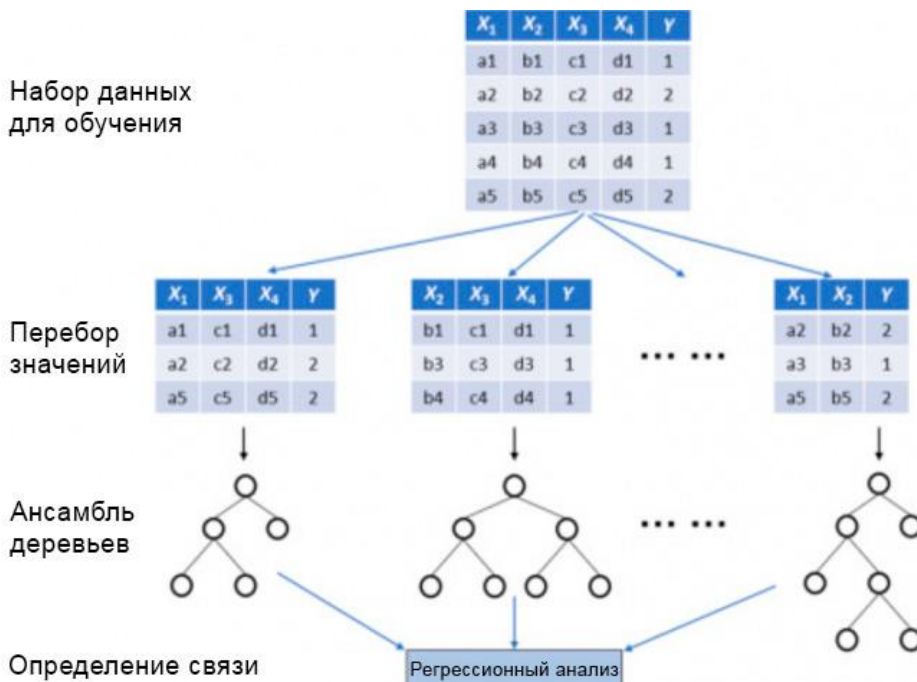


Рис. 2. Схема работы алгоритма "Случайный лес".

Fig. 2. Operational scheme of the "Random Forest" algorithm.

Обучение выполняется при расчете каждого наукаста (прогностического поля скорости порывов), при этом формируется 200 решающих деревьев регрессионного типа (рис. 3). Качество построения решающих правил оценивается сопоставлением восстановленных значений порывов с порывами, прогнозируемыми самой моделью COSMO-Ru2.2.

В настоящее время ЧПП все еще уступает простой экстраполяции текущей погоды ("лагранжевому" наукастингу) в локализации мезо-метеорологических объектов и структур на раннем этапе прогноза. Перенос в область радарного наукастинга состоит в том, что часть модельных предикторов заменяется результатами пространственно-временной экстраполяции полей максимальной отражаемости и высоты верхней границы облачности с помощью модели ruSTEPS (рис. 3).

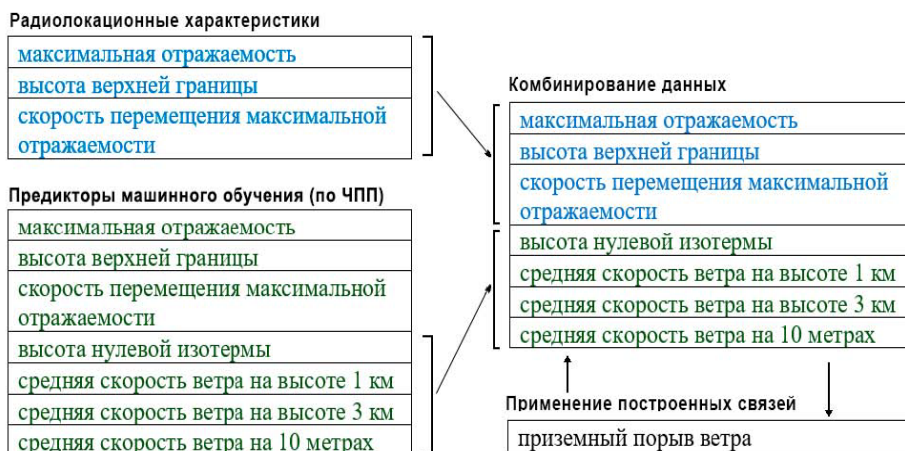


Рис. 3. Схема комбинирования полей предикторов при расчетах наукастинга скорости порывов.

Fig. 3. Combining scheme of predictor fields in calculations of gust speed nowcasting.

Общая блок-схема алгоритма наукастинга скорости порывов приземного ветра показана на рис. 4. Серым цветом обозначены блоки с возможностью распараллеливания вычислительных процессов. В верхней части схемы представлены модули, которые работают независимо друг от друга. Цепочка последовательных действий выполняется при расчете каждого наукаста. Распараллеливание сокращает время работы системы наукастинга в несколько раз, в результате чего расчет одного прогноза на 2.5 часа укладывается в 5–6 минут, что с запасом обеспечивает возможность обновления прогноза каждые 10 минут.

3. Описание, преобразование и сопоставление типов данных

Использованные типы прогностической и фактической информации различаются как по форматам, так и по значениям порывов ветра. Некоторые особенности использованных типов информации и возникающие при этом проблемы хорошо известны. Например, это точечный характер станционных наблюдений и двумерное представление прогнозов в виде полей,

причем поля ЧПП и наукастинга имеют различное разрешение. Разнообразие объемов выборок разных типов информации может критично отражаться на статистической значимости сравнительных оценок в парах сопоставляемых данных. Далее, большие объемы данных при испытании схем наукастинга по значительным периодам времени (с заблаговременностями до нескольких часов в 10-минутной дискретности) также способны затруднить их обработку и верификацию, вынуждая искать способы сокращения табличной и графической продукции без потери содержательной информации. Наконец, остаются проблемы отбора подходящих статистических характеристик для описания и сопоставления разнородных данных, а также для верификации качества прогнозов как в точках, так и в пространственных полях.



Рис. 4. Блок-схема алгоритма наукастинга скорости порывов приземного ветра.

Fig. 4. Flow chart of the nowcasting algorithm for surface wind gust speed.

Остановимся на проблеме "больших данных". Десятки тысяч прогностических полей (по миллиону точек каждое) со значениями скорости ветровых порывов удастся заметно "ужать" выделением порывов ветра в принятых градациях "неблагоприятного" (НЯ) и "опасного" явлений (ОЯ) [8]. При таком подходе количество "полей с порывами" уменьшается в несколько раз, а сами поля превращаются в удобные для обработки бинарные

(дихотомические) массивы, состоящие только из единиц и нулей. В проведенных испытаниях рассмотрены пороговые значения скорости ветра 15 м/с как "неблагоприятного метеорологического явления", в интервале 15–24 м/с как "сильного ветра" и в ранге ОЯ при скорости не менее 25 м/с как "очень сильного ветра". Для иллюстрации качества наукастинга приводятся оценки порывов скорости не менее 20 м/с. *Контрольные* стационарные наблюдения - ввиду их редкости и пространственной разреженности – градуировались порогами 15 и 20 м/с.

Ниже приняты следующие обозначения. Прогнозы и наблюдения производятся в 10-минутные сроки YYYYMMDDhhmi, где год YYYY = 2024, месяцы MM = {05, 06, 07, 08, 09}, дни DD = {01, 02, ..., 31}, часы hh = {00, 01, ..., 23}, десятки минут mi = {00, 10, 20, ...}. Продукция модели COSMO-Ru2.2 относится к срокам hh = {00, 03, 06, 09, 12, 15, 18, 21}. Подчеркнем, что всюду ниже, где указано реальное время, имеется в виду ВСВ.

