

## РЕЦЕНЗИЯ № 2

на статью «ПОТОЧЕЧНЫЕ И КОМПЛЕКСНЫЕ МЕРЫ КАЧЕСТВА В ИССЛЕДОВАНИЯХ АТМОСФЕРЫ И ОКЕАНА: ОБЗОР МЕТОДОВ И ПОДХОДОВ»

авторского коллектива: В. Ю. Резвов, М. А. Криницкий, Н. Д. Тилинина

Методы для оценки точности метеорологических и океанографических данных, представляющих собой результат измерений или численного моделирования, являются сейчас особенно востребованными, ввиду повышенного внимания исследователей к методам машинного обучения в приложении к задачам прогнозирования погоды и климата. Важным элементом технологий на основе методов машинного обучения является функция ошибки, посредством которой происходит оценка точности приближенного решения в сравнении с имеющимися данными.

Рассматриваемая статья содержит сводную информацию о способах оценки точности воспроизведения атмосферных и океанических процессов методами математического моделирования. Обзорные работы часто являются полезными, поскольку позволяют получить целостное представление о подходах, применяемых для решения конкретных задач. Полагаю, что данная статья может стать такой только после доработки и более предметного анализа существующих подходов.

К основным недостаткам статьи следует отнести:

1. Статья является композицией других работ. Заимствования, обычно, сопровождаются ссылкой на цитируемую публикацию. Мнение авторов статьи представлено в разделе Заключение и одним абзацем на строках 345–353, что суммарно составляет величину менее 10 % от общего объема статьи (без учета библиографии).

В разделе Заключение содержательным является только первый абзац. Остальная часть данного раздела представляет собой набор общих слов с ничем не подкрепленными рассуждениями о трендах и тенденциях «в развитии метрик качеств».

Строки 487–492 в Заключении являются перечислением методов и не могут быть отнесены к выводам по статье.

Собственное мнение автора является важной составляющей обзорной работы. Оно должно фокусировать внимание читателя на отдельных аспектах рассматриваемых методик в контексте конкретных задач и приложений. Авторам предлагается акцентировать содержание статьи на методах оценки точности алгоритмов в задаче повышения пространственного разрешения результатов численного моделирования.

2. Название работы носит общий характер и предполагает детальное, разноплановое и логичное описание существующих методов для оценки точности численного описания процессов в атмосфере и океане. В абстракте к статье указано, что в ней «представлена классификация наиболее часто встречающихся в научной литературе метрик качества».

Однако классификация ведется на уровне поточечных и комплексных категорий методов. Более детальная проработка классификации методов отсутствует, как и обсуждение идей и приложений, согласно которым это разделение существует. Например, по результатам работы международного проекта Spatial Forecast Verification Methods Inter-Comparison Project для задачи пространственной оценки ТОУНОСТИ моделирования выделено 2 категории (методы фильтрации и окрестные методы) и 4 подхода (scale separation, neighborhood (or fuzzy), features based (or objects based), field deformation).

К остальным недостаткам статьи следует отнести:

1. К частному замечанию, имеющему, тем не менее, важное значение, следует отнести объяснение термина «масштабирование» только в одном фрагменте статьи – в конце введения. Под масштабированием здесь понимается «...совокупность методик для преобразования данных с низким пространственным разрешением в данные с высоким разрешением...». Данный оборот сопровождается ссылкой на работу [Sun, 2024], где используется общепринятый термин “downscaling”, который, действительно, характеризует преобразование данных с грубым разрешением в их высоко детализированное представление. Однако, на практике также распространен термин “upscaling”, который характеризует обратное преобразование хорошо разрешенных данных в их огрубленное представление [Atkinson, 2000]. Используемый в представленной к рассмотрению статье термин «масштабирование» представляется не вполне удачным, поскольку объединяет в себе оба англоязычных – термина “downscaling” и “upscaling”. Автор вправе использовать собственную терминологию.

Однако, ссылка на работу [Sun,2024] в этом случае является не правильной.

