

УДК 004.8:551.515

ПРОГНОЗИРОВАНИЕ ДИНАМИКИ ТРОПИЧЕСКОГО ЦИКЛОНА НА ОСНОВЕ МЕТОДОВ ПРОСТРАНСТВЕННО-ВРЕМЕННОЙ ИНТЕГРАЦИИ РАЗНОРОДНЫХ И СЛАБОФОРМАЛИЗОВАННЫХ ДАННЫХ

Шкерин Алексей Леонидович

МИРЭА — Российский технологический университет, г. Москва, Российская Федерация

Аннотация. Тропические циклоны и связанные с ними опасные гидрометеорологические процессы представляют угрозу для прибрежных территорий, морской деятельности, объектов транспортной и энергетической инфраструктуры. В ходе исследования применялись метеорологические, спутниковые и ретроспективные данные, представленные различными форматами и различной структурой. Исследуемая задача сформулирована как бинарное вероятностное прогнозирование события циклогенеза для пространственно-временного окна наблюдений. Сформирован общий набор примеров для горизонтов прогноза 24, 48 и 72 ч, реализован механизм пространственно-временного согласования и интеграции слабоформализованных данных в единый согласованный набор. Проведена оценка результатов работы моделей по ряду метрик: площадь под кривой ошибки (ROC-AUC), площадь под кривой полноты и точности (PR-AUC), индекс Брайера (Brier Score), вероятность обнаружения (POD), критический индекс успеха (CSI) и др. По результатам численного эксперимента сформирован набор из 2 568 пространственно-временных примеров для горизонтов прогноза 24, 48 и 72 ч. Модель из семейства мультимодальных сверточных сетей краткосрочной долгой памяти показала наибольшее значение метрики площади под кривой полноты и точности на горизонте 72 ч (0,517) и критический индекс успеха (0,367). Исследованные модели и разработанные алгоритмы позволили получить воспроизводимые вероятностные оценки циклогенеза на горизонтах прогноза до 72 ч. Эксперимент показал высокую информативность пространственно-временной интеграции слабоформализованных данных, в частности метеорологических, спутниковых и ретроспективных в задачах мониторинга и поддержки принятия решений.

Ключевые слова: безопасность в чрезвычайных ситуациях, тропический циклон, циклогенез, модели на основе искусственного интеллекта, ERA5, мультимодальные модели

Для цитирования: Шкерин А. Л. Прогнозирование динамики тропического циклона на основе методов пространственно-временной интеграции разнородных и слабоформализованных данных // Техносферная безопасность. 2026. № 3 (52). С. 188–206.

FORECASTING THE DYNAMICS OF TROPICAL CYCLONES USING METHODS OF SPATIOTEMPORAL INTEGRATION OF HETEROGENEOUS AND LOOSELY FORMALIZED DATA

Alexey L. Shkerin

MIREA — Russian Technological University, Moscow, Russian Federation

Abstract. Tropical cyclones and the associated hazardous hydrometeorological processes pose a threat to coastal areas, maritime activities, and transportation and energy infrastructure. The study utilized meteorological, satellite, and retrospective data presented in various formats and with varying structures. The research problem is formulated as a binary probabilistic forecast of cyclogenesis for a spatiotemporal observation window. A general set of examples was formed for 24-, 48-, and 72-hour forecast horizons, and a mechanism for spatiotemporal harmonization and integration of weakly formalized data into a single harmonized dataset was implemented. The performance of the models was evaluated using a number of metrics, including Area Under the Receiver Operating Characteristic Curve (ROC-AUC), Area Under the Precision-Recall Curve (PR-AUC), Brier Score, Probability of Detection (POD), Critical Success Index (CSI), and several others. Based on the results of the numerical experiment, a set of 2,568 spatiotemporal examples was formed for forecast horizons of 24, 48, and 72 hours. A model from the family of multimodal Short-Term Long-Memory Convolutional Neural Networks showed the highest Area Under the Precision-Accuracy Curve metric at the 72-hour horizon, equal to 0.517, and the Critical Success Index, equal to 0.367. The investigated models and developed algorithms made it possible to obtain reproducible probabilistic estimates of cyclogenesis for forecast horizons up to 72 hours. The experiment demonstrated the high informational value of spatiotemporal integration of weakly formalized data, in particular meteorological, satellite, and retrospective data, in monitoring and decision support tasks.

Keywords: emergency safety, tropical cyclone, cyclogenesis, AI-based models, ERA5, multimodal models

For citation: Shkerin A. L. Forecasting the dynamics of tropical cyclones using methods of spatiotemporal integration of heterogeneous and loosely formalized data // Technosphere safety. 2026. No. 3 (52). P. 188–206.

Введение

Тропические циклоны являются источниками гидрометеорологических явлений, представляющих потенциальную опасность для критической инфраструктуры и жизней людей, проживающих на приморских территориях. Воздействие подобных природных

явлений проявляется в резком усилении ветра, штормовом нагоне, нарушении транспортного сообщения, повреждении инженерных сооружений и возникновении условий для развития чрезвычайных ситуаций природного характера. В связи с этим задача раннего выявления циклогенеза имеет прикладное значение для обеспечения безопасности

в чрезвычайных ситуациях, предупреждения ущерба и поддержки принятия управленческих решений.

В отечественных исследованиях вопросы анализа тропических циклонов рассматривались в контексте оценки опасности тайфунов, обработки метеорологических данных и влияния тропических циклонов на территорию Дальнего Востока России. В частности, были рассмотрены подходы к анализу накопленной энергии тайфунов при исследовании чрезвычайных ситуаций, обусловленных активностью тропических циклонов, вопросы анализа метео данных с упорядоченным аргументом применительно к задачам техносферной безопасности и особенности эволюции тропических циклонов, оказывавших влияние на погодные условия Дальнего Востока России [1–3].

В отечественной практике предупреждения чрезвычайных ситуаций природного характера прогнозирование опасных гидрометеорологических явлений является межведомственной задачей: результаты специализированных гидрометеорологических наблюдений и краткосрочных прогнозов использовались органами управления РСЧС, территориальными подразделениями МЧС России, дежурно-диспетчерскими службами и службами спасения. Методические рекомендации МЧС России по мониторингу и прогнозированию чрезвычайных ситуаций природного и техногенного характера определяют мониторинг и прогнозирование чрезвычайных ситуаций как элемент заблаговременного выявления угроз, оценки возможных последствий, подготовки решений по предупреждению и ликвидации чрезвычайных ситуаций [4].

Росгидромет и подведомственные ему учреждения выполняют функцию эксперт-

ного источника гидрометеорологической информации, включая сведения о фактических и ожидаемых погодных условиях, опасных и неблагоприятных гидрометеорологических явлениях. К опасным метеорологическим явлениям относятся явления, которые по интенсивности, продолжительности и времени возникновения представляют угрозу безопасности людей и могут нанести значительный ущерб инфраструктуре на месте происшествия.

