Cureus Journal of Computer Science
https://doi.org/10.7759/s44389-026-00050-w

Original Article

SafeFlow: Glacial Lake Outburst Floods Early Warning System With Probability-Based Risk Scoring

Asmit Patel [1] , Asha Rani Mishra [1, ✉] , Vedanshi . [1] , Krishna Yadav [1] , Molisha Agarwal [1]

1. *Artificial Intelligence and Machine Learning, Galgotias College of Engineering and Technology, Greater Noida, IND*

Received: February 28, 2026 | Review began: March 09, 2026 | Review ended: April 04, 2026 | Published: April 06, 2026

© **Copyright** 2026
This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Glacial Lake Outburst Floods (GLOFs) occur when ice dams break and glacial lakes suddenly burst open. These floods cause massive destruction as huge volumes of water rush downstream. Climate change is accelerating the formation and growth of glacial lakes, making GLOFs more frequent and increasingly unpredictable. However, traditional risk assessment techniques, such as field inventory and fixed hazard maps, are not capable of representing the dynamic nature of glacial hazards over space and time. To address this issue, the authors propose a new end-to-end deep learning framework called SafeFlow for GLOF risk estimation. The proposed framework combines meteorological information from ERA5, ice velocity information from NASA ITS_LIVE, and land surface temperature information from MODIS with local lake ablation characteristics such as area, elevation, and distance to upstream glaciers. The proposed framework provides a probabilistic risk score through a dashboard, which helps to effectively detect high-risk regions compared to traditional techniques.

Categories: Machine Learning (ML), Deep Learning, AI applications

Keywords: glacial lake outburst floods (glofs), artificial intelligence (ai), deep learning, transformer model, lstm, remote sensing, hybrid cnn lstm, probabilistic risk scoring, early warning system (ews)

Introduction

Glacial Lake Outburst Floods (GLOFs) are one of the most destructive types of hydro-geological hazards in high mountainous regions, caused by the catastrophic breakdown of natural ice or moraine dams, which suddenly discharge enormous amounts of water downstream [1]. This results in the massive destruction of villages, transportation infrastructure (roads and bridges), hydroelectric power stations, and extensive agricultural land, which constantly threatens the existence of at-risk communities, especially in the Himalayas, Karakoram, and Hindu Kush ranges [2].

The occurrence rate and intensity of GLOFs have been noticed to be rising in direct proportion to the increase in global warming, which is a major symptom of human-induced climate change as per Figure *1*, which illustrates an upward trend in occurrence frequency of GLOF [3,4]. This warming causes an increase in glacier melting, resulting in the creation of new glacial lakes and the rapid growth of existing ones, thus increasing the risk of hazardous flooding [5]. Since the beginning of the 20th century, the rate of recorded occurrences of GLOF events has been accelerating steadily, especially since the 1990s, which indicates the increasing severity of this threat [6].