2. В ключевых словах и во Введении упоминаются объектно-ориентированные методы оценки точности. В абстракте указывается, что «Отдельное внимание в данной работе посвящено росту популярности объектно-ориентированных метрик». Однако, в тексте статьи этому классу методов уделено недостаточно внимания (подраздел «Другие комплексные метрики»). Специфика подхода (оценка – способности – модели воспроизводить структурные элементы прогнозируемых полей вместе с их динамикой) остается не раскрытой, как не дается определение «объекта» – того, с чем работают объектно-ориентированные метрики.

3. Информация о методе деформации поля (field-deformation) в представленной статье отсутствует. Также нет упоминания о ряде классических методов оценки точности, таких как показатель мастерства Хайдке (Heidke skill score, HSS), например.

4. В задаче оценки точности ансамблевого моделирования наряду с CRPS часто применяется метод КОС-кривой, численным показателем которой является ее интегральная характеристика. Данный способ оценки точности моделирования в статье представлен как "BOCSS", а его описание ограничено общими словами вида "BOCSS часто используется для оценки качества вероятностных прогнозов и «..представляет собой сравнение частоты попаданий и частоты ложных «срабатываний»». Помимо орфографической ошибки смысл данного предложения остается не ясным. О каком сравнении идет речь (чего с чем)? Какие «срабатывания» и «попадания» имеются в виду? В чем специфика (отличие) данного подхода в контексте оценки качества прогнозирования?

5. Статья является, в целом, плохо проработанной. Например, формула на строке 400 приписана [Бундель, 2021]. Остается неясным, зачем сопровождать ссылкой утверждение на строках 410–411 («Идеальный прогноз имеет значение FSS=1, а в наиболее плохом случае моделирования FSS=0»), если это является неотъемлемой частью рассматриваемого метода.

Подраздел «показатель навыков дробей» включает неоднократные ссылки на работу [Бундель,2021], в которой используется другой общепринятый перевод англоязычного термина Fractions Skill Score – оценка по пространственным долям. «Показатель навыков дробей» является дословным переводом Fractions Skill Score. Правильный перевод этого термина встречается только в Таблице 2.

6. Заимствование в отрыве от контекста источника цитирования в некоторых случаях приводит к фактическим ошибкам. В качестве примера можно привести строки 99–101 и 105–112, которые являются заимствованием W3 работы [Sun, 2024]. Ниже выделены заимствованные предложения: **Gridded meteorological data refers to a type of dataset where meteorological variables (e.g., precipitation, surface temperature and wind) are systematically arranged into regularly spaced grids across a geographic region. This data serves a crucial role in a variety of applications such as climate change (Gilabert et al.,**

2015), quantitative remote sensing (Sun et al., 2020), and agricultural production (Moreira et al., 2019). Primarily derived from in-situ station observations using gauges, satellite remote sensing measurements, and reanalysis products (Yeggina et al., 2020, Zandler et al., 2020, Song et al., 2022), the resulting product is a spatially and temporally complete, gridded dataset of meteorological variables (Sun et al., 2019). **Due to the limitations of numerical models, platforms and observation sensors, the spatial resolution of gridded meteorological data typically ranges from tens to hundreds of kilometers (Kikumoto et al., 2015, Maina et al., 2020). The coarse spatial resolution is insufficient for regional or local studies to identify refined details. For example, high spatial-resolution meteorological datasets are required when determining theoretical limitations using radiation budget and energy balance equations (Leng et al., 2017).** To fulfill these needs, downscaling, the process to derive high-resolution (HR) fine-scale data from low-resolution (LR) large-scale conditions, is considered a significant way to improve spatial resolutions (Boe et al., 2007, De Caceres et al., 2018).