3.1. Поля наукастинга в области обзоров сети ДМРЛ-С

Поле наукастинга (FCAST) составлено из кругов радиолокационных обзоров с возможными пересечениями и содержит точки с константами отсутствия (NA) при отсутствии данных с каких-либо локаторов. В период испытаний количество отсутствующих значений в композитном поле достигало до трети общего размера поля в 987×987 точек. Заметим, что учет и обработка констант отсутствия в прогностических полях представляет собой непростую проблему, которая зависит от множества факторов и не имеет универсального решения [4].

Поля наукастинга и мезопрогнозов расположены в повернутых координатах модели COSMO с Москвой в центральной точке массива данных (рис. 5). Окаймляющий шестиугольник с вершинами в координатах некоторых городов построен для более наглядной иллюстрации поворота карты относительно стран света.

Рассмотрим дихотомические поля наукастинга со значениями порывов не меньше 20 м/с и представим квантильное распределение единиц по всему полю через 10 %. Для всех 15 заблаговременностей составлены таблицы, имеющие вид табл. 2.

Приведены части таблицы, характеризующие особенности этого распределения: видно, что количество точек с сильными порывами расположено в самом конце распределения и может изменяться в широких диапазонах (здесь от 3 до 10000). Отметим, что по таблицам такого вида удобно выделять развивающиеся ситуации, в которых количество точек со значительными скоростями порывов занимают постепенно растущие и постепенно уменьшающиеся области полей в последних квантилях распределения.

В табл. 3 отражена ситуация 20 июня около полудня с наукастами порывов до 28 м/с и с охватом до 17.5 тысяч точек поля (около 2 % общей области).

Таблица 3. Пространственное распределение значений порывов ветра в поле наукастинга на 10 мин для ситуации с порывами ветра 19 и 20 июня 2024 г.; обозначения как для таблицы 2

Table 3. Spatial distribution of wind gust values in the 10 min nowcasting field for the wind gust situation on 19 and 20 June 2024; notations as for Table 2

| | | | | | | | | | | | | | | | |
|---------------------------|------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|--------|--------|------|-------|
| FCAST_000_202406191330_10 | 2.19 | 3.80 | 4.42 | 5.15 | 5.76 | 6.60 | 8.01 | 8.80 | 9.67 | 10.45 | 20.42 | 956484 | 198693 | 20.8 | 29 |
| FCAST_000_202406200950_10 | 3.39 | 7.32 | 9.01 | 10.45 | 11.31 | 11.98 | 12.68 | 13.30 | 14.00 | 14.91 | 22.19 | 956484 | 208008 | 21.7 | 445 |
| FCAST_000_202406201000_10 | 3.23 | 7.13 | 8.86 | 10.41 | 11.32 | 11.98 | 12.68 | 13.27 | 13.98 | 14.90 | 20.73 | 956484 | 203571 | 21.3 | 82 |
| FCAST_000_202406201010_10 | 3.20 | 7.16 | 8.81 | 10.62 | 11.48 | 12.11 | 12.77 | 13.40 | 14.09 | 14.99 | 22.40 | 956484 | 204813 | 21.4 | 542 |
| FCAST_000_202406201020_10 | 3.27 | 7.33 | 9.15 | 10.82 | 11.59 | 12.23 | 12.90 | 13.49 | 14.22 | 15.12 | 25.40 | 956484 | 202743 | 21.2 | 3628 |
| FCAST_000_202406201030_10 | 3.19 | 7.46 | 9.30 | 10.92 | 11.73 | 12.37 | 13.04 | 13.61 | 14.26 | 15.15 | 24.47 | 956484 | 203328 | 21.3 | 5215 |
| FCAST_000_202406201040_10 | 3.84 | 7.28 | 9.52 | 10.59 | 11.28 | 11.81 | 12.65 | 13.46 | 14.34 | 15.61 | 24.30 | 956484 | 203004 | 21.2 | 5642 |
| FCAST_000_202406201050_10 | 4.09 | 7.44 | 9.53 | 10.56 | 11.22 | 11.90 | 12.80 | 13.59 | 14.57 | 16.47 | 24.55 | 956484 | 201690 | 21.1 | 8717 |
| FCAST_000_202406201100_10 | 4.30 | 7.31 | 9.49 | 10.51 | 11.18 | 11.78 | 12.67 | 13.50 | 14.36 | 16.43 | 25.33 | 956484 | 200574 | 21.0 | 10658 |
| FCAST_000_202406201110_10 | 4.56 | 7.47 | 9.58 | 10.68 | 11.31 | 11.94 | 12.81 | 13.62 | 14.49 | 16.62 | 28.03 | 956484 | 201042 | 21.0 | 12835 |
| FCAST_000_202406201120_10 | 4.44 | 7.60 | 9.60 | 10.90 | 11.47 | 12.13 | 12.84 | 13.68 | 14.56 | 16.84 | 25.83 | 956484 | 200682 | 21.0 | 14574 |
| FCAST_000_202406201130_10 | 4.26 | 7.83 | 9.71 | 11.06 | 11.56 | 12.19 | 12.86 | 13.73 | 14.72 | 17.07 | 25.20 | 956484 | 202212 | 21.1 | 13873 |
| FCAST_000_202406201140_10 | 3.86 | 7.27 | 8.84 | 9.94 | 10.87 | 11.66 | 12.57 | 13.34 | 14.57 | 16.84 | 25.17 | 956484 | 202050 | 21.1 | 14556 |
| FCAST_000_202406201150_10 | 3.56 | 7.34 | 8.88 | 9.94 | 10.85 | 11.67 | 12.56 | 13.35 | 14.57 | 16.98 | 26.18 | 956484 | 203121 | 21.2 | 15330 |
| FCAST_000_202406201200_10 | 4.38 | 7.26 | 8.87 | 9.76 | 10.66 | 11.43 | 12.23 | 13.12 | 14.30 | 16.45 | 25.76 | 956484 | 202626 | 21.2 | 5489 |
| FCAST_000_202406201210_10 | 4.41 | 7.32 | 8.86 | 9.81 | 10.71 | 11.60 | 12.38 | 13.26 | 14.33 | 17.08 | 23.44 | 956484 | 201987 | 21.1 | 6726 |
| FCAST_000_202406201220_10 | 3.79 | 7.33 | 8.88 | 9.86 | 10.81 | 11.63 | 12.42 | 13.31 | 14.42 | 17.36 | 23.69 | 956484 | 202302 | 21.2 | 10739 |
| FCAST_000_202406201230_10 | 3.82 | 7.31 | 8.93 | 9.91 | 10.76 | 11.63 | 12.52 | 13.32 | 14.55 | 17.62 | 23.08 | 956484 | 202482 | 21.2 | 15803 |
| FCAST_000_202406201240_10 | 3.58 | 6.62 | 8.12 | 9.34 | 10.26 | 11.12 | 12.03 | 13.17 | 14.43 | 17.26 | 24.01 | 956484 | 201249 | 21.0 | 17113 |
| FCAST_000_202406201250_10 | 3.56 | 6.54 | 7.91 | 9.10 | 10.06 | 11.07 | 12.01 | 13.07 | 14.35 | 16.99 | 24.43 | 956484 | 201654 | 21.1 | 17496 |
| FCAST_000_202406201300_10 | 3.33 | 6.07 | 7.56 | 8.62 | 9.63 | 10.43 | 11.30 | 12.37 | 13.31 | 14.81 | 23.22 | 956484 | 276291 | 28.9 | 3295 |
| FCAST_000_202406201310_10 | 2.95 | 5.94 | 7.43 | 8.36 | 9.83 | 10.65 | 11.63 | 12.60 | 13.63 | 14.97 | 23.44 | 956484 | 275841 | 28.8 | 4311 |
| FCAST_000_202406201320_10 | 2.96 | 5.90 | 7.38 | 8.52 | 9.88 | 10.61 | 11.59 | 12.47 | 13.56 | 15.16 | 23.22 | 956484 | 274725 | 28.7 | 6070 |
| FCAST_000_202406201330_10 | 2.90 | 5.89 | 7.27 | 8.66 | 10.15 | 10.92 | 11.79 | 12.49 | 13.57 | 14.99 | 23.23 | 956484 | 276696 | 28.9 | 5903 |
| FCAST_000_202406201340_10 | 2.95 | 5.58 | 6.74 | 8.69 | 9.66 | 10.38 | 11.13 | 11.78 | 12.61 | 14.43 | 22.93 | 956484 | 277767 | 29.0 | 5538 |
| FCAST_000_202406201350_10 | 2.86 | 5.70 | 6.62 | 8.64 | 9.60 | 10.28 | 10.98 | 11.64 | 12.49 | 14.46 | 22.28 | 956484 | 277632 | 29.0 | 3241 |
| FCAST_000_202406201410_10 | 1.97 | 5.59 | 6.48 | 8.57 | 9.42 | 10.04 | 10.58 | 11.23 | 12.23 | 14.00 | 22.14 | 956484 | 276138 | 28.9 | 545 |