Для задач обеспечения безопасности в чрезвычайных ситуациях важным является не только получение прогноза циклогенеза как атмосферного процесса, но и преобразование этого прогноза в информационно-аналитический признак возможного формирования источника чрезвычайных ситуаций. В настоящем исследовании рассмотрено применение вероятностной оценки циклогенеза в прикладных задачах, что позволяет предварительно выделить зоны повышенного внимания, требующие сопоставления с официальной информацией служб предупреждения и мониторинга чрезвычайных ситуаций.

Особое значение для задач предупреждения чрезвычайных ситуаций имеют системы раннего предупреждения, поскольку своевременная информация определяет возможность заблаговременного реагирования, уточнения зон возможного воздействия и подготовки сил и средств чрезвычайного реагирования. При этом эффективность раннего предупреждения зависит не только от факта прогнозирования, но и от корректной интерпретации спрогнозированных оценок, их применимости в действующих системах предупреждения чрезвычайных ситуаций и возможности преобразования результатов моделирования

в решения, обеспечивающие снижение риска негативных последствий чрезвычайных ситуаций [4].

Существенное значение для научного обоснования рассматриваемой задачи имели результаты отечественных исследователей Института физики атмосферы им. А. М. Обухова РАН, Института вычислительной математики им. Г. И. Марчука РАН и организаций Росгидромета. В работах, связанных с опасными атмосферными конвективными явлениями, внимание уделялось мезомасштабным конвективным системам, ливневым осадкам, граду, шквалам, смерчам и предикторам, значимым для прогноза опасных явлений.

Исследования Института вычислительной математики им. Г. И. Марчука РАН были связаны с развитием моделей климатической системы, общей циркуляции атмосферы, океана и сопряженных компонентов Земной системы. Результаты, представленные в этих исследованиях, сформировали важную методологическую основу численного моделирования атмосферных процессов, однако их основное назначение заключалось в физико-математическом описании климатической и метеорологической динамики. В настоящем исследовании ключевой задачей является прикладная задача раннего выделения источника чрезвычайных ситуаций и поддержка принятий решений в системе предупреждения чрезвычайных ситуаций [5–7].

Современное развитие методов искусственного интеллекта создает предпосылки для более полного использования пространственно-временной структуры метеорологических и спутниковых данных. В зарубежных исследованиях были предложены модели глубокого обучения для обнаружения тропи-

ческих циклонов, прогнозирования циклогенеза по спутниковым данным и данным реанализа, интерпретации результатов машинного обучения при прогнозировании тропического циклогенеза.

Отдельное направление составляют модели на основе глубокого обучения для прогнозирования траекторий тропических циклонов в северо-западной части Тихого океана. Данные модели были рассмотрены китайскими научными коллективами, в работе R. Zhang была предложена модель глубокого обучения для прогнозирования тропического циклогенеза по сеточным спутниковым данным в бассейне Тихого океана. Дальнейшие китайские исследования развивали это направление в сторону более сложных архитектур, в частности, модель TCGP-Net учитывала взаимодействие океана и атмосферы, а также включала элементы интерпретации результатов через причинный анализ и интегральные градиенты [8, 9].

В рамках проводимого исследования было выяснено, что отличия заключаются в цели применения и обучения подобных моделей, т. к. основное их применение рассматривается в прикладных задачах обеспечения безопасности в чрезвычайных ситуациях. Так же стоит отметить, что обучение моделей, представленных в настоящем исследовании, проводится на интегрированном наборе слабоформализованных и слабоструктурированных мультимодальных данных, что не представлено в работах китайских исследователей.

Полученные в описанных исследованиях результаты показали перспективность нейросетевых моделей в задачах предупреждения чрезвычайных ситуаций, однако выявили проблему согласования

разнородных источников данных, различающихся пространственным разрешением, временной дискретизацией, физическим смыслом переменных и степенью формализации.

Для обработки пространственно-временных последовательностей в метеорологических задачах особую значимость приобрели архитектуры, способные одновременно учитывать пространственную структуру метеорологических полей и их временную динамику. К подобным архитектурам относятся сверточные нейронные сети краткосрочной долгой памяти (ConvLSTM). Для задач предупреждения чрезвычайных ситуаций подобные модели представляют интерес, поскольку признаки формирования циклона проявляются не только в отдельных значениях метеорологических переменных, но и в согласованной эволюции облачно-конвективной структуры, полей давления, скорости ветра, влажности и температуры поверхности моря [11–15].

Следует отметить, что применение моделей глубокого обучения в задачах раннего предупреждения требует не только достижения высокой точности классификации, но и прикладной оценки качества вероятностного прогнозирования. Для практического применения разработанных моделей в системах предупреждения чрезвычайных ситуаций важно правильно определить методы оценки полученных результатов. В связи с этим целесообразно рассматривать в первую очередь метрики, отражающие прикладную полезность разработанной модели, лишь затем оценивать метрики, традиционно используемые в машинном обучении [16].

Материалы и методы

В настоящем исследовании использовали три источника разнородных данных.

Первым источником является последовательность метеорологических полей пятого поколения реанализа Европейского центра среднесрочных прогнозов погоды (ERA5), представляющая согласованные во времени и пространстве поля метеорологических переменных [17]. В ходе исследования были рассмотрены переменные температуры воздуха на высоте 2 м, точка росы на высоте 2 м, давление на уровне моря, компоненты ветра на высоте 10 м и температура поверхности моря.

Вторым источником данных является набор данных яркостной температуры в инфракрасном диапазоне по геостационарному каналу (GridSat-B1), в котором представлены данные геостационарных наблюдений в инфракрасном диапазоне [18].

Последним источником данных является набор данных из Международного архива наилучших треков в области рационального природопользования (IBTrACS), из которого были использованы характеристики: последовательность трека тропического циклона, географические координаты и время циклогенеза [19].

Выбранные данные представляют собой множество мультимодальных данных с разной временной и пространственной дискретизацией, что позволяет рассматривать комплексное состояние циклогенеза в целях раннего выявления чрезвычайных ситуаций.

Рассмотрим множество пространственно-временных окон наблюдений:

$$D = \{(X_i^{sat}, X_i^{era}, m_i, y_{i,H})\}_{i=1}^N, \quad H \in \{24, 48, 72\}, \quad (1)$$

где:

X_i^{sat} — последовательность спутниковых инфракрасных фрагментов из набора данных яркостной температуры в инфракрасном диапазоне по геостационарному каналу (GridSat-B1);

X_i^{era} — последовательность метеорологических полей пятого поколения реанализа Европейского центра среднесрочных прогнозов погоды (ERA5);

m_i — метаданные прогнозного окна, включающие в себя момент времени $t_{0,i}$, координаты, идентификатор события, бассейн зарождения;

$y_{i,H}$ — бинарная метка циклогенеза для горизонта прогноза H .

