frontiers in Climate

TYPE Original Research
PUBLISHED 30 March 2026
DOI 10.3389/fclim.2026.1779183

Check for updates

OPEN ACCESS

EDITED BY
Swadhin Kumar Behera,
Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Japan

REVIEWED BY
Dorian Abbot,
The University of Chicago, United States
Josep Ramon Miró,
Servei Meteorologic de Catalunya, Spain

*CORRESPONDENCE
Yury Maximov
yury@ieee.org

RECEIVED 01 January 2026
REVISED 23 February 2026
ACCEPTED 25 February 2026
PUBLISHED 30 March 2026

CITATION
Mozikov M, Taniushkina D, Bulkin A, Liu Y, Sotiriadi N, Osiptsov A, Sultimov R, Makarov I and Maximov Y (2026) Neural network-based approach to evaluate convective hazards frequency shift under climate change. *Front. Clim.* 8:1779183. doi: 10.3389/fclim.2026.1779183

COPYRIGHT
© 2026 Mozikov, Taniushkina, Bulkin, Liu, Sotiriadi, Osiptsov, Sultimov, Makarov and Maximov. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Neural network-based approach to evaluate convective hazards frequency shift under climate change

Mikhail Mozikov , Daria Taniushkina , Alexander Bulkin , Yuanchao Liu , Nazar Sotiriadi , Andrey Osiptsov , Roman Sultimov , Ilya Makarov and Yury Maximov *

This study addresses challenges in predicting extreme weather events, including intense rainfall, hail, and strong winds. Such events pose significant threats to infrastructure and human life, and their frequency has been increasing due to climate change. We show, that integrating climate forecasts with modern machine learning techniques improves prediction accuracy and helps identify regions where these events may become more frequent and dangerous. To achieve reliable predictions, we propose a robust neural network architecture that outperforms several common baselines in accuracy and reliability. Our model leverages problem-specific physics encapsulated in Coupled Model Intercomparison Project data. The analysis also highlight the impact of rugged terrain on the risk distribution of extreme weather events. The proposed model has significant potential to support mid-term adaptation strategies in response to climate change.

KEYWORDS
climate change, convective storm, machine learning, natural hazards, risk management

1 Introduction

Severe convective storms (SCS) are characterized by intense thunderstorms and can produce tornadoes, strong winds, hail, and heavy rainfall. Since 2020, the U.S. has incurred $146.2 billion in losses from such storm-related phenomena [(NOAA National Centers for Environmental Information (NCEI), 2024)], similar impacts are reported globally (Frame et al., 2020; Kron et al., 2019; Lesk et al., 2016).

Understanding changes in the frequency of SCS and related hazards is essential to minimize their harmful effects. Studies indicate that rising global temperatures may influence the behavior and intensity of convective storms (IPCC, 2022; Raupach et al., 2021), making accurate forecasting critical for infrastructure preparedness, urban planning, and improved disaster response (IPCC, 2022).

Machine learning (ML) have great potential for providing accurate forecasts, but the success strongly depends on data quality and volume. On one hand, collecting ground-truth events for analysis of SCS-related hazards is challenging due to the short duration of these

phenomena, the need for high-resolution observations, and limitations of established country-level data collection pipelines (Zabolotnikov, 2010). When working with observations alone, it is also important to consider potential population and clustering biases (Elmore et al., 2022). Although data collection methods and tools have improved significantly in recent years, obtaining historical event datasets with long time spans remains extremely difficult. On the other hand, predictors for SCS-associated hazards are often readily available, as climate modeling has received substantial attention in recent years. A particularly notable initiative in this area is the Coupled Model Intercomparison Project (CMIP), which offers a diverse range of models that facilitate assessment of essential climate indicators; however, the accuracy of these predictors remain limited.

Machine learning is resource-intensive, both in terms of data requirements and computational resources, yet it can be more computationally efficient than traditional numerical weather prediction (NWP) (Miyoshi et al., 2015). The integration of ML into meteorology has already led to notable advancements (Bochenek and Ustrnul, 2022), as evidenced by enhanced agro-meteorology (Kelley et al., 2020) and precipitation forecasting (Choi et al., 2018). Advanced ML methods applicable to weather forecasting include transformer models (Nguyen et al., 2023; Shevchenko et al., 2024; Pan et al., 2024), convolutional neural networks (Abdalla et al., 2021; Morozov et al., 2023), recurrent neural networks (Mozikov et al., 2023; Shevchenko et al., 2024; Zaytar and El Amrani, 2016; Grabar et al., 2023; Farmanifard et al., 2023), generative adversarial networks (Chen et al., 2019), and traditional ML approaches (Czibula et al., 2021; Pilyugina et al., 2025; Taniushkina et al., 2024; Tibau et al., 2021). Autoencoders are widely used for feature extraction and noise reduction, but are less commonly applied in climate analysis (Racah et al., 2017; Tibau et al., 2021).

An example of machine learning application for SCS-related hazard forecasting is given in Zhou et al. (2019), where the authors present a deep learning model for severe convective weather (including short-duration heavy rain, hail, convective gusts, and thunderstorms) based on NWP data with a 12-h lead time. In Czernecki et al. (2019), the authors combine real-time radar reflectivity, lightning frequency, and reanalysis datasets to predict significant hail events. Whereas previous work has offered relevant insights into various aspects of weather forecasting, our study contributes to long-term assessment of SCS-related hazard frequency. By leveraging ML techniques and the unique capabilities of CMIP climate projections, we aim to provide valuable insights into this underexplored area of research.

Our research contributions to SCS-risk assessment is two-fold:

- designing a neural network specifically tailored for long-term SCS-related hazard frequency risk,
- providing high-quality SCS-related hazard forecasts under climate change with limited available data.

The paper is organized as follows. Section 2 provides a summary of the research methodology, including details on the data used, the problem setup, the metrics applied, and the network architecture. Evaluation of the employed techniques is presented in Section 3. Following this, Section 4 assesses the results, and outlines directions for future research. The paper concludes with Section 5, discussing the limitations of the current study, summarizing our findings and offering final remarks.

2 Materials and methods

Our objective is to develop a machine learning model for assessing SCS-related hazard risk within a defined area. The approach leverages the physics encapsulated in CMIP models, which use mathematical representations of fundamental laws of physics and chemistry to simulate Earth climate system behavior. This helps improve accuracy and partially addresses challenges posed by limited data availability. Additionally, this synergy improves the robustness of the machine learning methodology, enabling more reliable predictions in the context of climate dynamics; however, CMIP physics alone, without machine learning adjustment, have limited power for hazards risk assessment due to intrinsic simplification of complex dependencies, spatial and geographical averaging. In this section, we provide a detailed description of our approach.