Рис. 6 с наукастами в географической "развертке" (без населенных пунктов) иллюстрирует табличные данные о ситуации 20 июня 2024 г., указывая координатное расположение областей с порывами не менее 20 м/с (желтый цвет). В центре средней панели рис. 6 заметна область желтого цвета (до 25 м/с) с вкраплением красного, сигнализирующего о порывах до 30 м/с

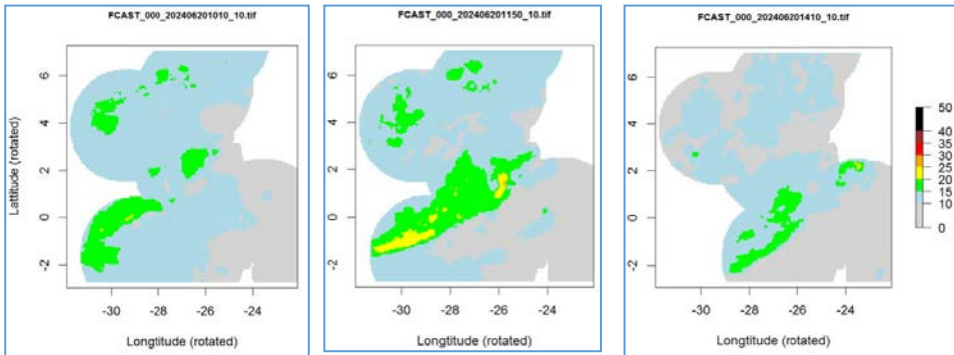


Рис. 6. Развитие ситуации около полудня 20 июня 2024 г. (слева направо) с порывами ветра в картах наукастинга на 10 минут

Fig. 6. Development of the situation around midday on June 20, 2024 (from left to right) with wind gusts in the 10 minutes nowcasting maps.

3.2. Поля прогнозов COSMO-Ru2.2: приведение к формату полей наукастинга

Основные характеристики прогностических полей представлены в выходных параметрах функции `gaster()` языка R, с помощью которой читаются файлы `.tiff` (или `.tif`) (табл. 4).

Таблица 4. Характеристики растровых изображений для наукастов (сверху) и мезопрогнозов (снизу)

Table 4. Raster image characteristics for nowcasts (upper) and mesoforecasts (bottom)

| | |
|--------------------|---|
| class: | RasterLayer |
| dimensions: | 978, 978, 956484 (nrow, ncol, ncell) |
| resolution: | 0.01, 0.01 (x, y) |
| extent: | -31.89, -22.11, -2.73, 7.05 (xmin, xmax, ymin, ymax) |
| values: | 1.098624, 7.956317 (min, max) |
| class: | RasterLayer |
| dimensions: | 601, 601, 361201 (nrow, ncol, ncell) |
| resolution: | 0.02, 0.02 (x, y) |
| extent: | -33.01, -20.99, -4.01, 8.01 (xmin, xmax, ymin, ymax) |
| values: | 0.01708796, 11.14982 (min, max) |

Размеры полей указаны в строке `dimensions`: количество строк, столбцов и точек (ячеек) соответственно. Разрешение сетки в строке `resolution` указывает на приблизительное количество километров на стороне ячейки при соответствии 0.01 ~ 1 км. Ориентация повернутая, в центре карт предполагаются данные по г. Москве. В строке `extent` указаны предельные значения координат по двум осям. Значения величин в строке `values` соответствуют наименьшей и наибольшей скорости порыва в данной карте. Обе части таблицы относятся к одному сроку прогноза. Как видно из таблицы, область мезомодели шире области наукастинга, но пространственное разрешение в два раза меньше.

Для сравнения наукастинга и мезопрогнозов карта мезомодели обрывается до пределов карты наукастинга и разрешение *увеличивается* вдвое методом "ближайшего соседа", при котором значения сильных порывов дублируются. Если для приведения полей к единому формату использовать *уменьшение* размерности поля FCAST, то его простейшая интерполяция эквивалентна отбору точек через одну. Можно предположить, что из-за относительной редкости экстремумов удвоение количества точек с экстремальными величинами незначительно влияет на суммарные оценки и может быть учтено в ходе дальнейших расчетов и интерпретации результатов.

3.3. Архив телеграмм автоматических станций (АМС)

Телеграммы АМС записаны в текстовом формате в виде таблиц (табл. 5), которые считываются функцией `read.table(*, header=TRUE)` языка R с заголовком столбцов, содержащим указания на Юникс-время наблюдения (`dt`) в секундах, индекс станции (`stan`), скорость порыва (`gust`) в м/с, среднюю скорость ветра за 10 минут (`ws`) в м/с, дату и время наблюдения (`dt_str`) в формате `YYYYMMDDhhmi`, географические координаты долготы и широты станции (`lon, lat`) и координаты в повернутой географической сетке мезомодели (`X, Y`).

Таблица 5. Структура телеграмм АМС в файле 202407010000.dat
Table 5. Structure of AMS telegrams in file 202407010000.dat

| dt | stan | gust | ws | dt_str | lon | lat | x | y |
|------------|-------|------|-----|--------------|--------|--------|---------|-------|
| 1719792000 | 19027 | 1.1 | 0.9 | 202407010000 | 37.775 | 55.508 | -26.607 | 1.945 |
| 1719792000 | 32466 | 1.1 | 0.2 | 202407010000 | 37.213 | 55.592 | -26.767 | 2.232 |
| ... | | | | | | | | |

Количество станций не превышает двухсот и может заметно меняться в зависимости от срока наблюдения. Телеграммы накапливаются в течение суток (YYYYMMDD), в то время как наблюдения привязаны к стандартному сроку в формате YYYYMMDDhhmi, ввиду чего файл телеграмм содержит порядка 25 тысяч строк. На рис. 7 демонстрируется расположение 196 станций, содержащихся в файле 20240625.dat. Для удобства обработки данных суточные данные АМС "рассыпаются" на 144 текстовых файла с наблюдениями всех отметившихся наблюдениями станций в стандартный суточный срок.

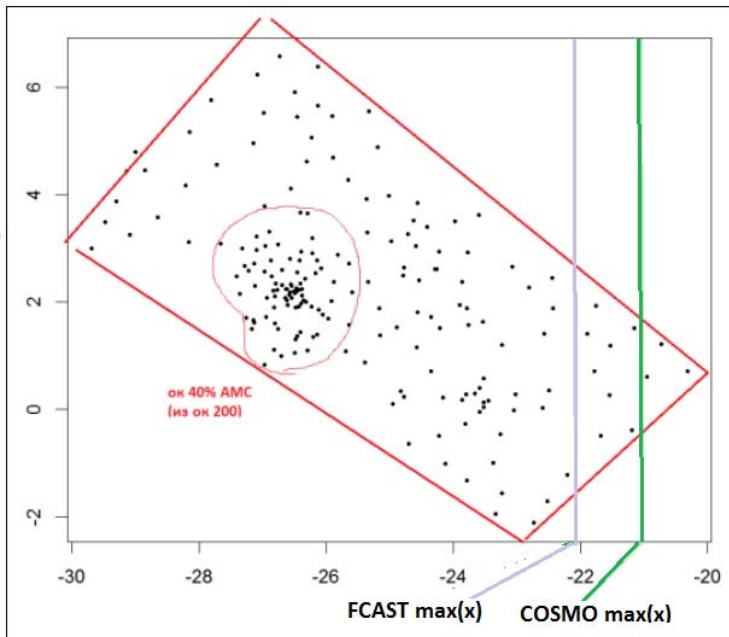


Рис. 7. Распределение АМС, приславших телеграммы 25 июня 2024 г., и окаймляющий четырехугольник (красный цвет). Сиреневой и зеленой вертикальными линиями выделены максимальные значения оси x зон наукастинга и мезопрогноза соответственно.