Бинарная метка определена следующим образом:

$$y_{i,H} = \begin{cases} 1, & 0 < t_i^{gen} - t_{0,i} \leq H, \\ 0, & t_i^{gen} - t_{0,i} > H. \end{cases} \quad (2)$$

где t_i^{gen} — момент циклогенеза, определенный по трековой информации из IBTrACS.

Задача, передаваемая модели, сформирована следующим образом:

$$p_{i,H} = P_\theta(y_{i,H} = 1 | X_i^{sat}, X_i^{era}, m_i, H), \quad (3)$$

где:

P_θ — вероятностная модель с параметрами θ ;

θ — параметры модели.

Для модели, принимающей на вход только спутниковые изображения, вероятность вычисляется как:

$$p_{i,H} = f_{sat}(X_i^{sat}, \theta_{sat}), \quad (4)$$

где:

f_{sat} — модель для спутниковой постановки;

θ_{sat} — обучаемые параметры модели.

Для модели, использующей только метеорологические данные, вероятность вычисляется следующим соотношением:

$$p_{i,H} = f_{era}(X_i^{era}, \theta_{era}), \quad (5)$$

где:

f_{era} — модель для реаналитической постановки;

θ_{era} — обучаемые параметры модели.

Для модели, основанной на мультимодальных данных, использовалась следующая формула:

$$p_{i,H} = f_{fused}(\varphi_{sat}(X_i^{sat}), \varphi_{era}(X_i^{era}), m_i, \theta), \quad (6)$$

где:

$p_{i,H}$ — прогнозная вероятность циклогенеза для i -го пространственно-временного окна на горизонте H ;

f_{fused} — модель для мультимодальной постановки;

$\varphi_{sat}, \varphi_{era}$ — отображения признаков соответствующих модальностей.

Для обучения нейросетевых моделей в качестве функции потерь была использована бинарная кросс-энтропия:

$$L(\theta) = \frac{1}{N_H} \sum_{i=1}^{N_H} [y_{i,H} \log(p_{i,H}) + (1 - y_{i,H}) \log(1 - p_{i,H})], \quad (7)$$

где:

θ — обучаемые параметры модели;

N_H — число обучающих примеров;

$y_{i,H}$ — истинная бинарная метка;

$p_{i,H}$ — предсказанная моделью вероятность положительного класса;

$\varphi_{sat}, \varphi_{era}$ — отображения признаков соответствующих модальностей.

В качестве основной исследуемой архитектуры моделей использовались сверточные сети краткосрочной долгой памяти

(ConvLSTM). В общем виде состояние одной ячейки модели может быть описано системой (8):

$$\begin{cases} i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + b_i) \\ f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + b_f) \\ o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + b_o) \\ \tilde{C} = \text{th}(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c) \\ C_t = f_t \circ C_{t-1} + i_t \circ \tilde{C}_{t-1} \\ H_t = o_t \circ \text{th}(C_t) \end{cases} \quad (8)$$

где:

i_t, f_t, o_t — входной, забывающий и выходной вентиля;

σ — сигмоидальная функция активации;

W_{x*}, W_{h*} — обучаемые сверточные ядра для входа и скрытого состояния;

$*$ — операция свертки;

X_t — входной пространственный тензор в момент времени t ;

H_t — скрытое состояние;

H_{t-1} — скрытое состояние на предыдущем временном шаге;

b_i, b_f, b_o, b_c — обучаемые смещения;

th — гиперболический тангенс;

C_t — состояние памяти;

$\circ$ — операция поэлементного произведения;

C_{t-1} — состояние памяти на предыдущем временном шаге;

$\tilde{C}_t$ — кандидатное состояние памяти.

Важными аспектами в применении описанной модели является воспроизводимость и достоверность полученных результатов, для чего в исследовании рассматривали следующие метрики верификации результатов: площадь под кривой ошибки (ROC-AUC), площадь под кривой полноты и точности (PR-AUC), индекс Брайера (Brier Score), ожидаемая ошибка калибровки (ECE), точность (Precision), полнота (Recall), F1-мера, вероятность обнаружения (POD), коэффициент ложных тревог (FAR), критический индекс успеха (CSI). Используемый при верификации набор исследуемых метрик позволяет удостовериться в точности полученных результатов с разносторонней оценкой. Обучение и проверку результатов моделей повторили несколько раз для установления статистической устойчивости полученных результатов.

Для преобразования рассчитанной вероятности к бинарной оценке использовали заданный порог τ_n , значение которого варьировали в зависимости от значения F1-меры на валидационной выборке.

Прикладной смысл введенного порога состоит в выделении окон, требующих повышенного внимания в системе мониторинга.

Для оперативного применения он может быть скорректирован в соответствии с требуемым значением допустимого уровня ложных тревог и требованиями конкретных метеорологических служб.

В актуальных системах мониторинга результат работы модели может быть представлен в виде формализованного сообщения для системы поддержки принятия решений: горизонт прогноза H , ч, координаты пространственного интервала, сектор потенциального воздействия, значения прогнозируемой вероятности, уровень риска. Подобная структура позволила использовать модель как вспомогательный прогнозно-аналитический компонент, дополняющий действующие процедуры мониторинга и прогнозирования чрезвычайных ситуаций природного характера.

Для расчета предложенных метрик использовали результаты бинарной классификации, определенные через количество истинно положительных (TP), ложноположительных (FP), истинно отрицательных (TN) и ложноотрицательных (FN) решений. Из данных значений сформировали метрики точности, полноты и F1-меры:

$$\begin{aligned} \text{Precision} &= \frac{TP}{TP + FP} \\ \text{Recall} &= \frac{TP}{TP + FN} \\ F1 &= \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \end{aligned} \quad (9)$$

Для задач раннего предупреждения ключевыми метриками являются POD, FAR и CSI; они задаются формулами (10):

$$\begin{aligned} \text{POD} &= \frac{TP}{TP + FN} \\ \text{FAR} &= \frac{FP}{TP + FP} \\ \text{CSI} &= \frac{TP}{TP + FN + FP}. \end{aligned} \quad (10)$$

Метрика POD характеризует долю обнаруженных опасных событий, FAR — долю ложных тревог среди выданных предупреждений, CSI — общий индекс успешности обнаружения с учетом пропусков и ложных тревог.

Данный набор метрик позволил корректно оценить пригодность разработанных моделей к предупреждению чрезвычайных ситуаций.