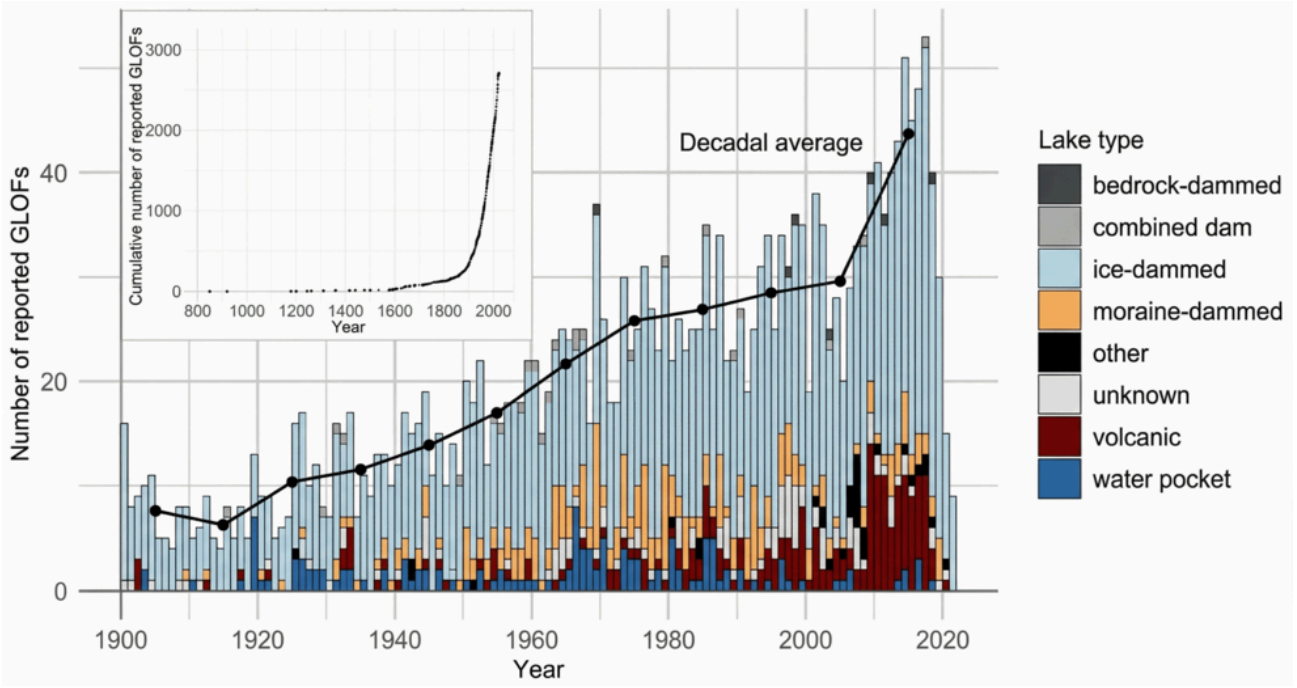


FIGURE 1: Number of Documented GLOF Events Globally from 1900 to 2020

GLOF, Glacial Lake Outburst Flood

The requirement for an effective early warning system is further highlighted by the fact that it is estimated that 15 million people around the world are living in areas that are highly vulnerable to GLOF hazard, with the largest number of people at risk residing in South Asian countries, specifically India, Pakistan, Nepal, and Bhutan. The objective of this study is to develop an early warning system for the hazard of GLOFs in the Chorabari Lake (Gandhi Sarovar) Glacier region of Uttarakhand, India, which is particularly vulnerable to the glacier's rapid retreat due to the rise in global temperatures.

Present methods for GLOF risk assessment usually involve static hazard maps, manual field observations, or simple threshold criteria [7]. These methods are inadequate because they do not account for the dynamic processes at work in glacial systems, such as rapid glacier retreat, changes in meltwater volumes, or moraine instability [8]. This makes current systems ineffective for real-time forecasting, thus preventing early warnings. Moreover, current systems are often based on simple binary warning systems (a "yes/no" warning), which often limit the effectiveness of disaster response [9]. In this case, if the system provides only binary warnings, disaster response managers are left with the option of either facing a potentially costly false alarm or a disastrous inaction. This makes it necessary to move towards a more adaptive and informative system [10,11].

Materials And Methods

Related work

Foundational Risk Assessments and Morphological Studies

Early researchers and scientists have long studied the phenomenon of GLOFs [12]. Initial research focused on defining GLOFs and understanding the processes involved in their occurrence, with the aim of identifying their causes. Consequently, inventories of GLOFs were compiled, and efforts were made to better understand the physical processes involved in their initiation [13].