Fig. 7. Distribution of AMS that sent telegrams on June 25, 2024, and the bordering quadrangle (red). The purple and green vertical lines highlight the maximum values of the x-axis of the nowcasting and mesoforecast zones, respectively.

Видно, что несколько станций не входят ни в одну область прогноза, при этом некоторые станции размещены только в зоне мезопрогноза и отсутствуют в полях наукастинга. Примерно 40 % всех станций расположены в центральной части карты, в основном Москва и Московская область.

Известно, что определение и фиксация порывов ветра производится вместе с определением и фиксацией средней скорости за выделенный период времени (здесь интервал осреднения равен 10 мин). Взаимосвязь этих двух величин хорошо изучена и используется в разнообразных схемах прогноза порывов, при этом средняя скорость является важнейшим предиктором в статистических схемах и непременным элементом параметризации порывов в гидродинамических ЧПП.

Продемонстрируем распределение повторяемостей средней скорости ветра и величины порывов, зафиксированных на 196 станциях 25 июня 2024 г. (рис. 8). Несмотря на ограниченность гистограммного представления прослеживаются такие особенности, как возможная аппроксимация средней скорости гамма-распределением, а порывов – распределением Вейбулла с тяжелым правым хвостом. То есть порывы обнаруживают статистические свойства экстремальной величины с указанным законом притяжения, что отмечается в множестве публикаций (например, [10, 11]). В данном случае *параметры формы* обоих распределений выше единицы.

Определение порыва в 10-минутных интервалах наблюдений обуславливает высокую повторяемость малых скоростей порывов: в приведенном примере максимальный порыв не превышает 15 м/с, что, в частности, не дотягивает до градации "неблагоприятного метеорологического явления" [8].

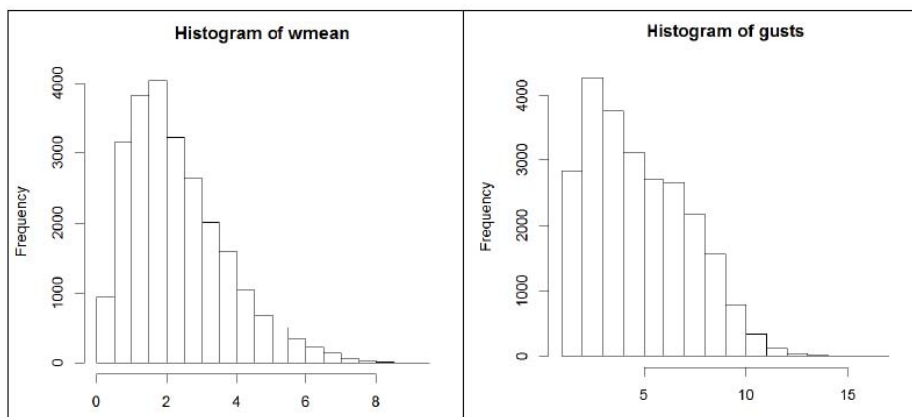


Рис. 8. Распределение средней скорости (левая панель) и порывов (правая панель) в 10-минутных интервалах по 196 станциям 25 июня 2024 г. (телеграммы файла 20240625.dat). Объем выборки – 27000 значений

Fig. 8. Distribution of average speed (left panel) and gusts (right panel) in 10-minute intervals at 196 stations on June 25, 2024 (telegrams of file 20240625.dat). Sample size – 27000 values.

Распределение *по времени* средней скорости ветра и порывов для отдельной станции в целом также сходно с гамма-распределением и распределением Вейбулла. Наличие порядка сотни наблюдений в сутки (из 144 максимальных) вполне достаточно для построения устойчивого распределения по паре месяцев года. Для иллюстрации приведем типичные гистограммы распределения для мая-июня 2025 г. На рис. 9 и 10 приводятся две типичные гистограммы распределения по времени средней скорости ветра и порывов, напоминающие гистограммы частот по всем станциям для одного дня (рис. 8). Заметно различие форм возможных аппроксимаций приведенных гистограмм: если для станции 19028 параметр формы гамма-распределения явно больше 1, а параметр формы распределения Вейбулла явно меньше 1, то для станции 19101 типы параметров меняются на противоположные.

Подчеркнем важное обстоятельство. Наблюдения АМС были использованы в том виде, в котором они поступали в базу данных Гидрометцентра – без предварительного контроля качества. В результате первичного статистического анализа было установлено, что имеется несколько десятков телеграмм с константами отсутствия (NA) в ячейках характеристик ветра, при этом некоторые станции генерировали явно бракованные наблюдения, хотя сами значения были реальными числами в привычном диапазоне изменения величины. Например, станция 19025 (первая в списке активно действующих АМС) прислала мае-июне 7803 приемлемых сообщений о средней скорости, но явно бракованные сообщения о порывах, что проявилось на соответствующих гистограммах плотности распределения (рис. 11).

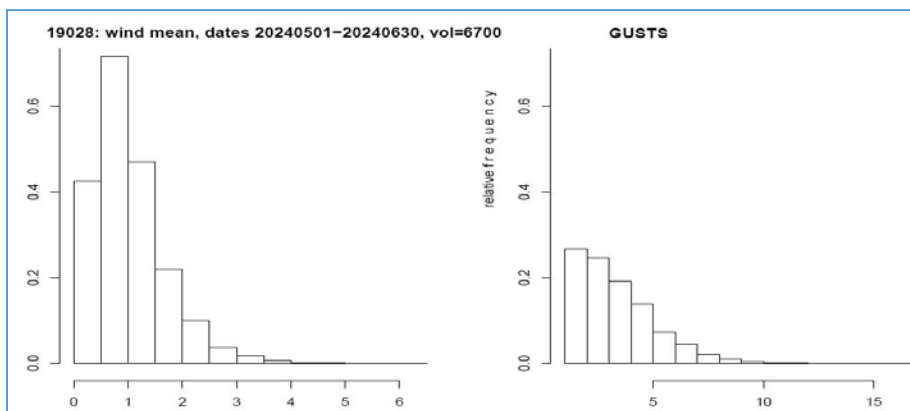


Рис. 9. Гистограммы относительных частот наблюдений средней скорости ветра (левая панель) и значений порывов (правая панель) на станции 19028 в период май-июнь 2024 г.

Fig. 9. Histograms of relative observation frequencies of average wind speed (left panel) and gust values (right panel) at station 19028 in the period May-June 2024.