2.1 Data

The study focuses on the region extending from the Caucasus Mountains to the Ural Mountains, an area historically prone to severe convective storms.

All SCS-related hazards are associated with cumulus congestus or cumulonimbus clouds, which can reach heights of over 18 km in some cases (Krauss et al., 2007). To capture this structure, we utilize data from multiple isobaric surfaces (i.e., different altitudes), resulting in a three-dimensional representation. Our climate data source is the Coupled Model Intercomparison Project Phase 5 (CMIP5) output (Taylor et al., 2012). A key advantage of CMIP models is rigorous validation against historical observations. To the best of our knowledge, CMIP outputs provide some of the most reliable openly available climate projections. Numerous universities and research institutions worldwide develop CMIP models based on varying assumptions regarding climate change scenarios. These scenarios range from adopting environmentally friendly practices to the opposite extreme of continuous increases in atmospheric emissions. The assumptions are categorized into four representative concentration pathways (RCPs) corresponding to various forcings.

For our study, we select Representative Concentration Pathway 8.5 (RCP 8.5), which is considered a worst-case option in CMIP 5. Recent events worldwide, such as the EU energy U-turn (Liu et al., 2023) and increased coal consumption in China (Chen et al., 2022), are likely consistent with this emissions scenario.

It is important to note that a significant portion of SCS-related hazards are typically associated with local weather patterns (Prein and Holland, 2018). To avoid losing valuable information due to low spatial density, we employ the simulation provided by

the Meteorological Research Institute (MRI-CGCM3) in CMIP5 (Yukimoto et al., 2012b), which offers the best available spatial resolution of $1^\circ \times 1^\circ$. Although this resolution does not explicitly resolve convective-scale processes or simulate deep convection directly, it effectively represents synoptic-scale patterns (i.e., moisture transport, jet positioning, baroclinic zones, etc.) that modulate environments favorable for SCS development. The features selected for model training include specific humidity, temperature, and the two components of the wind vector, all on 12 isobaric surfaces. These parameters are available at 6-h temporal resolution.

A map illustrating the orography of this region, playing a critical role in the prediction quality, is provided for visualization purposes in Figure 1.

The target variable data were obtained from Rosgidromet (All-Russia Hydrometeorological Data Research Institute, 2023), which provides them in the form of telegrams using the KN-01 format, a local adaptation of the SYNOP code (Zabolotnikov, 2010). Note that these data are only available at locations with meteorological stations. In contrast, CMIP5 outputs are presented on a regular grid. To ensure compatibility with the target data, we extract only the grid cells that contain meteorological stations and their neighboring cells from the full simulation output during preprocessing.

Figure 2 facilitates further analysis of the target data by illustrating the distribution of atmospheric phenomena categories classified by the KN-01 code. A considerable number of SCS-related hazards fall into the "Showers" and "Thunderstorms" categories. However, these categories occur less frequently than the cumulative sum of other events and days without any event, indicating class imbalance. Methods to address this issue are detailed in Section 2.2.

Figure 2 also visualizes the number of SCS-related hazards (millions per year). A notable aspect of these data is a distinct, and likely artificial, rise (Prein and Holland, 2018) in the number of events around 1995, plausibly related to improvements in observation tools and techniques. This spike is an important consideration in the analysis of the dataset. Although, the time span chosen for model training deliberately excludes this anomalous period, helping ensure that training data remain representative of typical conditions and improving model reliability.

Table 1 presents the complete list of variables and their sources used in our research. The training dataset covers 2009–2018, the



FIGURE 1
Orographic features of the Caucasus Mountains from https://opentopomap.org.

TABLE 1 Description of the data sources.

Dataset	Variables	Time span
Rosgidromet meteorological observations (All-Russia Hydrometeorological Data Research Institute, 2023)	SCS-related hazards	2009–2022
CMIP5 MRI-CGCM3 RCP8.5 (Yukimoto et al., 2012a,b)	u-component of wind	2009–2100
	v-component of wind	
	specific humidity	
	temperature	

CMIP5 MRI-CGCM3 RCP8.5 is utilized for feature extraction, and Rosgidromet data are used for the target.

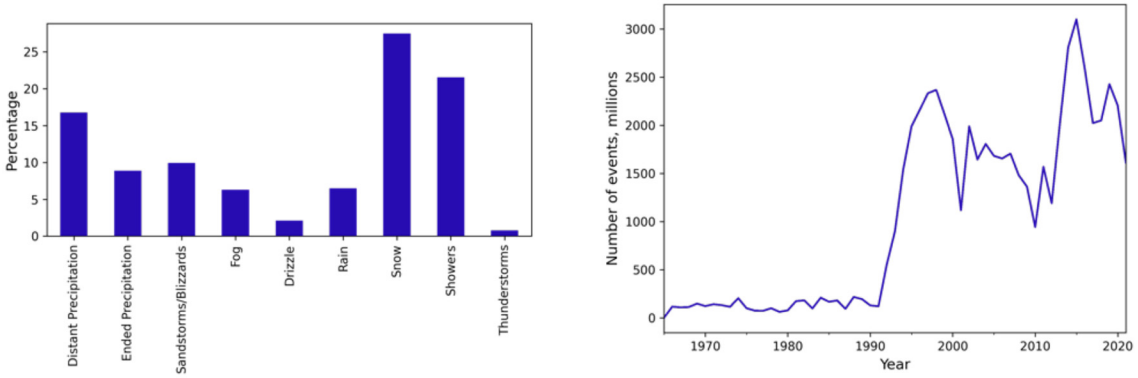


FIGURE 2
Left: distribution of main categories of atmospheric phenomena according to the KN-01 code. SCS-related hazard events predominantly fall into the "Showers" or "Thunderstorms" categories. The distribution reveals class imbalance, with a ratio of approximately 3:1 (no SCS : SCS). **Right**: the number of SCS-related hazards for the studied region (millions per year). An phase transition representing a substantial increase in the number of events happened around 1995.

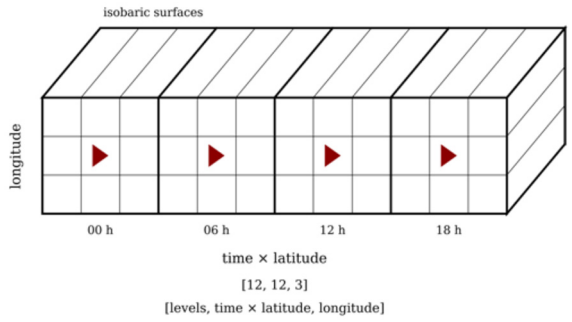


FIGURE 3
Input format. Gridded CMIP5 data are transformed into a form suitable for an autoencoder-based architecture. 3 × 3 cell blocks are cut and stacked along the time axis. Red triangles denote the location of the same station across different timestamps. Dimension *levels* represents the 12 isobaric surfaces used.

validation set covers 2019–2020, and the test set covers 2021–2022. Despite of climate change non-linearity, one-two decade long period can be considered as a stationary time series, since interannual variability substantially dominates non-linear climate change effect (Samset et al., 2023). We also believe that due to the high cost of capital and uncertainty in emissions scenario, a decade or a sub-decade planning for adaptation period is the target of this study.