Можно видеть, что в представленном фрагменте работы [Sun, 2024] оборот "Due to the limitations of numerical models.." относится к измеренным данным (in-situ station observations), данным – спутникового зондирования (satellite remote sensing measurements) и данным реанализа (reanalysis), которые, действительно, обладают в силу своей специфики достаточно грубым пространственным разрешением. Однако, в рассматриваемой статье сформулировано другое утверждение: «Характерные значения пространственного разрешения сеточных данных численных моделей составляют десятки и даже сотни километров..» Далее идет неверный перевод фрагмента "The coarse spatial resolution is insufficient for regional or local studies to identify refined details" следующего вида: "Такого разрешения недостаточно для детального регионального и местного моделирования погоды". В работе [Sun, 2024] речь идет о метеорологических данных, но не о результатах моделирования – об этом в [Sun, 2024] пишется ниже:

"Dynamical downscaling generally uses simulations of global climate models (GCM) to derive regional climate models (RCM) which can generate a set of fine-scale information".

Перевод из той же работы [Sun, 2024] фрагмента "For example, high spatial-resolution meteorological datasets are required when determining theoretical limitations using radiation budget and energy balance equations" также выполнен небрежно: "..высокое пространственное разрешение имеет определяющее значение в задачах со сложными вычислениями, например, энергетического баланса". Здесь опущен перевод оборота "theoretical limitations" и слова "equations". Из-за этого изначально туманное предложение потеряло всякий смысл: о какого рода вычислительных задачах идет речь? Энергетического баланса чего?

7. Абстракт к статье следует уточнить в части используемых слов и оборотов:

Строка 10: Не пояснен термин «масштабирование».

Строки 11–12: Процессы не могут лежать в основе моделей.

Строки 20–21: «Внимание» статьи не может быть «посвящено росту популярности».

Строки 16 и 22–23: Предложения «Для каждой группы методов качества приведены примеры их использования в научных задачах» и «Все рассматриваемые метрики дополняются существующими в научной литературе оценками их применения к научным задачам исследования атмосферы и океана» близки по смыслу, поэтому одно из них может быть опущено.

8. Строка 53: В предложении речь, вероятно, идет не о свойствах, но о способах применения метрик для обобщения, анализа и оценки.

9. Строки 32–35: Утверждение вида «Изменение климата оказывает значительное влияние на хозяйственную деятельность человечества, поэтому точное прогнозирование погоды важно не только для обеспечения безопасности населения, но и для мировой экономики в целом» должно быть подкреплено ссылками на публикации по результатам крупных международных проектов и/или статистическими отчетами международных организаций, имеющих прямое отношение к оценке влияния климатических изменений

на хозяйственную деятельность человека. Ссылки на работы "Statistical downscaling of numerical weather prediction based on convolutional neural networks" и "Approximation of high-resolution surface wind speed in the North Atlantic using discriminative and generative neural models based on RAS-NAAD 40-year hindcast" здесь выглядят неуместными.

Более того, в рамках данного предложения происходит смешение климатических и погодных масштабов прогнозирования: каждое из утверждений ("Изменение климата оказывает значительное влияние на хозяйственную деятельность человечества" и "точное прогнозирование погоды важно..") является верным, но их причинно-следственная связь не ясна.

Резюме рецензента: принять после доработки с учетом замечаний.

**Подпись. Рецензент № 2. 26.11.2024.**

**От редакции:** рецензия была направлена авторскому коллективу.

**Ответ рецензенту № 2 на Рецензию от 26.11.2024 а статью авторского коллектива: В. Ю. Резвов, М. А. Криницкий, Н. Д. Тилинина «ПОТОЧЕЧНЫЕ И КОМПЛЕКСНЫЕ МЕРЫ КАЧЕСТВА В ИССЛЕДОВАНИЯХ АТМОСФЕРЫ И ОКЕАНА: ОБЗОР МЕТОДОВ И ПОДХОДОВ».**

Уважаемый коллега!

Мы благодарны за ценные замечания, подчеркивающие существенные недостатки в нашей работе. Руководствуясь этими замечаниями, мы приложили все усилия, чтобы сделать работу более понятной и лучше структурированной. Далее мы представляем наши пошаговые ответы.

Замечания рецензента мы выделили полужирным. В новой версии статьи все исправления и добавления выделены желтым цветом.