Для оценки вероятностной точности был использован индекс Брайера:

$$BS = \frac{1}{N} \sum_{i=1}^N (p_i - y_i)^2. \quad (11)$$

Калибровка вероятностей оценивается через метрику ECE:

$$ECE = \sum_{b=1}^B \frac{|I_b|}{N} |acc(I_b) - conf(I_b)|, \quad (12)$$

где:

I_b — множество примеров, попавших в интервал вероятности b ;

$acc(I_b)$ — фактическая частота положительных событий в интервале;

$conf(I_b)$ — средняя предсказанная вероятность.

Основной областью исследуемых источников являются регионы в интервале долгот 110–180° в. д. и интервале широт 0–45° с. ш. В этой области проводилась пространственно-временная интеграция использованных данных.

Отдельно выделена область потенциального воздействия на Дальневосточный регион. Для секторной диагностики обозначили следующие районы Дальневосточного региона:

- 1) Приморье,
- 2) Японское море,

- 3) Охотское море,
- 4) юг Камчатки,
- 5) Курило-Тихоокеанский сектор.

При попадании события в несколько секторов применяется фиксированный приоритет, который в настоящей работе определен согласно приведенному нумерованному списку.

На рис. 1 представлено графическое отображение принятой схемы интервалов и размечены границы исследуемых регионов.

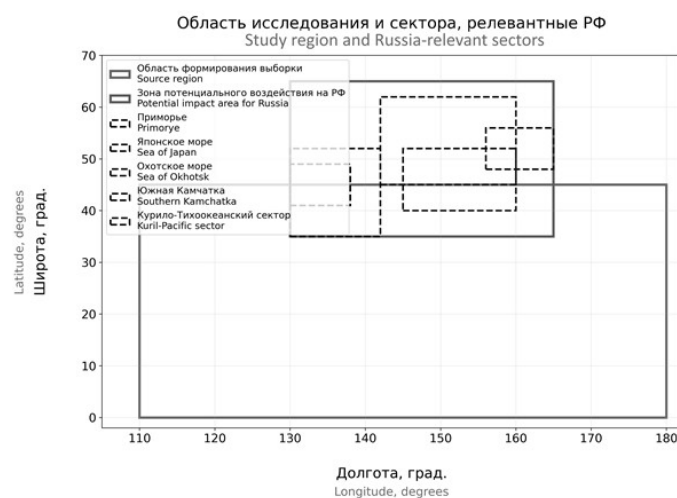


Рис. 1. Схема географической области исследования

Fig. 1. Map of the study area

В результате применения алгоритмов пространственно-временной интеграции сформировано 2 568 базовых пространственно-временных примеров. Для горизонтов прогноза 24, 48 и 72 ч сформирован набор из 7 704 строк, в исследуемой выборке отмечено 1 284 положительных события и 6 420 отрицательных.

Результаты и их обсуждение

В табл. 1 приведены основные результаты верификации моделей на тестовой выборке.

Полученные результаты показывают, что модели глубокого обучения превосходят традиционный метод логистической регрессии по большинству ранжирующих и вероятностных метрик. На горизонтах прогноза 24 и 48 ч лучшие значения ROC-AUC и PR-AUC демонстрирует исследуемая модель, обученная исключительно на данных реанализа ERA5. На горизонте 72 ч наибольший PR-AUC равен 0,517 и достигается при использовании исследуемой модели, обученной на пространственно-временной интеграции исследуемых данных.

Таблица 1
Значения метрик исследуемых моделей

Table 1
Metrics for the models under study

Модель Model	ROC-AUC 24/48/72 ч ROC-AUC 24/48/72 h	PR-AUC 24/48/72 ч PR-AUC 24/48/72 h	BS 24/48/72 ч BS 24/48/72 h	ECE 24/48/72 ч ECE 24/48/72 h	Доля обнаружений Detection rate
Логистическая регрессия исключительно на данных ERA5 Logistic regression trained exclusively on ERA5 data	0,798/ 0,781/ 0,819	0,145/ 0,248/ 0,368	0,039/ 0,071/ 0,097	0,003/ 0,029/ 0,045	0,615
Сверточная сеть исключительно на данных GridSat-B1 Convolutional network trained exclusively on GridSat-B1 data	0,887/ 0,866/ 0,852	0,238/ 0,395/ 0,444	0,035/ 0,065/ 0,088	0,007/ 0,015/ 0,038	0,677
Предложенная модель, обученная исключительно на данных GridSat-B1 Proposed model trained exclusively on GridSat-B1 data	0,881/ 0,866/ 0,852	0,183/ 0,372/ 0,429	0,038/ 0,065/ 0,089	0,015/ 0,020/ 0,034	0,625
Предложенная модель, обученная исключительно на данных ERA5 Proposed model trained exclusively on ERA5 data	0,889/ 0,881/ 0,867	0,317/ 0,469/ 0,498	0,034/ 0,058/ 0,082	0,014/ 0,026/ 0,019	0,771
Предложенная модель, обученная на интегрированном наборе данных Proposed model trained on integrated data	0,876/ 0,857/ 0,874	0,232/ 0,433/ 0,517	0,036/ 0,061/ 0,082	0,005/ 0,027/ 0,032	0,708

На рис. 2 представлено графическое отображение метрики ROC-AUC для исследуемых моделей.

На рис. 3 представлено графическое отображение метрики PR-AUC для исследуемых моделей.

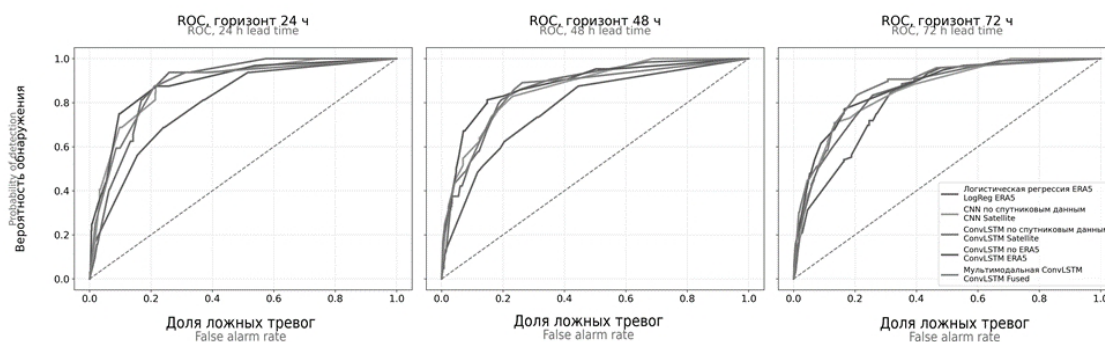


Рис. 2. ROC-кривые моделей на тестовой выборке для горизонтов прогноза 24, 48 и 72 ч. Сокращения, принятые в легенде: LogReg ERA5 — логистическая регрессия; CNN Satellite — сверточная сеть на спутниковых данных; ConvLSTM — предложенные модели на спутниковых, метеорологических и интегрированных данных