In the autoencoder-based setting, for convenience of data transmission to the neural network, one observation for one feature is represented as [levels, (time × latitude), longitude], where:

- *levels* (shape of 12) represent selected isobaric surfaces, acting as channels in image processing;
- *time* × *latitude* (shape of 4 × 3) represents features corresponding to one day, concatenated along the time axis;
- *longitude* (shape of 3) refers to the longitude axis.

We provide a visual representation of the input format in Figure 3. The station (represented by a red triangle) is centered in the 3 × 3 block for the corresponding observation period, and the surrounding cells are adjacent in the regular grid. Before feeding data into the model, each feature is normalized by its mean value. For a CatBoost baseline solution, a single observation is represented as a one-dimensional vector of shape [layers × latitude × longitude × time].

2.2 Problem statement

We aim to design a machine learning model for predicting SCS-related hazard frequency within a specific region. In our approach, we rely on CMIP5 climate projections and consider them representative of the real environment in which we estimate our target variable.

In the current setting, we define frequency as the ratio of the days in a month when an SCS-related hazard was observed to the total number of days in that month. Since our objective is not to forecast SCS-related hazards for specific days several years in advance, but rather to estimate monthly hazard frequency based on the CMIP5 scenario, we adopt an approach that differs from classic forecasting models. It includes:

- solving a binary classification task for each day to determine whether an SCS-related hazard occurs,
- aggregating daily classification outputs by averaging results within each month.

By employing this approach, we estimate monthly averaged SCS-related hazard frequencies based on CMIP5 climate projections. This enables us to focus on long-term changes in hazard frequency rather than forecasting individual events, which is crucial for understanding long-term trends and developing effective adaptation strategies.

In the initial step of our algorithm, we perform binary classification for each grid point that contains a meteorological station. Class 1 represents the presence of an SCS-related hazard, while Class 0 indicates its absence.

To formalize the optimization problem, we minimize Binary Cross Entropy (BCE) between model outputs and true target values. At this step, it is crucial to account for class imbalance, given that the Class 1 to Class 0 ratio is approximately 1:3 in the data used. To address this issue, we introduce a constant weight λ applied to Class 1 before computing the mean BCE loss:

$$\mathrm{BCE} = -\frac{1}{N}\sum_{n=1}^{N}\big(\lambda y \log(p) + (1-y)\log(1-p)\big),$$

where $n \in \mathbb{N}$ is an index of the meteorological station; $y \in \{0; 1\}$ is the true label, where class 1 corresponds to an SCS-related hazard; p is the predicted probability of class 1; and λ is a positive class weight. By incorporating this weighting scheme, we mitigate bias toward the majority class. In our experiments, we used grid search to select λ and set $\lambda = 2$.

2.3 Quality metrics

We use common classification metrics during training, including precision, recall, and F1-score. But, after monthly aggregation, we can also use regression metrics on vectors of true and predicted monthly mean SCS frequencies. At this stage, we evaluate performance using the standard Root Mean Square Error (RMSE) and the metric introduced in Prein and Holland (2018). The cited paper develops a model for hail risk estimation and proposes a normalized version of RMSE to reduce the influence of undersampled observations and artificial trends.

The metric is designed to capture the annual cycle typical of convective phenomena, which is applicable to SCS-related hazards in general. We use root mean squared error (RMSE) normalized over annual cycle:

$$\mathrm{RMSE}_{AC} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\|\mathbf{O_n} - \mathbf{P_n}\|^2},$$

where $\mathbf{P}_n$ and $\mathbf{O}_n$ are normalized predicted and observed monthly SCS hazard probabilities for station n, over N meteorological stations.

We also use a standard RMSE measure:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} \|\mathbf{X_n} - \mathbf{X_n^{pred}}\|^2},$$

where $\mathbf{X}_n$ and $\mathbf{X}_n^{pred}$ are observed and predicted monthly SCS hazard probabilities, respectively.

2.4 Baseline approach

At the beginning of our experimental analysis, we established a straightforward baseline model for comparative purposes. In this baseline, we simplify the spatial representation of every variable by aggregating nine adjacent grid cells (see Figure 3) into a single cell at each timestamp. This simplification is justified by the localized nature of SCS-related hazards, where observing a 24-h window with four records (at 6-h intervals) is expected to be sufficient to detect individual events.

Notably, ignoring spatial structure prevents the model from incorporating information about associated synoptic processes. However, due to uncertainty about whether $1^\circ \times 1^\circ$ data adequately preserve small-scale spatial features pertinent to SCS formation, this simplification was considered acceptable. In the subsequent section, we compare the results of this approach with our neural network solution. For classification within this framework, we employed the CatBoost model.

2.5 Neural network approach

In this subsection, we describe a neural network that, to the best of our knowledge, is the first developed for SCS-related hazard frequency estimation based on climate projections. The novelty of the proposed approach is the combination of a specifically designed neural network architecture with a shift from forecasting to classification, enabled by the use of climate projections.

The main building blocks of the architecture are separate pre-trained autoencoders for each of four features and a linear classifier that uses the resulting embeddings as input. The overall model is depicted in Figure 4. In the basic configuration, the neural network comprises two main parts: a feature extractor and a classification head. The feature extractor consists of four autoencoders, each preceded by a single convolutional layer to extract spatial information. Given that only a 3×3 cell block around each station is considered, additional convolutional layers are unnecessary. As mentioned in Section 2.1, grid parts corresponding to a 24-h period are stacked along the time axis during preprocessing. Therefore, we use a stride of 3 for the convolutional layer to avoid mixing information across the time axis. The embeddings are then concatenated and passed to the classification head.

The classification head uses the binary cross entropy loss defined in Section 2.2. For the reconstruction output of each autoencoder, we use mean squared error loss:

$$\text{MSE} = \frac{1}{N}\sum_{n=1}^{N} \left\| f(\mathbf{w}, \mathbf{X}_n) - \mathbf{X}_n \right\|^2,$$

where $n \in \mathbb{N}$ is an index over samples; $\mathbf{X}_n \in \mathbb{R}^{\text{levels} \times (\text{time x lat}) \times \text{lon}}$ is a tensor of climate variables at one timestamp; and $f(\mathbf{w}, \mathbf{X}_n)$ is the autoencoder reconstruction.