Shrestha et al. [14] provided a comprehensive study of GLOF lakes in Nepal and Bhutan, identifying more than 200 potentially hazardous lakes. The authors used case studies and satellite imagery to identify these lakes, which helped increase awareness of the associated risks in the region. Although geographically specific, this study paved the way for the identification and evaluation of GLOF-related risks. Later, Fujita et al. [15] proposed a single index to evaluate lakes using remotely sensed digital surface models. This index is based on the slope angle from the lakeshore. They studied the topography and maximum possible outflow volume by analyzing the topography of the Himalayas. However, their approach was limited by temporal constraints, as it assessed glacier and lake risks at a specific point in time and did not account for continuous changes in glacier dynamics. Fujita and his team acknowledged the high temporal variability of lakes, glaciers, and water volumes. Worni et al. [16] conducted a thorough risk assessment of three significant glacial lakes in the Indian Himalayas. Their assessment was based on fieldwork, remote sensing, and dynamic modelling. Ahmed et al. [17] extended previous research by simulating flood events and changes in lake volume using computational models. However, a key limitation of these above-mentioned studies is their inability to provide real-time predictive models for glacial lake systems.

Remote Sensing and Computer Vision in Lake Extraction

With the increasing availability of satellite images, research interest turned to the automated detection and segmentation of glacial lakes [18]. Pandey et al. [19] focuses on a potentially hazardous moraine-dammed proglacial lake that has been growing at an alarming rate over the past 20 years in the Bhilangna Valley in central Himalaya, India. Sharma et al. [20] employed the U-Net model, which greatly improved the accuracy of lake detection compared to Convolutional Neural Networks (CNNs). More recently, Maslov et al. [21] introduced the GlaViTU (Hybrid CNN-Transformer) model. Panwar et al. [22] show how the deep learning model like CNN can be used for real-time activity tracking using large amounts of labeled training data.

Integrated Frameworks and Machine Learning Advancements

The past decade has witnessed a paradigm shift from basic detection to advanced predictive modelling. Ahmed [23] used a methodical approach to examine the research on GLOF risk management and hazard mitigation. With a focus on integration with climate change adaptation, the study emphasises the significance of early warning systems, GLOF risk management abilities, and information sharing. Zhou et al. [24] also developed a hybrid machine learning model that integrates glacier evolution data with environmental factors to enhance debris flow prediction. However, the major limitation of these studies is that they consider static terrain variables, which sometimes overlook the dynamic characteristics of glaciers that may alter the hazard pattern in a particular year. Mohanty et al. [25] overcame this limitation by employing multi-modal data fusion approaches that combined Optical, Synthetic Aperture Radar, and Digital Elevation Model data. This enabled the improvement of glacier velocity and faces mapping accuracy, although the integration of multiple data sources remains a technically challenging task.

Early Warning Systems and Real-Time Monitoring

The main application target of GLOF research is the development and execution of functional early warning systems. Kumar et al. [26] conducted simulations to understand the effects of flash floods, specifically the South Lhonak GLOF, on infrastructure and human life. Wang et al. [27] enhanced transboundary monitoring through satellite communication systems, enabling real-time cross-border alerts, although these systems still require human verification to minimize false alarms. Rada et al. [28] developed a forecasting system that uses weather predictions and observational station data to identify periods of elevated GLOF risk.

Hydrodynamic Modeling and Socio-Economic Impact

Comprehending the implications of GLOFs is as crucial as forecasting their occurrence. In 2024, Wangchuk and Tsubaki [29] combined BREACH with HEC-RAS to assess the risk associated with the Thorthomi Glacier, predicting extensive flood areas and potential damage to buildings. In 2025, Acharya et al. [30] evaluated the potential socioeconomic impacts of GLOFs originating from the potentially hazardous Dudh Pokhari glacial lake (0.34 km^2) in the Everest region of

Nepal. However, there is a significant weakness in the hydrodynamic models, such as that of Chapagain et al. [31], which is recognized as the "clear-water assumption". GLOFs are normally large debris flows with a high sediment concentration, and therefore these models may not be able to represent the actual potential of the water-sediment mixture.

Conclusion and Future Directions

The literature indicates a clear path from manual observation to AI-based multi-sensor monitoring. There is a need for dynamic integration that prioritizes glacier movement over static terrain characteristics. Additionally, there is a pressing need for global coordination to address disparities in monitoring capacity across regions.