Fig. 2. ROC curves of the models on the test dataset for 24-, 48-, and 72-hour forecast horizons. Abbreviations used in the legend: LogReg ERA5 — logistic regression; CNN Satellite — Convolutional Neural Network based on satellite data; ConvLSTM — proposed models based on satellite, meteorological, and integrated data

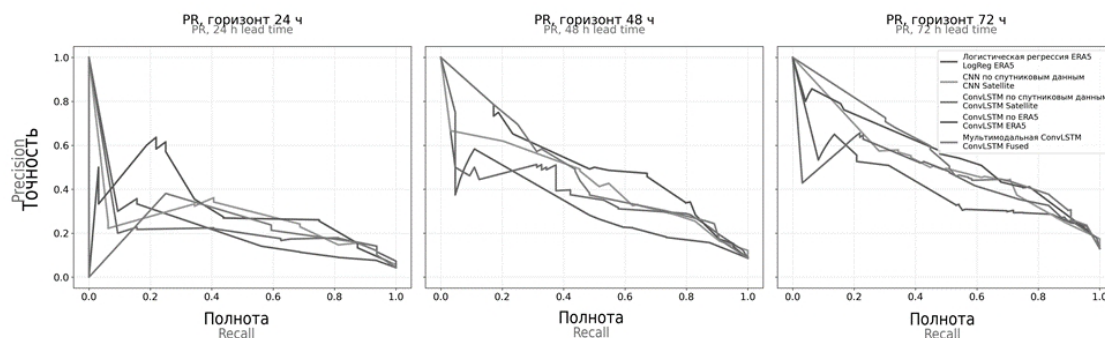


Рис. 3. PR-кривые моделей на тестовой выборке для горизонтов прогноза 24, 48 и 72 ч. Сокращения, принятые в легенде: LogReg ERA5 — логистическая регрессия; CNN Satellite — сверточная сеть на спутниковых данных; ConvLSTM — предложенные модели на спутниковых, метеорологических и интегрированных данных

Fig. 3. PR curves of the models on the test dataset for forecast horizons of 24, 48, and 72 hours. Abbreviations used in the legend: LogReg ERA5 — logistic regression; CNN Satellite — Convolutional Neural Network based on satellite data; ConvLSTM — proposed models based on satellite, meteorological, and integrated data

Из данных табл. 1, рис. 2 и 3 следует вывод о превосходстве предложенных моделей над традиционными. Для раннего преду-

ждения чрезвычайных ситуаций более информативны операционные метрики, показатели которых сведены в табл. 2.

Таблица 2
Значения операционных метрик для исследуемых моделей

Table 2
Values of operational metrics for the models under study

Модель Model	POD 24/48/72 ч 0,562/0,484/0,542	FAR 24/48/72 ч FAR 24/48/72 h	CSI 24/48/72 ч CSI 24/48/72 h
Логистическая регрессия исключительно на данных ERA5 Logistic regression trained exclusively on ERA5 data	0,562/ 0,484/ 0,542	0,859/ 0,718/ 0,673	0,127/ 0,217/ 0,256

Окончание таблицы 2

Модель Model	POD 24/48/72 ч 0,562/0,484/0,542	FAR 24/48/72 ч FAR 24/48/72 h	CSI 24/48/72 ч CSI 24/48/72 h
Сверточная сеть исключительно на данных GridSat-B1 Convolutional network trained exclusively on GridSat-B1 data	0,406/ 0,547/ 0,698	0,658/ 0,573/ 0,559	0,228/ 0,315/ 0,370
Предложенная модель, обученная исключительно на данных GridSat-B1 Proposed model trained exclusively on GridSat-B1 data	0,406/ 0,422/ 0,594	0,780/ 0,620/ 0,584	0,167/ 0,250/ 0,324
Предложенная модель, обученная исключительно на данных ERA5 Proposed model trained exclusively on ERA5 data	0,750/ 0,547/ 0,750	0,742/ 0,514/ 0,593	0,238/ 0,347/ 0,358
Предложенная модель, обученная на интегрированном наборе данных Proposed model trained on integrated data	0,594/ 0,484/ 0,677	0,765/ 0,617/ 0,555	0,202/ 0,272/ 0,367

Для приведения полученных результатов к прикладным значениям в задаче обеспечения безопасности в чрезвычайных ситуациях было выполнено сопоставление предложенной методики с существующими подходами, применяемыми в отечественной и зарубежной практике прогнози-

рования чрезвычайных ситуаций. Сравнение проводилось с открытыми и известными моделями и информационными технологиями, используемыми в задачах предупреждения чрезвычайных ситуаций природного характера, сравнение представлено в табл. 3.

Таблица 3
Сравнение разработанных моделей с актуальными аналогами

Table 3
Comparison of developed model with actual analogues

Класс подходов Solution class	Назначение Usage	Сильные стороны Advances	Ограничения Limitations
Синоптические прогнозы Росгидромета Roshydromet's weather forecasts	Подготовка официальной информации о гидрометеорологическом явлении Preparation of official information on a hydrometeorological phenomenon	Нормативная информация, экспертная проверка, интегрировано в современные системы Regulatory information and expert review, integrated into modern systems	Не предоставляет вероятностного прогнозирования по каждому пространственно-временному окну Does not provide probabilistic forecasts for each spatiotemporal window
Численные гидродинамические модели атмосферы Numerical hydrodynamic models of the atmosphere	Физически обоснованное прогнозирование атмосферной динамики Physically-Based Forecasting of Atmospheric Dynamics	Высокая физическая интерпретируемость High physical interpretability	Высокая вычислительная сложность, чувствительность к начальным условиям High computational complexity, sensitivity to initial conditions

Окончание таблицы 3

Класс подходов Solution class	Назначение Usage	Сильные стороны Advances	Ограничения Limitations
Статистические модели Statistical models	Оценка связи и признаков чрезвычайной ситуации Assessment of communication and signs of an emergency	Простота, интерпретируемость Simplicity, interpretability	Не учитывают временной динамики и метеорологических полей Do not take into account temporal dynamics and meteorological fields
Сверточные нейронные сети на спутниковых данных CNN applied to satellite data	Выявление облачно-конвективных признаков Identification of cloud-convective features	Эффективная обработка спутниковых изображений Efficient processing of satellite images	Ограниченный учет временной динамики и метеорологических полей Limited consideration of temporal dynamics and meteorological fields
Предложенная модель Proposed model	Вероятностная оценка риска циклогенеза на горизонтах прогнозирования 24, 48 и 72 ч Probabilistic assessment of cyclogenesis risk at 24-, 48-, and 72-hour forecast horizons	Интеграция слабоформализованных и слабоструктурированных данных, формирование прикладных оценок Integration of weakly formalized and weakly structured data; generation of applied estimates	Не предоставляет всесторонний прогноз, требует калибровки под конкретный регламент применения Does not provide a comprehensive forecast, requires calibration for specific application guidelines

Таким образом, разработанная модель решает прикладную задачу, недостаточно раскрытую в классических метеорологических методах, связанных с предупреждением чрезвычайных ситуаций. Наиболее устойчивые значения метрик показывает предложенная модель, обученная только на данных ERA5, значения метрики POD для нее составило 0,75, предложенная модель, обученная на интегрированных данных, показала значения CSI 0,367, что является высоким показателем для исследуемой области.