The two network blocks are shown in Figure 4: the autoencoder-based feature extractor and the classification head. These blocks are designed to operate independently, allowing for a modular and computationally efficient workflow. This approach, referred to as the "separate" method, involves freezing the weights of the feature extraction block during the training of the classification head, which is optimized using BCE loss. By isolating the training of these components, we not only streamline the process but also significantly reduce the computational overhead, making it a practical choice for large-scale applications.

Although this methodology already demonstrates superior quality compared to the CatBoost baseline, we hypothesize that further performance gains can be achieved by fine-tuning the feature extractor in conjunction with the classification head. This leads to our next approach, termed "end-to-end", where both blocks are trained simultaneously. Here, the loss function combines the autoencoder's MSE reconstruction loss with the BCE classification loss, enabling the model to refine its feature representations while optimizing for classification accuracy.

To further enhance the model performance, we explore the stacked hourglass architecture (Newell et al., 2016). This method deepens the model without encountering the vanishing gradient problem (Hochreiter, 1998). Inspired by skip connections (He et al., 2016), this architecture stacks three autoencoders sequentially, with the final reconstruction derived from their combined outputs via a 3D convolution. This configuration, referred to as the "hourglass" method, enhances the model's capacity to capture hierarchical features while maintaining stable gradient flow. Within our framework, each autoencoder generates three distinct embeddings for every input feature. While the original formulation (Newell et al., 2016) aggregates these embeddings through summation, we investigate alternative aggregation strategies, pixel-wise 1D convolution and concatenation followed by a linear layer, to identify the most effective approach for feature fusion. The systematic combinations of classification and aggregation methods, along with their corresponding performance metrics, will be presented in the results table.

Training was conducted using the Adam optimizer with weight decay 0.0001 and β parameters set to 0.9 and 0.999. We implemented an initial eight-epoch warm-up period for the learning rate, followed by the CyclicLR scheduler configured with a maximum learning rate of 0.0015, a minimum learning rate of 0.0001, and a step size of 5 epochs.

3 Results

During training of the hourglass setting, we observed that the autoencoders for atmospheric temperature and specific humidity

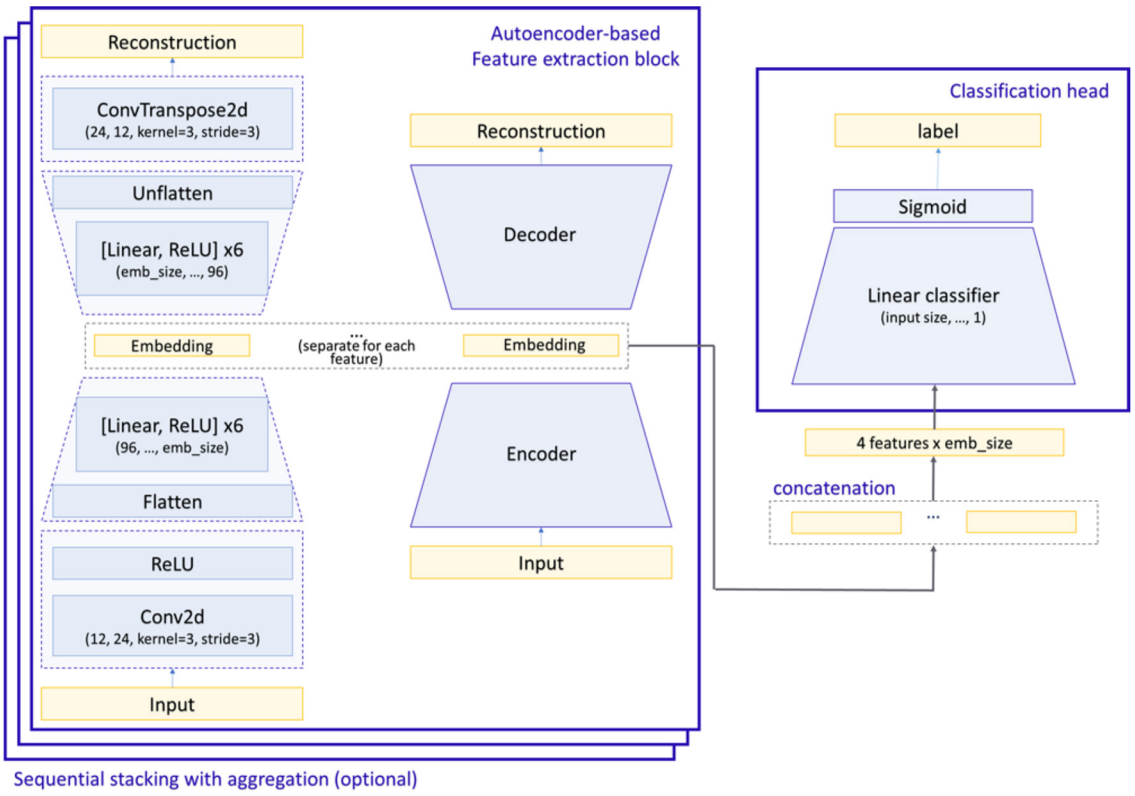


FIGURE 4
The model architecture. Each autoencoder is pre-trained separately for individual features, and the classification head is integrated during the fine-tuning phase.

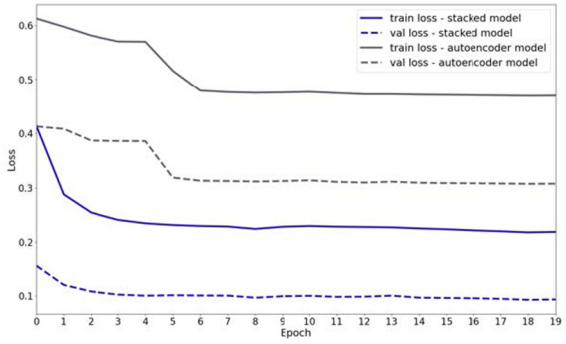


FIGURE 5
Comparison of reconstruction loss convergence across different aggregation approaches for the u-component of wind. The stacked (hourglass) architecture (in gray) outperforms the single-autoencoder architecture (in blue).

showed little improvement in reconstruction loss. This may be because these variables change slowly and exhibit low variance over a 24-h period, such that a single autoencoder is sufficient to capture the necessary information. In contrast, autoencoders for variables with more complex behavior, such as wind components, demonstrated a clear improvement in convergence. This supports our initial assumption; see Figure 5 for details.