Proposed work

SafeFlow estimates GLOF risk using a robust multi-modal forecasting framework that integrates projections from three distinct deep learning models: Terracast, Tempcast, and Raincast. A detailed description of the datasets used by these models is provided in Table *1*. The methodology was evaluated in the Chorabari Glacier (Gandhi Sarovar) region in Uttarakhand.

Model	Dataset(s)	Source	Key Input Variables
Raincast	ERA5-Land Daily Aggregates Dataset	ECMWF ERA5 Climate Reanalysis	Total precipitation, 2 m temperature, surface pressure, 10 m U/V wind components
Terracast	ITS_LIVE Glacier Velocity Dataset	Remote Sensing Products (Sentinel-1 SAR)	Latitude, longitude, date, velocity components
Tempcast	MODIS MOD11A1 LST Product (V6.1)	NASA MODIS (Terra/Aqua Satellites)	LST values, QC Flags, Latitude/Longitude, Seasonal features (sin/cos encodings)

TABLE 1: Data Sources and Input Variables for SafeFlow Multi-Modal Architecture

ECMWF, European Centre for Medium-Range Weather Forecasts; LST, Land Surface Temperature; SAR, Synthetic Aperture Radar

Dataset Description

Raincast model uses ERA5-Land Daily Aggregated Climate Dataset [32]. The European Centre for Medium-Range Weather Forecasts created this high-resolution global reanalysis climate dataset. It uses numerical weather models in conjunction with observational data to generate hourly and daily estimates of land surface and atmospheric variables. By offering greater spatial resolution land surface parameters, ERA5-Land enhances the original ERA5 dataset and is appropriate for hydrological and environmental modelling. Variables in this dataset are total precipitation, snowfall, surface runoff, soil moisture, air temperature, and evaporation. These variables directly affect glacial lake water accumulation, which increases the probability of a GLOF event.

Terracast model uses the Inter-Mission Time Series of Land Ice movement and Elevation (ITS_LIVE) dataset [33].Using satellite imagery, collection offers a long-term worldwide record of glacier surface movement. The dataset was created using optical and radar imagery from the Landsat (4-9), Sentinel-1, and Sentinel-2 satellites. The autoRIFT feature tracking technique, which monitors displacement between image pairs to predict glacier motion, is used to construct surface velocity fields. Monitoring glacier dynamics, ice flow acceleration, and glacier mass movement - all important markers of glacial instability and possible risk of GLOF - is made possible by this dataset. The dataset is one of the most extensive

glacier velocity datasets accessible, with over 36 million image pairings spanning glaciers worldwide from 1982 forward. Variables used in the dataset are glacier surface velocity (m/year), ice flow direction, velocity uncertainty, and time-series displacement vectors. These variables help detect abnormal glacier motion, which can trigger glacial lake dam failure.

The Moderate Resolution Imaging Spectroradiometer (MODIS) sensor on NASA's Terra satellite is the source of the MOD11A1 Land Surface Temperature (LST) dataset, which is used by Tempcast Model [34]. It offers daily emissivity and LST readings. The split-window technique, which estimates surface temperature under clear skies by combining many thermal infrared bands, is used to retrieve temperature. The dataset is extensively utilised in hydrological models, glacier melt research, and climate monitoring. Globally, the product offers daily per-pixel temperature readings with a spatial resolution of 1 km. Variables used in this dataset are LST (day), LST (night), surface emissivity, quality assurance flags, and observation time. Surface temperature strongly influences glacier melt dynamics and lake expansion, which are key precursors of GLOF events.

Data Preparation and Structuring

The reliability of open data platforms provides access to high-resolution satellite data and climate reanalysis datasets for ongoing research, which is the primary goal of this study, as they are inexpensive and globally accessible. In addition, the long-term datasets, including MODIS data from 2000 to the present, allow models to learn from and make predictions about long-term climate trends rather than just short-term anomalies. One of the most technically challenging issues with the Tempcast LST dataset was the high proportion of low-quality data, up to 60%, caused by cloud interference affecting the calibration of the satellite sensor. To mitigate this issue, multiple imputation and quality-filtering methods were applied to fill in missing data and ensure continuity, as required for the LSTM model. The original time-series data were transformed into continuous cyclic features representing the day of the year using sinusoidal and cosinusoidal functions, enabling the model to identify seasonal trends, such as summer glacier surges. All numerical variables were scaled using Min-Max techniques and aggregated into fixed-length time series, specifically 30-day intervals for each seasonal timeframe.