Таким образом, по данным табл. 1–3 сделаны выводы о превосходстве предложенных моделей среди исследуемых моделей по ряду важных метрик. Примечательными являются частные случаи превосходства модели, обученной только на данных реанализа, над моделью, обученной на интегрированных данных.

Для оценки статистической устойчивости были проведены повторные запуски обучения и тестирования исследуемых моделей. На горизонте прогноза 72 ч для предложенной модели, обученной только на данных реанализа, метрика PR-AUC составила 0,505 с 95-процентным доверительным интервалом [0,389; 0,599], для модели, обученной на интегрированных данных, значение той же метрики составило 0,521 с 95-процентным доверительным интервалом [0,383; 0,621]. Для метрики ROC-AUC для тех же моделей на том же горизонте прогноза получены значения 0,868 и 0,875 с 95-процентным доверительным интервалом, [0,829; 0,913] и [0,834; 0,922], для моделей, обученных только на данных реанализа и на интегрированных данных соответственно.

Для Дальневосточного региона выборка данных содержала 77 окон на каждый прогнозный горизонт и 9 событий циклогенеза,

результаты исследования моделей приведены в табл. 4.

На исследуемой выборке наибольшую вероятность обнаружения так же получила предложенная модель, обученная только на спутниковых данных: модель показала наивысший результат на горизонте 72 ч, что говорит о высокой значимости и информативности облачно-конвективной струк-

туры при анализе отдельных российских регионов.

Результаты калибровки вероятности показывают улучшение согласования прогнозных вероятностей с наблюдаемыми частотами. Для предложенной модели на горизонте прогноза 72 ч значения метрик BS и ECE улучшились, результаты калибровки представлены на диаграмме достоверности (рис. 4).

Таблица 4

Значения метрики PR-AUC для Дальневосточного региона

Table 4

PR-AUC metric values for the Far East region

Модель Model	PR-AUC 24 ч PR-AUC 24 h	PR-AUC 48 ч PR-AUC 48 h	PR-AUC 72 ч PR-AUC 72 h	Доля обнаружений Detection rate
Логистическая регрессия исключительно на данных ERA5 Logistic regression trained exclusively on ERA5 data	0,390	0,647	0,898	0,556
Сверточная сеть исключительно на данных GridSat-B1 Convolutional network trained exclusively on GridSat-B1 data	0,349	0,709	0,905	0,593
Предложенная модель, обученная исключительно на данных GridSat-B1 Proposed model trained exclusively on GridSat-B1 data	0,380	0,688	0,960	0,630
Предложенная модель, обученная исключительно на данных ERA5 Proposed model trained exclusively on ERA5 data	0,515	0,716	0,889	0,704
Предложенная модель, обученная на интегрированном наборе данных Proposed model trained on integrated data	0,466	0,692	0,911	0,667

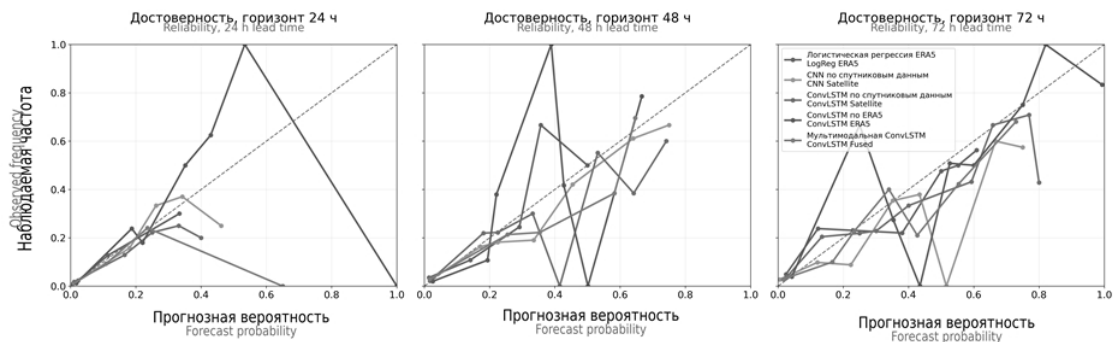


Рис. 4. Диаграммы достоверности исследуемых моделей на тестовой выборке для горизонтов прогноза 24, 48 и 72 ч. Сокращения, принятые в легенде: LogReg ERA5 — логистическая регрессия; CNN Satellite — сверточная сеть на спутниковых данных; ConvLSTM — предложенные модели на спутниковых, метеорологических и интегрированных данных

Fig. 4. Confidence plots for the models under study on the test dataset for 24-, 48-, and 72-hour forecast horizons. Abbreviations used in the legend: LogReg ERA5 — logistic regression; CNN Satellite — Convolutional Neural Network based on satellite data; ConvLSTM — proposed models based on satellite, meteorological, and integrated data

На рис. 4. представлено, что наиболее достоверными результатами обладают предложенные модели, обученные на данных ре-анализа и на интегрированных данных.

Заключение

В ходе исследования были разработаны модели с пространственно-временной интеграцией слабоформализованных и слабоструктурированных данных для вероятностной оценки риска тропического циклогенеза на горизонтах 24, 48 и 72 ч. Предложенные модели были рассмотрены в прикладном контексте обеспечения безопасности в чрезвычайных ситуациях: как инструмент раннего выявления чрезвычайной ситуации и информационно-аналитической поддержки принятия решений.

В результате применения алгоритмов пространственно-временной интеграции был сформирован набор из 2 568 пространственно-временных интервалов. Для исследуемых горизонтов прогнозирования 24, 48 и 72 ч был получен общий набор из 7 704 строк, включавший 1 284 положительных и 6 420 отрицательных примеров. Полученные результаты показали, что нейросетевые модели оказались эффективнее по ряду метрик, ключевых для оценки применимости моделей в задачах раннего предупреждения чрезвычайных ситуаций. На горизонте 72 ч предложенная модель, обученная на интегрированном наборе данных, показала наибольшее значение метрики PR-AUC, равное 0,517, и высокое значение метрики CSI, равное 0,367. Данные результаты указывают на прикладную значимость и применимость разработанной модели в актуальных системах мониторинга и предупреждения чрезвычайных ситуаций.