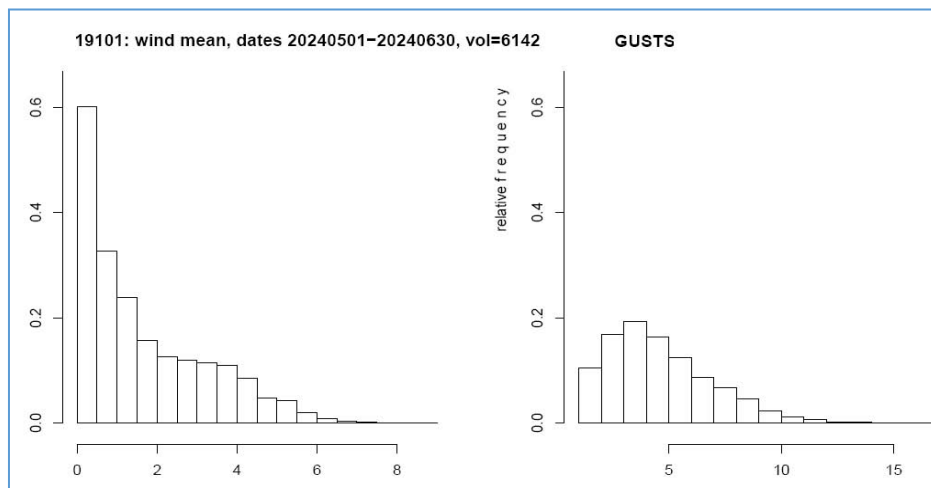


Рис. 10. Гистограммы относительных частот наблюдений средней скорости ветра (левая панель) и значений порывов (правая панель) на станции 19101 в период май-июнь 2024 г. Объем выборки 6142 телеграммы.

Fig. 10. Histograms of relative observation frequencies of average wind speed (left panel) and gust values (right panel) at station 19101 in the period May-June 2024. Sample size 6142 telegrams.

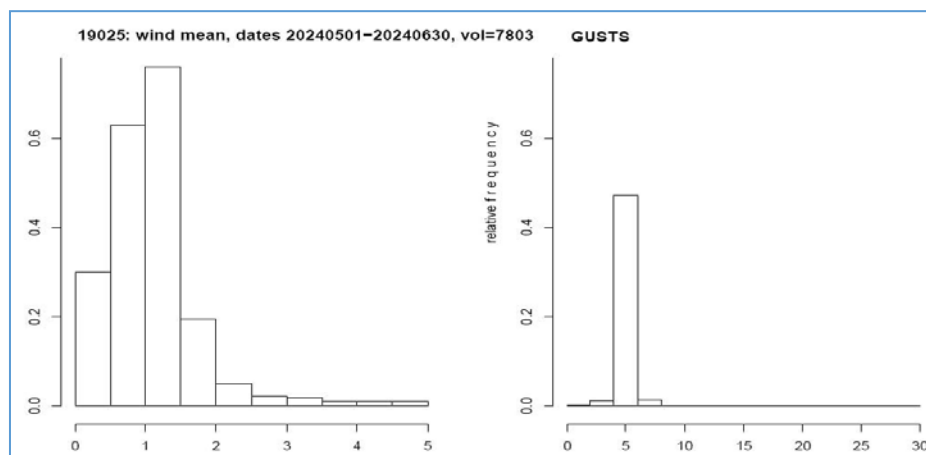


Рис. 11. Гистограммы относительных частот наблюдений средней скорости ветра (левая панель) и значений порывов (правая панель) на станции 19025 в период май-июнь 2024 г. Объем выборки 7803 телеграммы.

Fig. 11. Histograms of relative observation frequencies of average wind speed (left panel) and gust values (right panel) at station 19025 in the period May-June 2024. Sample size 7803 telegrams.

Ниже будут рассмотрены данные по всему архиву испытаний с 1 мая до 23 сентября 2024 года.

3.4. Сопоставление разнородных данных: синхронизация и дихотомизация

По всему периоду испытаний в полях наукастинга и мезопрогнозов, а также в сообщениях АМС выделяются сроки, в которые наблюдались "явления порыва" хотя бы в одной точке поля или хотя бы в одной станции соответствующей телеграммы. Напомним, под порывом понимается превышение определенного порога скорости: для наукастинга и мезопрогнозов 20 м/с, а для АМС 15 и 20 м/с.

Количество синхронизированных пар (и троек) полей и станционных наблюдений свидетельствует об информационном покрытии интервала испытаний и о статистической значимости результатов верификации. Приведем эти объемы, напомнив максимально возможные объемы полей и наборов телеграмм. В интервале от 1 мая до 23 сентября 2024 г. содержится 146 суток по 144 десятиминутных сроков в каждом сутках. Таким образом, максимально возможное количество наукастов одной заблаговременности и сроков наблюдений равно 21024. Максимальное количество мезопрогнозов одной заблаговременности равно $146 \times 8 = 1168$.

Виды и результаты проделанной синхронизации выглядят следующим образом (с округлениями в десятых).

– *Синхронизация пар: поля наукастинга FCAST и телеграммы наблюдений АМС.* Количество пар для прогнозов от 10 до 100 минут в 10-минутной дискретности равно около 20640 для каждой заблаговременности.

– *Синхронизация пар: поля наукастинга FCAST и поля мезопрогноза COSMO.* Количество пар равно около 1150 для каждой заблаговременности.

– *Синхронизация троек: поля наукастинга FCAST, поля мезопрогноза COSMO и телеграммы АМС.* Количество троек для заблаговременностей от 10 до 100 мин равно около 1140.

– *Синхронизация пар: поля наукастинга FCAST и поля наукастинга "первого приближения" (наукастинга на 10 минут).* Количество пар лишь незначительно отличается от соответствующего количества пар FCAST – АМС, т. е. около 20640.

Перечень не полон: не рассматривались синхронизированные пары *поле наукастинга и поле мезопрогноза "первого приближения"* (мезопрогноза на 10 мин).

Очевидно, что большая разница объемов выборок пар и троек наборов данных усложняет статистическое сопоставление качества прогнозов между синхронизированными парами и тройками.

Также следует учитывать условности и ограничения, которые накладываются исследуемой величиной – ветровыми порывами. Основными свойствами порывов являются мелкомасштабность и турбулентность, т. е. кратковременность развития (до минуты для порывов и до нескольких минут для шквалов), относительная редкость и случайность проявления в пространстве и времени.

Мелкомасштабность и редкость заставляют прибегать к обобщениям (агрегированию) как по пространству, так и по времени, а турбулентный характер явления вынуждает использовать вероятностные определения и термины. При анализе качества моделирования и прогнозирования редкого явления следует ожидать чрезмерного количества как случаев "отсутствия явления", так и "отсутствия прогнозов явления". Малоинформативными оказываются или даже теряют смысл такие стандартные оценки качества, как общая оправдываемость и критерий Пирса – Обухова: первый показатель стремится к единице, второй – к доли попаданий. Так как теряется значимость "отсутствия явления", из дихотомического анализа часто исключают ячейку "нет-нет", или (2,2), в таблице сопряженности [9]. Именно по указанным причинам одной из наиболее подходящих мер качества считается критерий "отношения пространственных долей", FSS (Fractions Skill Score), учитывающий, как пространственный масштаб явления, так и его вероятностный характер [1, 5].

На этом этапе удастся частично решить и проблему "больших данных" переходом от величины порыва (действительное число) к категории "порыва" (0 или 1) и сохранением для обработки лишь тех дихотомических полей, в которых имеются метки "порыва", привычно обозначаемые единицей. Экономия на дихотомизации при верификации прогнозов существенно упрощает расчеты оценки FSS, так как данный показатель не вырождается лишь в тех случаях, когда *каждое* из двух полей (прогноз – контроль) имеют хотя бы одну единицу. Также существенно упрощается процедура обобщения полей на часовые интервалы, когда в точку обобщенного поля единица записывается дизъюнктивно – если среди шести последовательных полей в данной точке имеется хотя бы одна единица.

Имеет смысл дать короткую справку о наличии такого рода полей, об их количестве и о количестве единиц, назвав такие порывы "существенными".

– *Существенные порывы в сообщениях AMS.* Среди примерно 20600 станционных сообщений было обнаружено 549 сообщений с порывами не менее 15 м/с (чуть менее 3 %) и 119 сообщений с порывами не менее 20 м/с (менее 0.1 %).