Deep Learning Models Architectures

The research relies solely on high-resolution open-source satellite imagery and climate reanalysis data. This is a crucial starting point for the development of a cost-effective solution that has global applicability. The fact that there is historical data available for decades, as in the case of MODIS data from 2000, allows the models to learn and predict long-term climate patterns rather than just short-term trends. Detail of the region of interest, Chorabari Lake, Kedarnath, Uttarakhand, in terms of latitude and longitude is shown in Figure *2*.

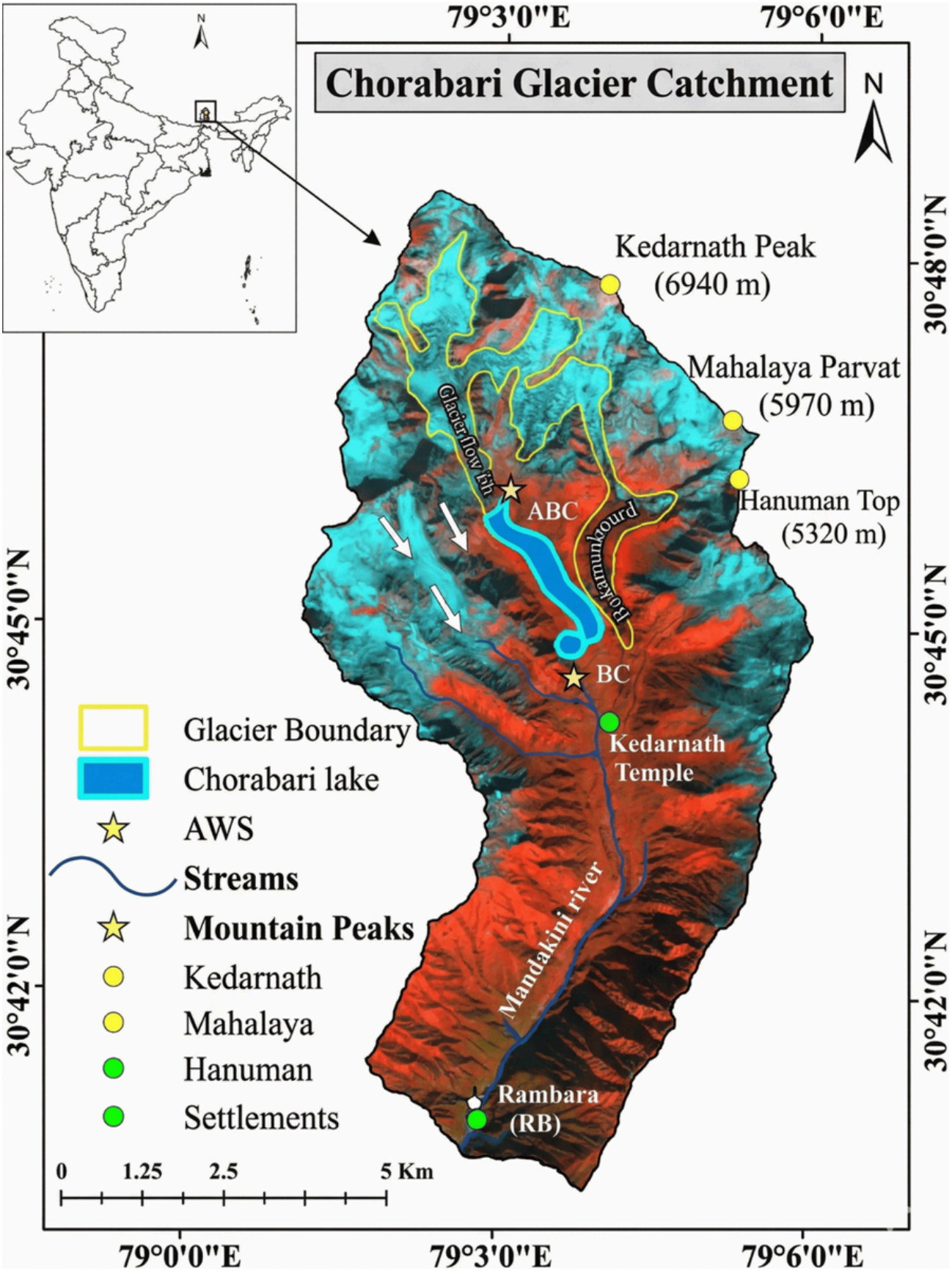

Chorabari Glacier Catchment
N
79°3'0"E
79°6'0"E
30°48'0"N
30°45'0"N
30°42'0"N
79°0'0"E
Kedarnath Peak
(6940 m)
Mahalaya Parvat
(5970 m)
Hanuman Top
(5320 m)
ABC
BC
Kedarnath
Temple
Rambara
(RB)
Mandakini river
Glacier flow ith
Boukamunkaourd
Glacier Boundary
Chorabari lake
AWS
Streams
Mountain Peaks
Kedarnath
Mahalaya
Hanuman
Settlements
0
1.25
2.5
5 Km

FIGURE 2: Region of Interest (30.74° North Latitude and 79.05° West Longitude)

1. Raincast: Hybrid CNN-LSTM Precipitation Level Forecasting: Raincast monitors changes in hydrology and forecasts heavy rains that could rapidly change the water will, and hence the dam condition, and these changes account for 20% of the overall risk score.

Layer (type)	Output Shape	Parameter
conv1d (Conv1D)	(None, 28, 64)	2,368
max_pooling1d	(None, 14, 64)	0
lstm (LSTM)	(None, 50)	23,000
dropout (Dropout)	(None, 50)	0
dense (Dense)	(None, 1)	51

TABLE 2: Raincast Architecture Module

Conv1D, 1-Dimensional Convolutional Layer; LSTM, Long Short-Term Memory

The architecture design, as shown in Table *2*, has been implemented using a hybrid CNN-LSTM model to predict sequentially. The initial layer of the input sequence is a Conv1D layering of input data followed by a MaxPooling1D. These layers allow for local recognition of characteristics through down sampling the number of dimensions within the multivariate data (surface pressure, temperature, and wind data) being used as the input to the CNN. With