Проведенный эксперимент продемонстрировал, что модель показала устойчивые и воспроизводимые результаты на горизонтах прогнозирования 24 и 48 ч, на горизонте прогнозирования 72 ч достоверность результата уменьшается согласно увеличению неопределенности данных.

Полученные в ходе проведенного исследования результаты позволяют сформулировать ряд важных выводов.

Разработанные в ходе исследования алгоритмы показали высокие результаты по метрикам, важным для оценки пригодности моделей в задаче предупреждения чрезвычайных ситуаций. Важно отметить, что предложенная модель обладает рядом преимуществ: объединение слабоформализованных и слабоструктурированных данных, формирование вероятностной оценки циклогенеза, прикладная интерпретация результатов, высокие результаты сценарной апробации на регионах Дальнего Востока.

Стоит отметить ограничения и недостатки модели: необходимость дополнительной проверки устойчивости на больших горизонтах прогнозирования, зависимость итоговой оценки от определяемого порога, необходимость использования большого количества данных для получения разносторонней информации о чрезвычайной ситуации, отсутствие готовых решений для интеграции разработанной модели в актуальные системы мониторинга и предупреждения чрезвычайных ситуаций.

Практическая значимость 72-часового прогнозирования риска природных опасных явлений заключается в возможности заблаговременного выделения территорий, требующих повышенного внимания со стороны органов управления, дежурно-

диспетчерских служб и служб спасения. В организационном отношении разработанные модели могут быть использованы для предварительной мобилизации резервов материально-технических ресурсов, оценки возможной нагрузки на транспортную, энергетическую и прибрежную инфраструктуру.

В условиях современного этапа комплексных и гибридных угроз, включая сценарии преднамеренного воздействия на природно-технические системы, в частности технологии искусственного воздействия

на гидрометеорологические процессы, значение заблаговременного прогнозно-аналитического обеспечения возрастает.

Дальнейшие исследования целесообразно направить на сопряжение разработанных моделей с геоинформационными слоями уязвимости территорий, данными о населении, критической инфраструктуры. Такое развитие позволит перейти от вероятностной оценки циклогенеза к всесторонней оценке риска чрезвычайной ситуации и повысить прикладную значимость прогнозируемых значений.

СПИСОК ИСТОЧНИКОВ

1. Малыгин И. Г., Калач А. В., Парамонов А. А. Почти-периодический анализ накопленной энергии тайфунов при исследовании развития чрезвычайных ситуаций, обусловленных активностью тропических циклонов // Проблемы управления рисками в техносфере. 2024. № 3 (71). С. 55–62. DOI: 10.61260/1998-8990-2024-3-55-62. EDN: IBPLCG.
2. Калач А. В., Парамонов А. А., Крынецкий Б. А. К вопросу анализа метеоданных с упорядоченным аргументом в аспекте решения задач техносферной безопасности // Технологии техносферной безопасности. 2024. № 3(105). С. 170–180. DOI: 10.25257/TTS.2024.3.105.170-180. EDN: HDQEJN.
3. Озерова Н. А. Исследование особенностей эволюции тропических циклонов, влияющих на погоду Дальнего Востока России // European Journal of Natural History. 2023. № 1. С. 29–34. EDN: ETWEOW.
4. Методические рекомендации МЧС России по мониторингу и прогнозированию чрезвычайных ситуаций природного и техногенного характера / под ред. К. Д. Моськина. М., 2024. 168 с. URL: <https://clck.ru/aBgfJ> (дата обращения: 10.07.2026).
5. Towards the accurate automatic detection of mesoscale convective systems in remote sensing data: from data mining to deep learning models and their applications / M. Krinitskiy [et al.] // Remote Sensing. 2023. Vol. 15. No. 14. Art. 3493. DOI: 10.3390/rs15143493.
6. Система краткосрочного численного прогноза высокой детализации COSMO-Ru, ее развитие и приложения / Г. С. Ривин [и др.] // Гидрометеорологические исследования и прогнозы. 2019. № 4 (374). С. 37–53.
7. Воспроизведение современного климата в новой версии модели климатической системы ИВМ РАН / Е. М. Володин [и др.] // Известия РАН. Физика атмосферы и океана. 2017. Т. 53, № 2. С. 164–178. DOI: 10.7868/S0002351517020122.

8. Predicting tropical cyclogenesis using a deep learning method from gridded satellite and ERA5 reanalysis data in the Western North Pacific Basin / R. Zhang [et al.] // *IEEE Transactions on Geoscience and Remote Sensing*. 2022. Vol. 60. P. 1–10. DOI: 10.1109/TGRS.2021.3069217.
9. A deep learning-based global tropical cyclogenesis prediction model and its interpretability analysis / B. Mu [et al.] // *Science China Earth Sciences*. 2024. Vol. 67. P. 3671–3695. DOI: 10.1007/s11430-023-1383-6.
10. Alcántara-Ayala I., Oliver-Smith A. Early Warning Systems: Lost in Translation or Late by Definition? A FORIN Approach // *International Journal of Disaster Risk Science*. 2019. Vol. 10. P. 317–331. DOI: 10.1007/s13753-019-00231-3.
11. Deep learning approach for detecting tropical cyclones and their precursors in the simulation by a cloud-resolving global nonhydrostatic atmospheric model / D. Matsuoka [et al.] // *Progress in Earth and Planetary Science*. 2018. Vol. 5. Art. 80. DOI: 10.1186/s40645-018-0245-y.
12. Loi C. L., Wu C.-C., Liang Y.-C. Prediction of Tropical Cyclogenesis Based on Machine Learning Methods and Its SHAP Interpretation // *Journal of Advances in Modeling Earth Systems*. 2024. Vol. 16. Art. e2023MS003637. DOI: 10.1029/2023MS003637.
13. A Review on the Application of Machine Learning Methods in Tropical Cyclone Forecasting / Z. Wang [et al.] // *Frontiers in Earth Science*. 2022. Vol. 10. Art. 902596. DOI: 10.3389/feart.2022.902596.
14. Forecasting tropical cyclone tracks in the northwestern Pacific based on a deep-learning model / L. Wang [et al.] // *Geoscientific Model Development*. 2023. Vol. 16. P. 2167–2179. DOI: 10.5194/gmd-16-2167-2023.
15. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting / X. Shi [et al.] // *Advances in Neural Information Processing Systems*. 2015. Vol. 28. P. 802–810.
16. Loveday N., Taggart R., Khanarmuei M. A User-Focused Approach to Evaluating Probabilistic and Categorical Forecasts // *Weather and Forecasting*. 2024. Vol. 39. No. 8. P. 1163–1180. DOI: 10.1175/WAF-D-23-0201.1.
17. The ERA5 global reanalysis / H. Hersbach [et al.] // *Quarterly Journal of the Royal Meteorological Society*. 2020. Vol. 146. No. 730. P. 1999–2049. DOI: 10.1002/qj.3803.
18. Globally gridded satellite observations for climate studies / K. R. Knapp [et al.] // *Bulletin of the American Meteorological Society*. 2011. Vol. 92. No. 7. P. 893–907. DOI: 10.1175/2011BAMS3039.1.
19. The International Best Track Archive for Climate Stewardship (IBTrACS): Unifying tropical cyclone best track data / K. R. Knapp [et al.] // *Bulletin of the American Meteorological Society*. 2010. Vol. 91. No. 3. P. 363–376. DOI: 10.1175/2009BAMS2755.1.