Table 2 summarizes the metrics for all considered models, including the three aggregation approaches: summation, pixel-wise 1D convolution, and concatenation followed by a linear layer. The basic two-part neural network already outperforms the CatBoost baseline. Moving to the end-to-end setting yields a gradual improvement in most metrics. Notably, the end-to-end model performs best in terms of RMSE, while showing comparatively worse performance in capturing the normalized annual cycle of SCS-related hazard activity, as indicated by $RMSE_{AC}$.

If the ultimate goal is to minimize potential losses from SCS-related hazards, it is reasonable to prioritize recall for Class 1 (SCS) during the classification step, while maintaining strong performance across other metrics. Based on our analysis, the hourglass model with the `conv1d` aggregation approach is the most suitable for this problem setting.

4 Discussion

In this section, we present a visual example of the model forecast and analyze the results in the context of existing findings in the field. The aim is to evaluate model performance and gain insights into underlying patterns and mechanisms of SCS-related hazards.

While the CatBoost baseline offers a robust, interpretable, and computationally efficient solution by aggregating features

TABLE 2 Model evaluation.

Model methods		Class 0: no SCS			Class 1: SCS			RMSE	$RMSE_{AC}$
Aggregator	Classificator	Precision	Recall	F1	Precision	Recall	F1		
n/a	Catboost (baseline)	0.82	0.34	0.48	0.31	0.79	0.40	0.0100	0.0121
n/a	Separate	0.82	0.53	0.64	0.32	0.66	0.43	0.0086	0.0138
n/a	End-to-end	0.83	**0.54**	**0.65**	0.33	0.67	0.44	**0.0080**	0.0136
Sum	Hourglass	0.83	0.53	0.65	**0.34**	0.69	**0.45**	0.0084	0.0125
conv1d	Hourglass	0.84	0.50	0.63	0.33	0.71	**0.45**	0.0093	0.0154
concat (linear)	Hourglass	**0.85**	0.38	0.53	0.31	**0.81**	**0.45**	0.0125	**0.0114**
Support		68,015			23,360				

Columns 3 to 8 show metrics for the classification stage. Metrics for Class 1 are more important because this class corresponds to underrepresented severe convective phenomena. The last two columns represent metrics for the aggregation stage. Lower values of $RMSE_{AC}$ indicate a better ability to capture the annual cycle of SCS. Support denotes the number of elements in subsets corresponding to Class 0 and Class 1. Bold values indicate the best results.

to a one-day timescale, it inherently overlooks temporal dynamics and spatial dependencies critical for SCS detection. In contrast, our neural network approaches (particularly the end-to-end and hourglass architectures) explicitly model these dependencies, capturing spatial patterns and hierarchical feature interactions.

Topographic characteristics are crucial for convection development, as consistently emphasized in climatological research (Xie et al., 2006; Tang et al., 2012; Hagen et al., 2011). These studies frequently highlight the susceptibility of mountainous regions to convection given their interaction with atmospheric dynamics. In line with these findings, our forecast illustrates how terrain morphology, particularly in mountainous areas, can act as a catalyst for severe convection.

Figure 6 shows the forecast for spring and early summer (March–June) 2025. The figure indicates that high SCS-related hazard frequency is expected in the Black Sea coast–Caucasus region. Starting in April, the mountainous terrain of the Caucasus provides favorable conditions for severe convection. The mountains cause orographic lifting, leading to cloud formation and enhanced convection. In addition, rapid elevation changes create localized instability that contributes to intense thunderstorm development. Proximity to the Black Sea provides the atmospheric moisture needed for convection in this area. These findings align with the analysis of historical climate data in Inyukhin et al. (2014). Retrospective assessment of precision/recall in 2025, demonstrated expected quality (F1 for class 0 is 0.67, F1 for class 1 is 0.46, RMSE was 0.0082 for the proposed approach).

We also observe a narrow, extended area of activity in May and June in the center of the figure, corresponding to the southern part of the Ural Mountains. The rugged terrain of the Urals contributes to orographic lifting and the development of severe convection. During these months, there is a notable expansion of SCS-active areas and an increase in SCS-related hazard frequencies to the northeast of the Black Sea, indicating the peak season of SCS activity in the studied region. Figure 7 demonstrates risk evolution by 2035 under the most aggressive climate change scenario.

5 Conclusions

This work presents a neural network-based framework for translating climate-model projections into actionable, long-term estimates of severe convective storm (SCS)-related hazard frequency. Instead of attempting to predict individual events years in advance, we cast the problem as daily hazard classification followed by monthly aggregation. This design directly targets the decision-relevant quantity for climate-risk assessment: how often hazardous convection is expected to occur under a given emissions scenario.

Across a range of architectures, our approach consistently outperforms the CatBoost baseline, with the best-performing configuration (the hourglass stack with `conv1d` embedding aggregation) offering a strong practical trade-off: high recall for the underrepresented hazard class while maintaining competitive overall error after aggregation. Importantly, the resulting frequency maps reproduce physically plausible spatial patterns, including enhanced activity over complex terrain. In the example forecast for March–June 2025, the model highlights elevated SCS-related hazard frequency along the Black Sea coast and the Caucasus, as well as a distinct corridor over the southern Urals, consistent with established evidence that orography and moisture availability can amplify convection.

These capabilities are directly relevant to climate-resilient planning and risk management. Sectors such as agriculture, renewable energy, and insurance are increasingly exposed to severe wind and hail impacts (Kim et al., 2023; Ranasinghe et al., 2021). By coupling CMIP-driven projections with a lightweight ML pipeline, our method can help identify emerging hotspots, support stress-testing of assets and portfolios, and inform adaptation measures (e.g., infrastructure hardening, operational planning, and insurance pricing). The computational footprint is modest (8 h of training on two GTX 1080 Ti GPUs), making iterative retraining and scenario testing feasible.