the local features captured through Cam, the model is able to recognise the pattern made by the input data to show consistent changes in the input data as evidenced by a reduction in surface pressure with a rise in wind characteristics as a leading indicator of significant rainfalls in the near future. After obtaining the local features from the CNN, they are passed to the LSTM layer (i.e. LSTM(128)) to assess the temporal dependencies and estimate the overall precipitation amounts expected. This hybrid model is appropriate because of the local meteorological model characteristics observed in mountainous regions. The model is trained to recognise rapid increases in rain; rainfall amounts of greater than 50 mm can present a very rapid risk for flash flooding.

2. Terracast: Transformer-Based Glacier Velocity Modeling: Terracast is the kinetic monitoring part of the system, which gives predictions on the velocity of the glacier terminus. This velocity directly affects the mechanical stress on the dam structure and is the most important physical precursor to failure, which is given the highest weightage (50%) in the final risk assessment.

The modelling technique uses a Transformer architecture, as shown in Table *3*, which models glacier surge dynamics as a complex, non-local time-series problem that requires efficient management of long-range dependencies. The architecture consists of six Transformer encoder layers, each of which includes an eight-head multi-head attention mechanism. This attention mechanism allows simultaneous weighting of input features over the entire 32-day window and captures complex spatial dependencies, thus representing the dynamics of the glacier's accumulation zone and its effect on the terminus.

The model predicts a single velocity measure in meters per year, which is trained using Mean Squared Error loss and measured using Mean Absolute Error. The most important output is the identification of surge events, which are velocities above 500 m/yr, that are a high-confidence indicator of imminent mechanical failure risk.