**Рецензент:** *...Мнение авторов статьи представлено в разделе Заключение и одним абзацем на строках 345-353, что суммарно составляет величину менее 10 % от общего объема статьи (без учета библиографии).*

*В разделе Заключение содержательным является только первый абзац. Остальная часть данного раздела представляет собой набор общих слов с ничем не подкрепленными рассуждениями о трендах и тенденциях «в развитии метрик качества». Строки 487–492 в Заключении являются перечислением методов и не могут быть отнесены к выводам по статье.*

**Ответ:** Спасибо, мы переработали и сделали более последовательным и структурированным раздел Заключение, обобщив выводы, отойдя от общих слов и простого перечисления описанных методов.

**Рецензент:** *Собственное мнение автора является важной составляющей обзорной работы. Оно должно фокусировать внимание читателя на отдельных аспектах рассматриваемых методик в контексте конкретных задач и приложений. Авторам предлагается акцентировать содержание статьи на методах оценки точности алгоритмов в задаче повышения пространственного разрешения результатов численного моделирования.*

**Ответ:** Добавили по тексту работы собственное мнение авторов. Например, строки 126–136, 172–177, 205–208, 566–584, и раздел Заключение.

**Рецензент:** *Название работы носит общий характер и предполагает детальное, разноплановое и логичное описание существующих методов для оценки точности численного описания процессов в атмосфере и океане. В абстракте к статье указано, что в ней «представлена классификация наиболее часто встречающихся в научной литературе метрик качества». Однако, классификация ведется на уровне поточечных и комплексных категорий методов. Более детальная проработка классификации методов отсутствует, как и обсуждение идей и приложений, согласно которым это разделение существует. Например, ... для задачи пространственной оценки точности моделирования выделено 2 категории и 4 подхода.*

**Ответ:** Спасибо за акцент на недостаточной проработанности классификации.

Введение было сильно переработано и расширено. Оставлена информация, касающаяся непосредственно настоящей работы (актуальность и контекст в науках об

атмосфере и океане). Добавлена формулировка цели работы (строки 55–57) и указана ее структура (строки 58–68).

Мы переработали все разделы для сохранения единообразия в описании разных типов метрик.

Был переименован и существенно сокращен раздел «1 Общие свойства метрик в исследованиях атмосферы и океана» для того, чтобы избежать многократного повторения дних и тех же идей и слишком общих рассуждений.

В раздел «2 Классификация метрик» кратко добавили информацию о применимости разных типов метрик к разным типам величин (строки 150–156, 163–171).

Также упомянули о существовании иных, более подробных классификаций комплексных метрик, на основании проекта Spatial Forecast Verification Methods Inter-Comparison Project (ICP) – строки 172–208.

Таблица 1 помещена, что более логично, в раздел «3 Поточечные метрики», так как в ней содержится описание именно основных поточечных метрик, для симметрии с расположением таблицы 2 в разделе «4 Комплексные метрики».

Аналогично, для единообразия разделов 3 и 4 рассуждения об ограничениях классических поточечных метрик вместо отдельного раздела помещены в «3 Поточечные метрики».

**Рецензент:** *К частному замечанию, имеющему, тем не менее, важное значение, следует отнести объяснение термина «масштабирование» только в одном фрагменте статьи - в конце введения. Под масштабированием здесь понимается «совокупность методик для преобразования данных с низким пространственным разрешением в данные с высоким разрешением...» Данный оборот сопровождается ссылкой на работу [Sun, 2024], где используется общепринятый термин "downscaling", который, действительно, характеризует преобразование данных с грубым разрешением в их высокодетализированное представление. Однако, на практике также распространен термин "upscaling", который характеризует обратное преобразование хорошо разрешенных данных в их огрубленное представление [Atkinson, 2000]. Используемый в представленной к рассмотрению статье термин "масштабирование" представляется не вполне удачным, поскольку объединяет в себе оба англоязычных термина "downscaling" и "upscaling". Автор вправе использовать собственную терминологию.*

*Однако, ссылка на работу [Sun, 2024] в этом случае является не правильной.*

**Ответ:** Спасибо за внимание, обращенное на путаницу терминов в данной статье. Мы добавили объяснение используемых в статье терминов в строках 119–130: Проблема противопоставления вычислительной эффективности и требуемого пространственного разрешения результатов численного моделирования может быть решена с помощью специальных методов повышения их разрешения. В англоязычной литературе такие методы объединены под общим названием downscaling (Hewitson, Crane, 1996; Wilby, Wigley, 1997; Schoof, 2013; Sun et al., 2024). Преобразование данных с грубым разрешением в их представление с высокой детализацией повышает полезность сеточных данных для различных применений.