Our results also clarify the main constraints of the problem. First, the model is currently tailored to the station network and

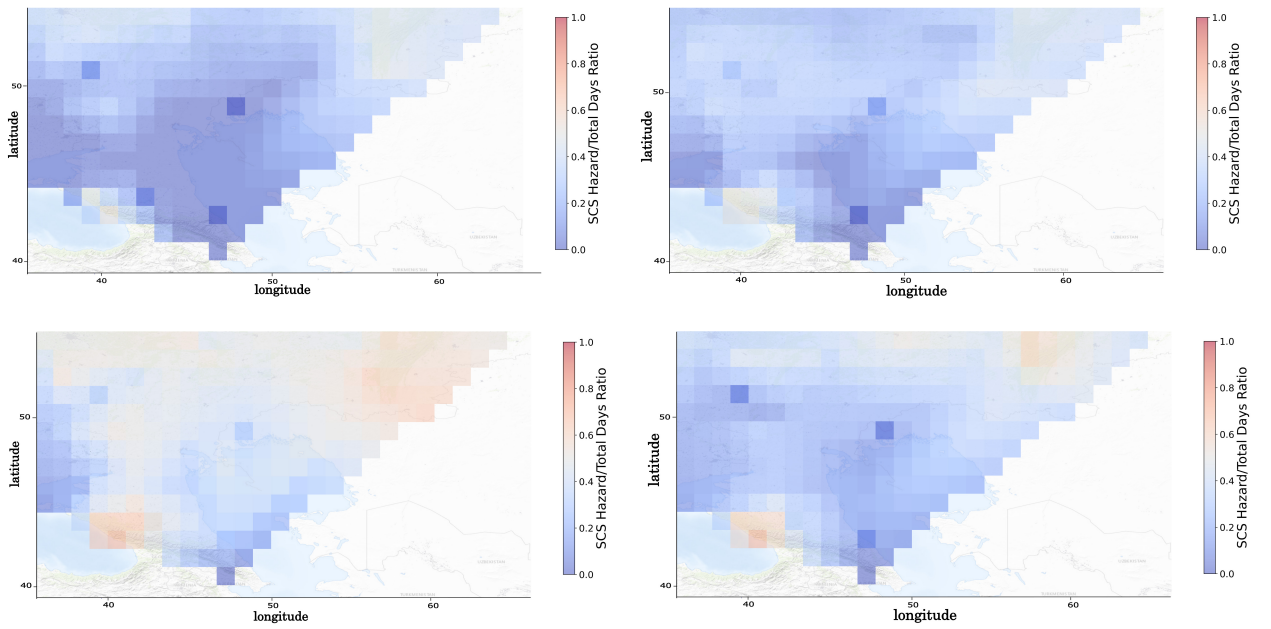


FIGURE 6
Monthly averaged forecasted SCS-related hazard frequencies for March, April, May, and June 2025. The forecast covers the region from the Caucasus to the Ural Mountains. Areas with high SCS activity are highlighted in red.

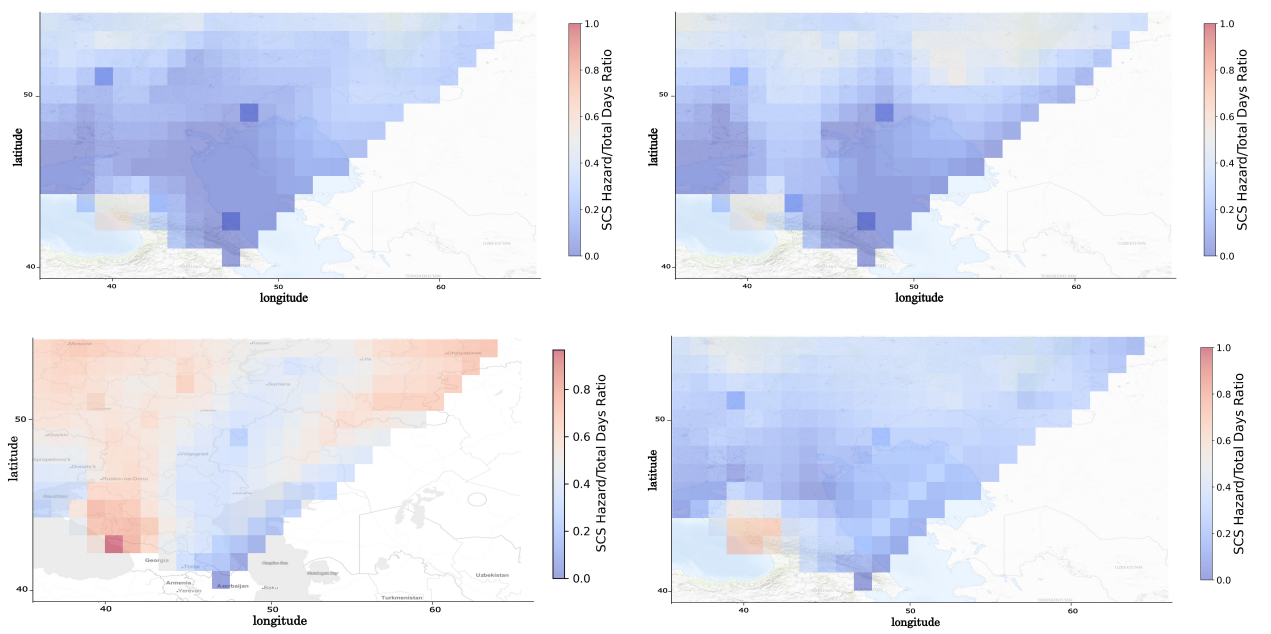


FIGURE 7
Monthly averaged forecasted SCS-related hazard frequencies for March, April, May, and June 2035 under the most aggressive emissions scenario RCP 8.5. The forecast covers the region from the Caucasus to the Ural Mountains. Areas with high SCS activity are highlighted in red.

region used for training; deployment in new regions requires retraining and careful validation. Second, the classifier tends to be conservative, favoring false positives over false negatives (high recall with lower precision). While this may be acceptable and often preferable in risk mitigation settings, it highlights the need for probability calibration and cost-sensitive evaluation aligned with end-user objectives. Third, despite exploring multiple architectural variants (including the hourglass stack, Newell et al., 2016, performance appears to plateau around an F1-score of about 0.45, suggesting that data limitations (label noise, reporting inhomogeneity, and class imbalance), likely being the primary bottleneck rather than model capacity.

Note, that despite of a significant change in extreme event frequency during the 2025–2035 decade, the mean values and even 90% quartile of distributional parameters (u, v, q, T) do not exhibit substantial changes at least until 2045–2050 and demonstrate consistent behavior in the area of study across various RCP scenarios. In particular, u and v components of the wind change consistently by 2045 within 5%, while v component exhibit almost 50% change under RCP 8.5 by 2100. Similarly, the temperature change in the area of interest increase on more than 2C by 2100, while by 2045 the expected change is less than 0.5 C. The latter result is consistent with the state-of-the-art results for some other regions (Hausfather and Peters, 2020; McCrystall et al., 2021).

Future work should focus on (i) improving label quality and consistency, (ii) incorporating multi-model and multi-scenario ensembles (e.g., broader CMIP coverage) to quantify uncertainty, and (iii) enhancing spatial generalization via domain adaptation or transfer learning. Together, these directions can strengthen confidence in long-term convective hazard frequency estimates and expand the applicability of ML-enabled climate-risk assessment.