Рассмотрим распределение сообщений по количеству станций, одновременно зафиксировавших порыв в интервале 10 мин. По одному порыву (т.е. на одной станции) было зафиксировано в 429 сообщениях ($429/549=78\%$). По два порыва содержались в 68 сообщениях ($68/549=12\%$), по три порыва – в 19 сообщениях, по четыре – в 15, по 5 и 6 порывов – в 13 сообщениях. На 7, 10 и 13 станциях порывы были зафиксированы единожды и в двух сообщениях данные о порывах получены с 11 станций. Намного скуднее количество сообщений о порывах не менее 20 м/с: общее количество сообщений с порывами – 119, из них 1 порыв отмечен в 111 сообщениях (93 %), 2 – в 4, 3 и 4 порыва отмечены в двух сообщениях. Таким образом, основная масса срочных сообщений содержит наблюдения порыва на единственной станции (в 78 % случаев для порога 15 м/с и в 93 % случаев – для порога 20 м/с). При этом максимальное

количество станций, одновременно зафиксировавших существенные порывы, оказалось равным 13 (в двух сообщениях) для порога 15 м/с и 4 станции (в двух сообщениях) для порога 20 м/с.

– *Существенные порывы в наукастах (FCAST)*. Табл. 6 содержит полную информацию о пространственном распределении прогнозируемых порывов не менее 20 м/с в зависимости от заблаговременности.

Наблюдаются следующие особенности характеристик распределения в зависимости от заблаговременности. Количество полей с наличием порывов растет от 400 до 450, затем падает до 376. Минимальное количество (порыв в одной точке поля) не изменяется, квартили до 75-го изменяются без особого тренда. Квартиль q_{75} в основном снижается от 2334 до 1269, а максимум растет от 19500 до 28500 (почти на 10 тысяч). При этом среднее количество порывов в поле для каждой заблаговременности существенно не меняется. Снижение квартиля 75 % и рост максимального количества точек с порывами с ростом заблаговременности систематично, но вряд ли интерпретируемо.

Таблица 6. Квартильное распределение порывов в наукастах (FCAST): заблаговременность (ld минут), количество полей в выборке (#fields), минимальное количество единиц (vmin), квартильные значения количества единиц в 25% полей (q_{25}), в 50 % полей (median) в 75% полей (q_{75}), максимальное количество единиц (vmax) и среднее количество единиц по всем полям выборки
Table 6. Quartile distribution of gusts in nowcasts (FCAST): lead time (ld minutes), number of fields in the sample (#fields), minimum number of units (vmin), quartile values of the number of units in 25% of fields (q_{25}), in 50% of fields (median) in 75% of fields (q_{75}), maximum number of units (vmax) and average number of units across all fields in the sample

| ld | #fields | vmin | q_{25} | median | mean | q_{75} | vmax |
|-----|---------|------|----------|--------|------|----------|-------|
| 10 | 406 | 1 | 100 | 547 | 1799 | 2334 | 19426 |
| 20 | 403 | 1 | 161 | 581 | 1884 | 2366 | 20322 |
| 30 | 425 | 1 | 153 | 537 | 1846 | 2142 | 20868 |
| 40 | 444 | 1 | 112 | 522 | 1807 | 1989 | 21274 |
| 50 | 454 | 1 | 111 | 496 | 1789 | 1953 | 22159 |
| 60 | 449 | 1 | 118 | 481 | 1806 | 1898 | 23123 |
| 70 | 449 | 1 | 128 | 445 | 1807 | 1904 | 24070 |
| 80 | 446 | 1 | 134 | 484 | 1800 | 2012 | 25269 |
| 90 | 448 | 1 | 116 | 462 | 1754 | 1884 | 26308 |
| 110 | 428 | 1 | 100 | 452 | 1742 | 1618 | 28179 |
| 120 | 405 | 1 | 138 | 510 | 1800 | 1559 | 28595 |
| 130 | 398 | 1 | 124 | 478 | 1784 | 1438 | 28308 |
| 140 | 391 | 1 | 108 | 458 | 1757 | 1342 | 28662 |
| 150 | 376 | 1 | 126 | 446 | 1762 | 1269 | 28529 |

– *Существенные порывы в мезопргнозах (COSMO-Ru2.2)*. Табл. 7 содержит информацию о пространственном распределении прогнозируемых порывов не менее 20 м/с в зависимости от заблаговременности.

Напомним, количество порывов в урезанном и переинтерполированном поле примерно вдвое превышает количество порывов в исходном поле из-за интерполяции методом "ближайшего соседа". Размах максимальных порывов в полях наукастинга в интервале составляет 20–29 тысяч (табл. 6) намного меньше аналогичного размаха от 7500 до 40000 точек в мезопрогнозах (табл. 7). Сужение размаха в наукастинге можно объяснить особенностями алгоритма построения дерева решений, в котором используется скользящий пространственный фильтр масштаба около 10 км (раздел 1); в общем *поля порывов* в наукастинге оказываются более сглаженными, чем *поля среднего ветра* в мезопрогнозе.

Таблица 7. Квартильное распределение порывов в мезопрогнозах (COSMO-Ru/2.2): заблаговременность (ld минут), количество полей в выборке (#fields), минимальное количество единиц (vmin), квартильные значения количества единиц в 25% полей (q25), в 50% полей (median) в 75% полей (q75) и максимальное количество единиц (vmax) и среднее количество единиц по всем полям выборки(mean)

Table 7. Quartile distribution of gusts in mesoforecasts (COSMO-Ru/2.2): lead time (ld minutes), number of fields in the sample (#fields), minimum number of units (vmin), quartile values of the number of units in 25% of fields (q25), in 50% of fields (median) in 75% of fields (q75) and maximum number of units (vmax) and average number of units for all fields in the sample (mean)

| ld | #fields | vmin | q25 | median | mean | q75 | vmax |
|-----|---------|------|-----|--------|------|------|-------|
| 10 | 315 | 4 | 28 | 144 | 583 | 550 | 13420 |
| 20 | 352 | 4 | 28 | 160 | 820 | 835 | 19076 |
| 30 | 375 | 4 | 32 | 188 | 1036 | 988 | 23800 |
| 40 | 395 | 4 | 32 | 244 | 1230 | 1134 | 27072 |
| 50 | 411 | 4 | 36 | 244 | 1394 | 1270 | 29656 |
| 60 | 420 | 4 | 36 | 238 | 1553 | 1357 | 31924 |
| | | | | | | | |
| 70 | 255 | 4 | 26 | 112 | 553 | 458 | 7412 |
| 80 | 299 | 4 | 32 | 156 | 791 | 682 | 12372 |
| 90 | 326 | 4 | 32 | 190 | 1020 | 896 | 17084 |
| 100 | 339 | 4 | 36 | 228 | 1246 | 1074 | 21712 |
| 110 | 352 | 4 | 40 | 228 | 1435 | 1316 | 26284 |
| 120 | 362 | 4 | 44 | 250 | 1609 | 1408 | 30652 |
| | | | | | | | |
| 130 | 236 | 4 | 28 | 158 | 564 | 461 | 9744 |
| 140 | 250 | 4 | 40 | 222 | 867 | 775 | 16288 |
| 150 | 264 | 4 | 40 | 250 | 1138 | 1009 | 21648 |
| 160 | 277 | 4 | 36 | 276 | 1377 | 1280 | 27480 |
| 170 | 285 | 4 | 36 | 300 | 1622 | 1512 | 33948 |
| 180 | 294 | 4 | 40 | 332 | 1848 | 1631 | 40144 |

В табл. 7 пустые строки вставлены намеренно, чтобы подчеркнуть "пилообразное" изменение квартильных характеристик с часовым циклом в столбцах количества полей, медианы, средней величины, квартиля 75 % и максимума. Изменение явно систематическое и связано с особенностями прогнозирования порывов в модели COSMO-Ru2.2.

Часовой цикл проявляется и в количестве точек с порывами, и в расположении очагов порывов, иногда он выражен и на большем интервале времени. Так, в прогнозе в 10:00 20 июня при переходе от заблаговременности 60 мин к 70 мин количество точек с существенными порывами уменьшилось с 25080 до 6092, при переходе от 120 мин к 130 мин – с 23676 до 4692 точек, т. е. область порывистости сузилась в 4 и 5 раз соответственно. При переходе от 180 мин к 190 мин эта область уменьшилась почти в 8 раз!