REFERENCES

1. Malygin I. G., Kalach A. V., Paramonov A. A. Near-Periodic Analysis of Typhoon Accumulated Energy for Emergencies Caused by Tropical Cyclone Activity // *Risk Management Problems in the Technosphere*. 2024. No. 3 (71). P. 55–62. DOI: 10.61260/1998-8990-2024-3-55-62. EDN: IBPLCG.
2. Kalach A. V., Paramonov A. A., Krynetsky B. A. On the Analysis of Meteorological Data with an Ordered Argument for Solving Technosphere Safety Problems // *Technosphere Safety Technologies*. 2024. No. 3 (105). P. 170–180. DOI: 10.25257/TTS.2024.3.105.170-180. EDN: HDQEJN.

3. Ozerova N. A. Evolution of Tropical Cyclones Affecting the Russian Far East Weather // *European Journal of Natural History*. 2023. No. 1. P. 29–34. EDN: ETWEOW.
4. *Methodological Recommendations of EMERCOM of Russia on Monitoring and Forecasting Natural and Technogenic Emergencies* / edited by K. D. Moskin. Moscow, 2024. 168 p. URL: <https://clck.su/aBgfJ> (accessed 10.07.2026).
5. Towards the accurate automatic detection of mesoscale convective systems in remote sensing data: from data mining to deep learning models and their applications / M. Krinitskiy [et al.] // *Remote Sensing*. 2023. Vol. 15. No. 14. Art. 3493. DOI: 10.3390/rs15143493.
6. COSMO-Ru High-Resolution Short-Range Numerical Weather Prediction System: Its Development and Applications / G. S. Rivin [et al.] // *Hydrometeorological Research and Forecasting*. 2019. No. 4 (374). P. 37–53.
7. Simulation of Modern Climate with the New Version of the INM RAS Climate Model / E. M. Volodin // *Izvestiya, Atmospheric and Oceanic Physics*. 2017. Vol. 53. No. 2. P. 164–178. DOI: 10.7868/S0002351517020122.
8. Predicting tropical cyclogenesis using a deep learning method from gridded satellite and ERA5 reanalysis data in the Western North Pacific Basin / R. Zhang [et al.] // *IEEE Transactions on Geoscience and Remote Sensing*. 2022. Vol. 60. P. 1–10. DOI: 10.1109/TGRS.2021.3069217.
9. A deep learning-based global tropical cyclogenesis prediction model and its interpretability analysis / B. Mu [et al.] // *Science China Earth Sciences*. 2024. Vol. 67. P. 3671–3695. DOI: 10.1007/s11430-023-1383-6.
10. Alcántara-Ayala I., Oliver-Smith A. Early Warning Systems: Lost in Translation or Late by Definition? A FORIN Approach // *International Journal of Disaster Risk Science*. 2019. Vol. 10. P. 317–331. DOI: 10.1007/s13753-019-00231-3.
11. Deep learning approach for detecting tropical cyclones and their precursors in the simulation by a cloud-resolving global nonhydrostatic atmospheric model / D. Matsuoka [et al.] // *Progress in Earth and Planetary Science*. 2018. Vol. 5. Art. 80. DOI: 10.1186/s40645-018-0245-y.
12. Loi C. L., Wu C.-C., Liang Y.-C. Prediction of Tropical Cyclogenesis Based on Machine Learning Methods and Its SHAP Interpretation // *Journal of Advances in Modeling Earth Systems*. 2024. Vol. 16. Art. e2023MS003637. DOI: 10.1029/2023MS003637.
13. A Review on the Application of Machine Learning Methods in Tropical Cyclone Forecasting / Z. Wang [et al.] // *Frontiers in Earth Science*. 2022. Vol. 10. Art. 902596. DOI: 10.3389/feart.2022.902596.
14. Forecasting tropical cyclone tracks in the northwestern Pacific based on a deep-learning model / L. Wang [et al.] // *Geoscientific Model Development*. 2023. Vol. 16. P. 2167–2179. DOI: 10.5194/gmd-16-2167-2023.
15. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting / X. Shi [et al.] // *Advances in Neural Information Processing Systems*. 2015. Vol. 28. P. 802–810.
16. Loveday N., Taggart R., Khanarmuei M. A User-Focused Approach to Evaluating Probabilistic and Categorical Forecasts // *Weather and Forecasting*. 2024. Vol. 39. No. 8. P. 1163–1180. DOI: 10.1175/WAF-D-23-0201.1.
17. The ERA5 global reanalysis / H. Hersbach [et al.] // *Quarterly Journal of the Royal Meteorological Society*. 2020. Vol. 146. No. 730. P. 1999–2049. DOI: 10.1002/qj.3803.

18. Globally gridded satellite observations for climate studies / K. R. Knapp [et al.] // Bulletin of the American Meteorological Society. 2011. Vol. 92. No. 7. P. 893–907. DOI: 10.1175/2011BAMS3039.1.

19. The International Best Track Archive for Climate Stewardship (IBTrACS): Unifying tropical cyclone best track data / K. R. Knapp [et al.] // Bulletin of the American Meteorological Society. 2010. Vol. 91. No. 3. P. 363–376. DOI: 10.1175/2009BAMS2755.1.

ИНФОРМАЦИЯ ОБ АВТОРЕ

Шкерин Алексей Леонидович, ассистент кафедры прикладной математики Института информационных технологий МИРЭА — Российского технологического университета (119454, Российская Федерация, г. Москва, просп. Вернадского, д. 78); РИНЦ ID: 1212098; Researcher ID: 154232; ORCID: 0009-0004-2839-4731; e-mail: le.hkerin@yandex.ru

INFORMATION ABOUT THE AUTHOR

Alexey L. Shkerin, Assistant at the Department of Applied Mathematics of Institute of Information Technologies, MIREA — Russian Technological University (78 Vernadsky Avenue, Moscow, 119454, Russian Federation); ID RSCI: 1212098; Researcher ID: 154232; ORCID: 0009-0004-2839-4731; e-mail: le.hkerin@yandex.ru