Data availability statement

The datasets for this study can be found online in the following repositories: 1. World Climate Research Program data, in particular MRI MRI-CGCM3 rcp85 dataset used in the study, is accessible at: https://www.wdc-climate.de/ui/cmip5?acronym=MRMCr8. 2. All-Russia Hydrometeorological Data Research Institute severe convective storm dataset is accessible at: http://aisori-m.meteo.ru/waisori/. 3. NOAA National Centers for Environmental Information weather and climate disasters dataset is accessible at: https://www.ncei.noaa.gov/access/billions.

Author contributions

MM: Methodology, Software, Writing – original draft, Conceptualization. DT: Writing – original draft, Software, Data curation, Methodology. AB: Methodology, Validation, Writing – original draft, Formal analysis. YL: Writing – original draft, Validation, Formal analysis, Visualization. NS: Validation, Methodology, Writing – original draft. AO: Validation, Writing – original draft. RS: Visualization, Funding acquisition, Validation, Writing – original draft. IM: Writing – review & editing, Methodology. YM: Writing – review & editing.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abdalla, A. M., Ghaith, I. H., and Tamimi, A. A. (2021). "Deep learning weather forecasting techniques: literature survey," in *Proceedings of International Conference on Information Technology (ICIT)*, 622–626. doi: 10.1109/ICIT52682.2021.9491774

All-Russia Hydrometeorological Data Research Institute (2023). *Specialized arrays for climate research*. Available online at: http://aisori-m.meteo.ru/waisori/ (Accessed January 6, 2023).

Bochenek, B., and Ustrnul, Z. (2022). Machine learning in weather prediction and climate analyses–applications and perspectives. *Atmosphere* 13:180. doi: 10.3390/atmos13020180

Chen, H., Zhang, X., Liu, Y., and Zeng, Q. (2019). Generative adversarial networks capabilities for super-resolution reconstruction of weather radar echo images. *Atmosphere* 10:555. doi: 10.3390/atmos10090555

Chen, J., Li, Z., Song, M., Wang, Y., Wu, Y., and Li, K. (2022). Economic and intensity effects of coal consumption in China. *J. Environ. Manage.* 301:113912. doi: 10.1016/j.jenvman.2021.113912

Choi, C., Kim, J., Kim, J., Kim, D., Bae, Y., and Kim, H. S. (2018). Development of heavy rain damage prediction model using machine learning based on big data. *Adv. Meteorol.* 2018:5024930. doi: 10.1155/2018/5024930

Czernecki, B., Taszarek, M., Marosz, M., Półrolniczak, M., Kolendowicz, L., Wyszogrodzki, A., et al. (2019). Application of machine learning to large hail prediction–the importance of radar reflectivity, lightning occurrence and convective parameters derived from ERA5. *Atmosphere. Res.* 227, 249–262. doi: 10.1016/j.atmosres.2019.05.010

Czibula, G., Mihai, A., Albu, A.-I., Czibula, I.-G., Burcea, S., and Mezghani, A. (2021). Autonowp: an approach using deep autoencoders for precipitation nowcasting based on weather radar reflectivity prediction. *Mathematics* 9:1653. doi: 10.3390/math9141653

Elmore, K. L., Allen, J. T., and Gerard, A. E. (2022). Sub-severe and severe hail. *Weather Forec.* 37, 1357–1369. doi: 10.1175/WAF-D-21-0156.1

Farmanifard, S., Alesheikh, A. A., and Sharif, M. (2023). A context-aware hybrid deep learning model for the prediction of tropical cyclone trajectories. *Expert Syst. Appl.* 231:120701. doi: 10.1016/j.eswa.2023.120701

Frame, D., Rosier, S., Noy, I., Harrington, L., Carey-Smith, T., Sparrow, S., et al. (2020). Climate change attribution and the economic costs of extreme weather events: a study on damages from extreme rainfall and drought. *Clim. Change* 162, 781–797. doi: 10.1007/s10584-020-02729-y

Grabar, V., Marusov, A., Zaytsev, A., Maximov, Y., Sotiriadi, N., and Bulkin, A. (2023). Long-term drought prediction using deep neural networks based on geospatial weather data. *arXiv:2309.06212*.

Hagen, M., Van Baelen, J., and Richard, E. (2011). Influence of the wind profile on the initiation of convection in mountainous terrain. *Q. J. R. Meteorol. Soc.* 137, 224–235. doi: 10.1002/qj.784

Hausfather, Z., and Peters, G. P. (2020). Rcp8. 5 is a problematic scenario for near-term emissions. *Proc. Natl. Acad. Sci.* 117, 27791–27792. doi: 10.1073/pnas.2017124117

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. doi: 10.1109/CVPR.2016.90

Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *Int. J. Uncert. Fuzziness Knowl. Based Syst.* 6, 107–116. doi: 10.1142/S0218488598000094

Inyukhin, V. S., Kushchev, S. A., and Liev, K. B. (2014). "Distribution of hail origin sites in the central part of the North Caucasus," in *Proceedings of the All-Russian Open Conference on Physics of Clouds and Active Influences on Hydrometeorological Processes, dedicated to the 80th Anniversary of Elbrus*, 261.

IPCC (2022). *Climate Change 2022: Impacts, Adaptation and Vulnerability*. Summary for Policymakers. Cambridge, UK and New York, NY, USA: Cambridge University Press.

Kelley, J., McCauley, D., Alexander, G. A., Gray, W. F., Siegfried, R., Oldroyd, H. J., et al. (2020). Using machine learning to integrate on-farm sensors and agro-meteorology networks into site-specific decision support. *Trans. ASABE* 63, 1427–1439. doi: 10.13031/trans.13917

Kim, M., Lee, J., and Lee, S.-J. (2023). Hail: Mechanisms, monitoring, forecasting, damages, financial compensation systems, and prevention. *Atmosphere* 14:1642. doi: 10.3390/atmos14111642

Krauss, T. W., Sinkevich, A. A., Veremey, N. E., Dovgalyuk, Y. A., and Stepanenko, V. D. (2007). Study of the development of an extremely high cumulonimbus cloud (Andhra Pradesh, India, september 28, 2004). *Russ. Meteorol. Hydrol.* 32, 19–27. doi: 10.3103/S1068373907010037

Kron, W., Löw, P., and Kundzewicz, Z. W. (2019). Changes in risk of extreme weather events in Europe. *Environ. Sci. Policy* 100, 74–83. doi: 10.1016/j.envsci.2019.06.007

Lesk, C., Rowhani, P., and Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature* 529, 84–87. doi: 10.1038/nature16467

Liu, Y., Xie, X., and Wang, M. (2023). Energy structure and carbon emission: Analysis against the background of the current energy crisis in the EU. *Energy* 280:128129. doi: 10.1016/j.energy.2023.128129

McCrystall, M. R., Stroeve, J., Serreze, M., Forbes, B. C., and Screen, J. A. (2021). New climate models reveal faster and larger increases in arctic precipitation than previously projected. *Nat. Commun.* 12:6765. doi: 10.1038/s41467-021-27031-y

Miyoshi, T., Kondo, K., and Terasaki, K. (2015). Big ensemble data assimilation in numerical weather prediction. *Computer* 48, 15–21. doi: 10.1109/MC.2015.332

Morozov, V., Galliamov, A., Lukashevich, A., Kurdukova, A., and Maximov, Y. (2023). CMIP X-MOS: Improving climate models with extreme model output statistics. *arXiv:2311.03370*.