Layer (type)	Output Shape	Parameters
embedding (Linear)	(None, seq_len, d_model)	input_dim × d_model + d_model
positional_encoding (Parameter)	(1, 1000, d_model)	1,000 × d_model
transformer_encoder (TransformerEncoder)	(None, seq_len, d_model)	num_layers × TransformerEncoderLayer params
multihead_attention	(None, seq_len, d_model)	$4 \times d_model^2 + 4 \times d_model$
feedforward (Linear ×2)	(None, seq_len, d_model)	2 × (d_model × dim_ff + dim_ff)
layer_norm ×2	(None, seq_len, d_model)	4 × d_model
fc (Linear)	(None, 1)	d_model + 1

TABLE 3: Terracast Architecture Module

3. Tempcast: LSTM-Based LST Forecasting: Thermal stress is considered one of the leading causes of meltwater generation and compromise of internal dam integrity; Tempcast assigns a 30% risk weight for this predictor.

Layer	Output Shape	Parameter
bidirectional_4 (Bidirectional)	(None, 10, 100)	22,400
dropout_5 (Dropout)	(None, 10, 100)	0
bidirectional_5 (Bidirectional)	(None, 100)	60,400
dropout_6 (Dropout)	(None, 100)	0
dense_8 (Dense)	(None, 1)	101

TABLE 4: Tempcast Architecture Module

The model employs a three-layer LSTM architecture [LSTM (128)-LSTM (64)-LSTM (32)] with dropout rates of 0.4 for the first two layers and 0.2 for the final layer, to capture complex, multi-layer temporal patterns in the highly volatile LST data, as shown in Table *4*. The proposed model will have the ability to identify long-term warming trends, which average an increase of 0.8 K per decade, as well as short-term seasonal peaks, which are critical for predicting melt events. The proposed approach is expected to achieve relatively low prediction errors (less than ±3 K) for seasonally based melting peak predictions. Temperature predictions from Tempcast are reported in Kelvin, with warming peaks above 279.15 K indicating high melt rates. The Tempcast results validate the efficacy of the LSTM as a predictive forecasting tool for LST values, particularly when optical sensor data are unavailable, such as under conditions of high cloud cover.

How to cite this article:
Patel A, Mishra A, . V, et al. (April 06, 2026) SafeFlow: Glacial Lake Outburst Floods Early Warning System With Probability-Based Risk Scoring. Cureus J Comput Sci 3 : es44389-026-00050-w. DOI https://doi.org/10.7759/s44389-026-00050-w