При этом в полях среднего ветра часовой цикличности в 10-минутных прогнозах не обнаруживается ни в картах, ни в таблицах квантилей (которые здесь не приводятся). К этой "часовой пиле" оказались устойчивыми и наукасты, однако некоторые сходные явления проявились в оценках качества наукастинга относительно мезопрогнозов.

4. Анализ случая и результаты верификации

Ввиду отмеченной систематической особенности прогноза порывов в модели COSMO, результаты сопоставления синхронизованных полей наукастинга и мезопрогнозов нельзя считать корректными. Резкое снижение количества пар полей наукастинга с синхронизованными полями мезопрогнозов "первого приближения" снижает статистическую устойчивость оценок качества, в связи с чем остаются лишь варианты сравнения пар *наукастинг – станции* и *наукастинг – наукастинг первого приближения*.

Несомненно, сопоставление в один срок менее двухсот станционных наблюдений и прогностических полей с одним миллионом узлов имеет мало смысла со статистической точки зрения, если требуется оценка прогнозов по всей области. В таких случаях проводится подробный анализ отдельных синоптических ситуаций (case studies) с наличием областей сильных порывов, подтвержденных и несколькими станциями наблюдательной системы, и прогностическими полями. Такого рода "анализ случая", помимо убедительной иллюстрации, характеризует общую полезность схемы наукастинга. Рассмотрим в этой связи синоптическую ситуацию 20 июня в ЦФО, в которой вокруг Москвы к 10:00 сформировался очаг порывов (не менее 15 м/с), зафиксированных несколькими АМС.

4.1. Порывы ветра 20 июня 2024 г, на территории Московской области

На рис. 12 последовательно расположены карты прогноза системы наукастинга порывов на 90, 60, 30 и 10 минут, зафиксированных станциями в центральной части Московской области.

На всех картах цветом магента выделены значения станционных наблюдений выше 15 м/с в условный срок 00 минут (эти значения неизменны на всех картах). В центральной части карты на нескольких станциях зафиксированы скорости порывов выше 15 м/с, при этом максимальный порыв (22.9 м/с) сообщен станцией, расположенной к востоку от Калуги.

В полях FCAST максимальная скорость растет с уменьшением заблаговременности: 16.9 (90 мин), 17.6 (60 мин), 20.9 (30 мин) и 22.2 м/с (10 мин). Однако превышение порога 20 м/с в поле наукастинга (желтый цвет) прогнозируется в районе Курска, т. е. далеко за пределами станционного скопления порывов. При этом наиболее точное указание на район возможной порывистости ветра (в центральной части карты) дается лишь в прогнозах от часа и меньше; максимально отчетливо – лишь на карте прогноза заблаговременности 30 минут.

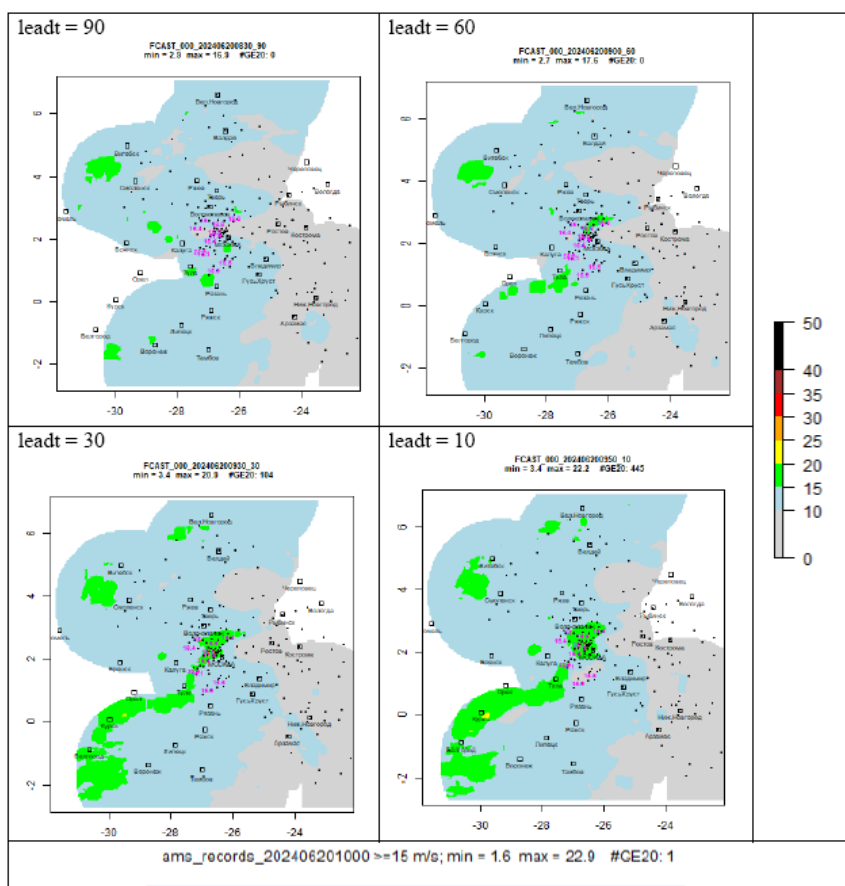


Рис. 12. Карты наукастинга в 10:00 20 июня 2024 г. на 90, 60, 30 и 10 мин (leadt).

Fig. 12. Nowcasting maps at 10:00 on June 20, 2024 for 90, 60, 30 and 10 min (leadt).

В панелях рис. 12 заметны более масштабные особенности, никак не фиксируемые станциями. Например, система наукастинга довольно быстро заполняет очагом существенной скорости широкую полосу в левом нижнем углу карты вне станционного массива. Наличие развивающейся ситуации большого масштаба подтверждается 60-минутным прогнозом COSMO-Ru2.2 со стартом в 09:00 (рис. 13).

Как показал предыдущий анализ архива испытаний, обоснованные выводы можно сформулировать лишь при двух вариантах сопоставлений: при визуальном анализе всех трех типов информации в отдельных ситуациях, как показано выше в "анализе случая", и при использовании наукастов первого приближения. При сопоставлении полей рекомендуются высокоинформативные критерии пространственной верификации, в частности, упомянутые выше оценки качества по "пространственным долям" FSS. Результаты оценок качества по этому показателю собраны в табл. 8 [2]. Формулировка вывода такова. Если считать прогнозы системы наукастинга на 10 минут наиболее точно отражающими реальность и использовать их в виде "первого приближения", то качество прогностической системы (по значению FSS > 0.5) в пространственном масштабе до 40×40 км и в часовом обобщении превышает 50 % в заблаговременности до полу- часа.

Таблица 8. Количество пар с оценкой FSS >0.5 в окнах со сторонами 0–40 км; "сумма" всех случаев FSS > 0.5 для условных заблаговременностей 10–90 мин

Table 8. Number of pairs with FSS >0.5 in windows with sides of 0–40 km; "sum" of all cases of FSS > 0.5 for conditional lead times of 10–90 min

| Условная заблаговрем. | Кол-во пар | Окна со стороной (в км) | | | | | Сумма |
|-----------------------|------------|-------------------------|-----|-----|-----|-----|-------|
| | | 0 | 10 | 20 | 30 | 40 | |
| 10 | 453 | 245 | 277 | 289 | 293 | 297 | 1401 |
| 20 | 419 | 182 | 209 | 221 | 233 | 242 | 1087 |
| 30 | 398 | 135 | 171 | 183 | 194 | 201 | 884 |
| 40 | 372 | 101 | 126 | 134 | 154 | 164 | 679 |
| 50 | 357 | 91 | 106 | 118 | 127 | 135 | 577 |
| 60 | 335 | 59 | 86 | 94 | 109 | 114 | 462 |
| 70 | 307 | 43 | 55 | 68 | 76 | 84 | 326 |
| 80 | 290 | 32 | 42 | 49 | 64 | 69 | 256 |
| 90 | 270 | 23 | 31 | 39 | 45 | 53 | 191 |

Сделаем одно замечание, касающееся свойств показателя FSS в тех ситуациях, когда явление одновременно редкое и наблюдается на границе домена верификации, что приводит к вырождению показателя, т. е. к резкому росту до единицы или к падению до нуля. Универсального рецепта против вырождения показателя FSS не существует, если не считать простого правила устанавливать пороговое (наименьшее) количество точек поля и не учитывать в суммарном представлении качества оценки вблизи границ домена верификации.