Mozikov, M., Lukyanenko, I., Makarov, I., Bulkin, A., and Maximov, Y. (2023). Long-term hail risk assessment with deep neural networks," in *Advances in Computational Intelligence. IWANN 2023. Lecture Notes in Computer Science* (Springer Nature Switzerland), 288–301. doi: 10.1007/978-3-031-43085-5_23

Newell, A., Yang, K., and Deng, J. (2016). "Stacked hourglass networks for human pose estimation," in *Computer Vision-ECCV 2016: 14th European Conference* (Springer), 483–499. doi: 10.1007/978-3-319-46484-8_29

Nguyen, T., Brandstetter, J., Kapoor, A., Gupta, J. K., and Grover, A. (2023). "Climax: a foundation model for weather and climate," in *Proceedings of the 40th International Conference on Machine Learning (ICML 2023)*, *volume 202 of Proceedings of Machine Learning Research* (PMLR), 25904–25938.

NOAA National Centers for Environmental Information (NCEI) (2024). *U.S. billion-dollar weather and climate disasters*. Available online at: https://www.ncei.noaa.gov/access/billions/ (Accessed January 6, 2023).

Pan, X., Jiang, T., Sun, W., Xie, J., Wu, P., Zhang, Z., et al. (2024). Effective attention model for global sea surface temperature prediction. *Expert Syst. Appl.* 254:124411. doi: 10.1016/j.eswa.2024.124411

Pilyugina, P., Chernikov, T., Smirnova, M., Zaytsev, A., Bulkin, A., Burnaev, E., et al. (2025). A physics-informed machine learning framework for permafrost stability assessment. *IEEE Access* 13, 96423–96433. doi: 10.1109/ACCESS.2025.3573072

Prein, A. F., and Holland, G. J. (2018). Global estimates of damaging hail hazard. *Weather Clim. Extr.* 22, 10–23. doi: 10.1016/j.wace.2018.10.004

Racah, E., Beckham, C., Maharaj, T., Ebrahimi Kahou, S., Prabhat, and Pal, C. (2017). "Extremeweather: a large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events," in *Advances in Neural Information Processing Systems 30 (NeurIPS 2017)* (Curran Associates, Inc.), 3402–3413.

Ranasinghe, R., Ruane, A. C., Vautard, R., Arnell, N., Coppola, E., Cruz, F. A., et al. (2021). "Climate change information for regional impact and for risk assessment," in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press), 1767–1926.

Raupach, T. H., Martius, O., Allen, J. T., Kunz, M., Lasher-Trapp, S., Mohr, S., et al. (2021). The effects of climate change on hailstorms. *Nat. Rev. Earth Environ.* 2, 213–226. doi: 10.1038/s43017-020-00133-9

Samset, B. H., Zhou, C., Fuglestvedt, J. S., Lund, M. T., Marotzke, J., and Zelinka, M. D. (2023). Steady global surface warming from 1973 to 2022 but increased warming rate after 1990. *Commun. Earth Environ.* 4:400. doi: 10.1038/s43247-023-01061-4

Shevchenko, V., Lukashevich, A., Taniushkina, D., Bulkin, A., Grinis, R., Kovalev, K., et al. (2024). Climate change impact on agricultural land suitability: an interpretable machine learning-based Eurasia case study. *IEEE Access* 12, 15748–15763. doi: 10.1109/ACCESS.2024.3358865

Tang, X.-D., Yang, M.-J., and Tan, Z.-M. (2012). A modeling study of orographic convection and mountain waves in the landfalling typhoon Nari (2001). *Q. J. R. Meteorol. Soc.* 138, 419–438. doi: 10.1002/qj.933

Taniushkina, D., Lukashevich, A., Shevchenko, V., Belalov, I. S., Sotiriadi, N., Narozhnaia, V., et al. (2024). Case study on climate change effects and food security in southeast asia. *Sci. Rep.* 14:16150. doi: 10.1038/s41598-024-65140-y

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* 93, 485–498. doi: 10.1175/BAMS-D-11-00094.1

Tibau, X.-A., Reimers, C., Requena-Mesa, C., and Runge, J. (2021). "Spatio-temporal autoencoders in weather and climate research," in *Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science, and Geosciences* (Wiley), 186–203. doi: 10.1002/9781119646181.ch13

Xie, S.-P., Xu, H., Saji, N. H., Wang, Y., and Liu, W. T. (2006). Role of narrow mountains in large-scale organization of Asian monsoon convection. *J. Clim.* 19, 3420–3429. doi: 10.1175/JCLI3777.1

Yukimoto, S., Adachi, Y., Hosaka, M., Sakami, T., Yoshimura, H., Hirabara, M., et al. (2012a). MRI-CGCM3 model output prepared for cmip5 historicalext, served by esgf. doi: 10.1594/WDCC/CMIP5.MRMChn

Yukimoto, S., Adachi, Y., Hosaka, M., Sakami, T., Yoshimura, H., Hirabara, M., et al. (2012b). A new global climate model of the meteorological research institute: MRI-CGCM3 model description and basic performance. *J. Meteorol. Soc. Japan* 90, 23–64. doi: 10.2151/jmsj.2012-A02

Zabolotnikov, G. V. (2010). *Tutorial on Working with Meteorological Information Codes KN-01*. Russian State Hydrometeorological University.

Zaytar, M. A., and El Amrani, C. (2016). Sequence to sequence weather forecasting with long short-term memory recurrent neural networks. *Int. J. Comput. Applic.* 143, 7–11. doi: 10.5120/ijca2016910497

Zhou, K., Zheng, Y., Li, B., Dong, W., and Zhang, X. (2019). Forecasting different types of convective weather: a deep learning approach. *J. Meteorol. Res.* 33, 797–809. doi: 10.1007/s13351-019-8162-6