Поскольку русскоязычная терминология в данной прикладной области наук о Земле еще не устоялась, вместо калькированного «даунскейлинг» мы будем пользоваться более привычным для русского языка термином «пространственное масштабирование», понимая под ним не изменение пространственного разрешения данных вообще, а именно повышение такого разрешения.

**Рецензент:** *В ключевых словах и во Введении упоминаются объектно-ориентированные методы оценки точности. В абстракте указывается, что «Отдельное внимание в данной работе посвящено росту популярности объектно-ориентированных метрик». Однако, в тексте статьи этому классу методов уделено недостаточно внимания (подраздел «Другие комплексные метрики»). Специфика подхода (оценка способности — модели воспроизводить структурные элементы прогнозируемых полей вместе с их динамикой) остается не раскрытой, как не дается определение «объекта» - того, с чем работают объектно-ориентированные метрики.*

**Ответ:** Спасибо! По всему тексту неудачный термин «объектно-ориентированные метрики» был заменен на термин «феноменологические метрики». Данному типу мер мы выделили отдельный подраздел с более подробным раскрытием темы (строки 526–584).

**Рецензент:** *Информация о методе деформации поля в представленной статье отсутствует. Также нет упоминания о ряде классических методов оценки точности, таких как показатель мастерства Хайдке, например.*

**Ответ:** Спасибо, мы добавили информацию о методе деформации поля в строках 194–203, 577–580.

HSS (Критерий Багрова-Хайдке в русскоязычной литературе) упомянут в строках 508–513.

Мы также посчитали нужным добавить подробную информацию про характеристики оправдываемости прогноза в целом для бинарных переменных (включая критерий Пирси-Обухова – строки 440–483) и кратко про параметр асимметрии (строки 491-496).

**Рецензент:** *В задаче оценки точности ансамблевого моделирования ... часто применяется метод ROC-кривой, численным показателем которой является ее интегральная характеристика. Данный способ оценки точности моделирования в статье представлен как "ROCSS", а его описание ограничено общими словами....*

**Ответ:** Описание ROC Skill Score переработано и представлено в строках 598–612.

**Рецензент:** *Статья является, в целом, плохо проработанной. Например, формула на строке 400 приписана [Бундель, 2021]. Остается неясным, зачем сопровождать ссылкой утверждение на строках 410-411, если это является неотъемлемой частью рассматриваемого метода.*

*Подраздел "показатель навыков дробей" включает неоднократные ссылки на работу [Бундель, 2021], в которой используется другой общепринятый перевод англоязычного термина fractions skill score - оценка по пространственным долям.*

**Ответ:** Спасибо, по всему тексту статьи был принят более правильный перевод «оценка по пространственным долям».

**Рецензент:** *Займствование в отрыве от контекста источника цитирования в некоторых случаях приводит к фактическим ошибкам. В качестве примера можно привести строки 99–101 и 105–112, которые являются заимствованием из работы [Sun, 2024].*

**Ответ:** Спасибо, эти заимствования были переработаны, для правильного помещения в контекст статьи, как и в целом переделано Введение и раздел 1.