Выводы

На основе представленной табличной и графической информации можно сформулировать выводы об особенностях и качестве отдельных типов использованных данных, а также об успешности наукастинга порывов

в заблаговременности до полутора часов по результатам испытаний в период май–сентябрь 2024 года.

Во-первых, выявлены специфические свойства трех типов использованных данных, которые необходимо учитывать и по возможности устранять при усовершенствовании системы наукастинга порывов ветра с опорой на отечественные ресурсы:

- алгоритм машинного обучения порождает *сильно сглаженные* прогностические поля порывов (наукасты);
- в прогнозах скорости порывов модели COSMO-Ru2.2 версии 2024 г. прослеживается *систематическое смещение*;
- стационарные наблюдения порывов нуждаются в серьезном *контроле качества*.

Во-вторых, в условиях проведенных испытаний и на основе пространственной оценки качества "по полю первого приближения" итоговый результат таков: информативный (или полезный) интервал наукастинга сильных порывов (не менее 20 м/с) в часовом обобщении и в масштабе окна со стороны до 40 км *не превышает 30 минут*.

Работа выполнена в рамках темы 1.1.5 Плана НИТР Росгидромета на 2020–2024 гг.

Список литературы

1. Киктев Д.Б., Муравьев А.В., Бундель А.Ю. Методические рекомендации по верификации метеорологических прогнозов. М.: АМА ПРЕСС, 2021. 94 с.
2. Киктев Д.Б., Муравьев А.В., Смирнов А.В. Технология наукастинга скорости порывов ветра в летний период на основе данных дистанционного зондирования атмосферы и численных прогнозов погоды // Гидрометеорологические исследования и прогнозы. 2025. №2 (396). С. 169-173.
3. Методическое письмо об итогах работы сети ДМРЛ Росгидромета в 2020 г. Долгосрочный: Центральная аэрологическая обсерватория, 2021. 88 с.
4. Муравьев А.В., Бундель А.Ю., Киктев Д.Б., Смирнов А.В. Опыт пространственной верификации радиолокационного наукастинга осадков: определение и статистика объектов, ситуаций и условных выборов // Гидрометеорологические исследования и прогнозы. 2022. № 2 (384). С. 6-52.
5. Муравьев А.В., Киктев Д.Б. Качество, предсказуемость и полезность в задачах радиолокационного наукастинга осадков // Метеорология и гидрология. 2024. № 7. С. 93-107.
6. Муравьев А.В., Киктев Д.Б., Смирнов А.В. Наукастинг порывов ветра с помощью численных прогнозов, данных радиолокации и машинного обучения: определения и термины, инструменты наблюдений и модели // Гидрометеорологические исследования и прогнозы. 2025. №3 (397). С. 8-31.
7. Программный комплекс вторичной обработки информации доплеровского метеорологического радиолокатора ДМРЛ-С (шифр «ГИМЕТ-2010») версия 02: Свидетельство о государственной регистрации программы для ЭВМ. Правообладатель: ФГБУ «ЦАО», номер регистрации: 2018665447, дата регистрации: 05.12.2018.
8. РД 52.27.724-2019 Наставление по краткосрочным прогнозам погоды общего назначения. М.: Гидрометцентр России», 2019. 66 с.
9. Forecast Verification in Atmospheric Science. A Practitioner's Guide: Second Ed. / I. Jolliffe, D. Stephenson (Eds.). John Wiley & Sons Ltd. 2012. 274 p.
10. Kislov A., Matveeva T. An Extreme Value Analysis of Wind Speed over the European and Siberian Parts of Arctic Region // Atmospheric and Climate Sciences. 2016. Vol. 6. P. 205-223.

11. Palutikof J.P., Brabson B.B., Lister D.H., Adcock S.T. A review of methods to calculate extreme wind speeds // *Meteorol. Appl.* 1999. Vol. 6. P.119-132 .

References

1. Kiktev D.B., Murav'ev A.V., Bundel' A.Yu. Metodicheskie rekomendatsii po verifikatsii meteorologicheskikh prognozov. Moscow, AMA PRESS Publ., 2021, 94 p. [in Russ.].

2. Kiktev D.B., Muravyev A.V., Smirnov A.V. Technology for nowcasting wind gusts in summer based on the data of remote sensing of the atmosphere and numerical weather prediction. *Gidrometeorologicheskie issledovaniya i prognozy* [*Hydrometeorological Research and Forecasting*]. 2025, vol. 396, no. 2, pp. 169-173 [in Russ.].

3. Metodicheskoe pis'mo ob itogah raboty seti DMRL Rosgidrometa v 2020 g. Dolgoprudnyy, Tsentral'naya aerologicheskaya observatoriya, 2021, 88 p. [in Russ.].

4. Muravev A.V., Bundel A.Yu., Kiktev D.B., Smirnov A.V. Expertise in spatial verification of radar precipitation nowcasting: identification and statistics of objects, situations and conditional samples. *Gidrometeorologicheskie issledovaniya i prognozy* [*Hydrometeorological Research and Forecasting*]. 2022, vol. 384, no. 2, pp. 6-52 [in Russ.].

5. Muravev A.V., Kiktev D.B. Quality, Predictability, and Utility in Radar Precipitation Nowcasting Applications. *Russ. Meteorol. Hydrol.* 2024, vol. 49, no. 7, pp. 627-637. DOI: 10.3103/S1068373924070070.

6. Muravev A.V., Kiktev D.B., Smirnov A.V. Naukasting poryvov vetra s pomoshch'yu chislennykh prognozov, dannykh radiolokatsii i mashinnogo obucheniya: opredeleniya i terminy, instrumenty nablyudeniya i modeli. *Gidrometeorologicheskie issledovaniya i prognozy* [*Hydrometeorological Research and Forecasting*]. 2025, vol. 397, no. 3, pp. 8-31 [in Russ.].

7. Programmnyy kompleks vtorichnoy obrabotki informatsii doplerovskogo meteorologicheskogo radiolokatora DMRL-S (shifr «GIMET-2010») versiya 02: Svidetel'stvo o gosudarstvennoy registratsii programmy dlya EVM. Pravoobladatel': FGBU «TSAO», nomer registratsii: 2018665447, data registratsii: 05.12.2018 [in Russ.].

8. RD 52.27.724-2019 Nastavlenie po kratkosrochnym prognozam pogody obshchego naznacheniya. Moscow, Gidromettsentr Rossii [Hydrometcenter of Russia], 2019, 66 p. [in Russ.].

9. Forecast Verification in Atmospheric Science. A Practitioner's Guide: Second Ed. John Wiley & Sons Ltd., 2012, 274 p.

10. Kislov A., Matveeva T. An Extreme Value Analysis of Wind Speed over the European and Siberian Parts of Arctic Region. *Atmospheric and Climate Sciences*. 2016, vol. 6, pp. 205-223.

11. Palutikof J.P., Brabson B.B., Lister D.H., Adcock S.T. A review of methods to calculate extreme wind speeds. *Meteorol. Appl.* accepted for publication 1999, vol. 6, pp.119-132.

Поступила 17.08.2025; принята в печать 26.11.2025.
Submitted 17.08.2025; accepted for publication 26.11.2025.