**Рецензент:** *Абстракт к статье следует уточнить в части используемых слов и оборотов.*

**Ответ:** Абстракт был существенно переработан.

**Рецензент:** *Строка 53: В предложении речь, вероятно, идет не о свойствах, но о способах применения метрик для обобщения, анализа и оценки.*

**Ответ:** Данный оборот был исключен из раздела «Общие свойства метрик в исследованиях атмосферы и океана»

**Рецензент:** *Строки 32–35: Утверждение вида «Изменение климата оказывает значительное влияние на хозяйственную деятельность человечества, поэтому точное прогнозирование погоды важно не только для обеспечения безопасности населения, но и для мировой экономики в целом» должно быть подкреплено ссылками на публикации по результатам крупных международных проектов и/или статистическими отчетами международных организаций, имеющими прямое отношение к оценке влияния климатических изменений на хозяйственную деятельность человека.*

...  
Более того, в рамках данного предложения происходит смешение климатических и погодных масштабов прогнозирования: каждое из утверждений («Изменение климата оказывает значительное влияние на хозяйственную деятельность человечества» и «точное прогнозирование погоды важно...») является верным, но их причинно-следственная связь не ясна.

**Ответ:** Спасибо, данный абзац мы исключили из Введения, как не имеющий отношения к последующим рассуждениям об актуальности исследования метрик качества.  
**С уважением, авторский коллектив. 06.12.2024.**

**От редакции:** ответ и доработанная версия статьи были направлены редакцией рецензенту.

## **Подтверждение Рецензента № 2 на публикацию:**

Авторы существенно доработали статью в части логики изложения и полноты представленных в обзоре методов. Все замечания учтены.

### Несущественные замечания:

1. строка 43: Вместо слова «поспособствовать» лучше использовать «способствовать».

2. строки 58–59: Определение основного для статьи понятия об атмосферных и океанических данных сформулировано нечетко: «данные часто представляют», «поля некоторых переменных», «в определенном .. регионе».

3. строка 81: Неясно, почему сопоставление «сеточных геофизических данных» должно проводиться «на регулярной сетке». Данное утверждение необходимо либо изменить, либо подтвердить путем указания источника.

4. строка 89: Применительно к термину «метрика» оборот «ни одна из них не может .. проанализировать» является неудачным. |

5. строка 89: Сленг «прогнозное поле осадков» лучше не использовать. Предлагается: прогнозируемое/предсказываемое моделью...

6. строка 90: Оборот «корректируется ассимиляцией» требует пояснения, поскольку вырван из контекста. Ассимиляция чего? Русскоязычным аналогом термина «ассимиляция» является усвоение. Если так, то каким образом усвоение применяется для «коррекции» прогнозного поля осадков?

7. строка 103: Оборот «статистической стационарности данных» требует уточнения, поскольку речь, видимо, идет об ошибках, но в контексте предложения может быть интерпретирован как относящийся к результатам предсказания. Идет ли речь о систематических ошибках метода или имеется в виду что-то еще?

8. строка 111: Для анализа изменений климата.

9. строка 115: Что имеется в виду под словом «детальный» в контексте первой части предложения, где определен характерный масштаб («десятки и даже сотни километров»)? Разрешение современных моделей регионального прогноза погоды около 1 км., глобальных – 7 км. Для сценарных экспериментов hires smip, предположительно, будут заявлены модели с шагом расчетной сетки менее 10 км.

10. строка 135: «Высокочастотная пространственная изменчивость» – чего?

11. строка 146: «Средняя арифметическая ошибка» – имеется в виду стандартная ошибка среднего?

12. строка 336: Во фразе «позволяет приблизительно аппроксимировать» выделенное слово является лишним.

13. строка 370: Фраза «Как упоминалось выше» является лишней.

14. строки 526–530: Предложение необходимо переформулировать: «между прогнозом и экспертной оценкой с учетом имеющихся у него данных» – у кого «у него»? «хорошее соответствие между прогнозом и соответствующими наблюдениями» соответствующими чему? «относительные преимущества» – относительно чего?

15. строка 539: Во фразе «но смещенные в пространстве» необходимо уточнить: относительно чего происходит смещение.

16. Мнение авторов в разделе заключение не обязательно выделять личными местоимениями.

**Резюме рецензента:** Принять.

**Подпись. Рецензент № 2. 10.12.2024.**

**От редакции:** Высказаны также рекомендации стилистического характера. Они учтены в финальной версии статьи.