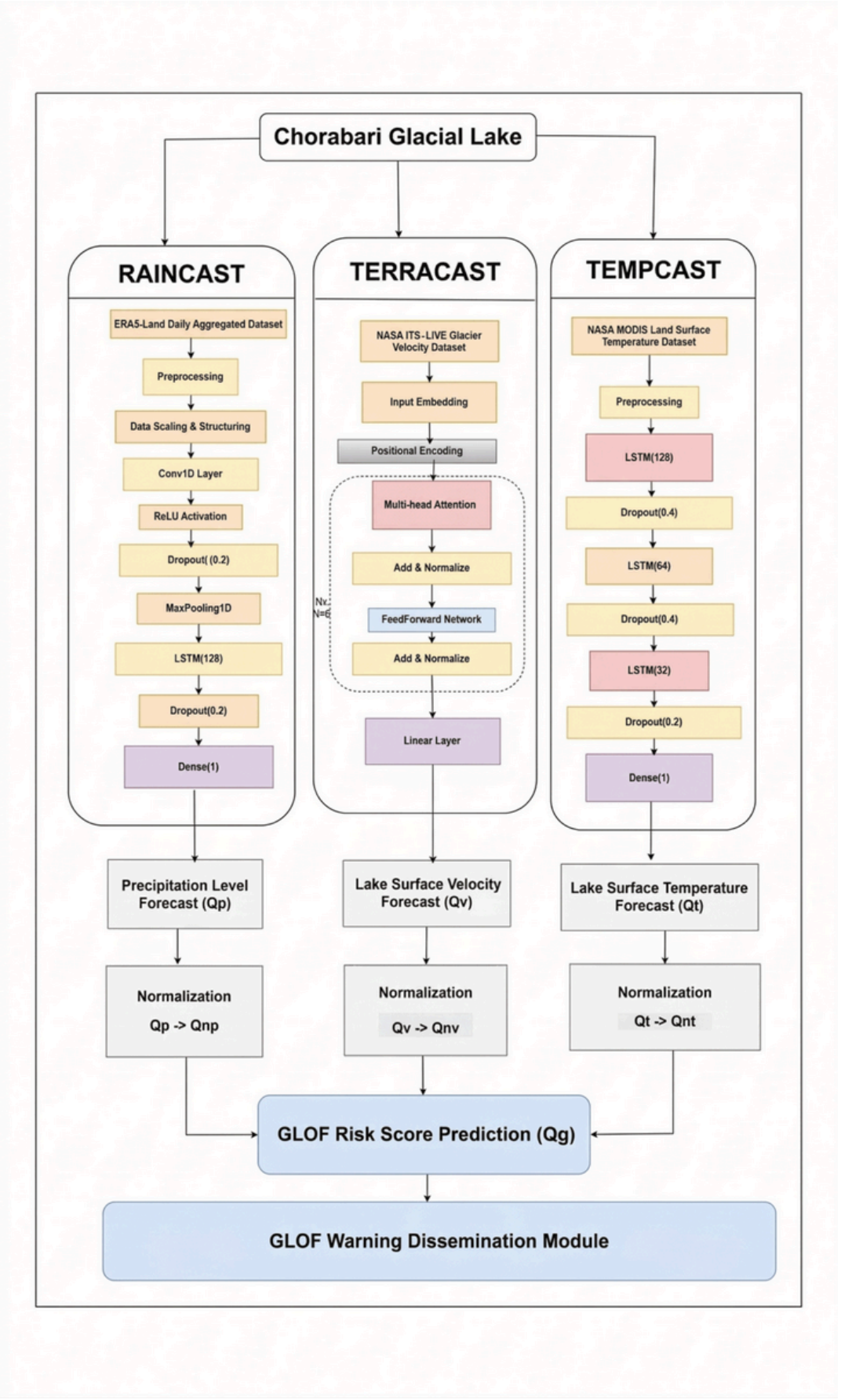

Chorabari Glacial Lake
RAINCAST
ERA5-Land Daily Aggregated Dataset
Preprocessing
Data Scaling & Structuring
Conv1D Layer
ReLU Activation
Dropout((0.2)
MaxPooling1D
LSTM(128)
Dropout(0.2)
Dense(1)
TERRACAST
NASA ITS-LIVE Glacier Velocity Dataset
Input Embedding
Positional Encoding
Multi-head Attention
Add & Normalize
FeedForward Network
Add & Normalize
Linear Layer
TEMPCAST
NASA MODIS Land Surface Temperature Dataset
Preprocessing
LSTM(128)
Dropout(0.4)
LSTM(64)
Dropout(0.4)
LSTM(32)
Dropout(0.2)
Dense(1)
Precipitation Level Forecast (Qp)
Lake Surface Velocity Forecast (Qv)
Lake Surface Temperature Forecast (Qt)
Normalization
Qp -> Qnp
Normalization
Qv -> Qnv
Normalization
Qt -> Qnt
GLOF Risk Score Prediction (Qg)
GLOF Warning Dissemination Module

FIGURE 3: Internal Layer Architecture Diagram

GLOF, Glacial Lake Outburst Flood; LSTM, Long Short-Term Memory; ReLU, Rectified Linear Unit

Dynamic Risk Scoring and Alert Generation

The risk score metric in the proposed Safeflow methodology is not derived from a single model but is instead calculated by combining the results of multiple deep learning models. Individual predictions from Tempcast (temperature), Terracast (ice velocity), and Raincast (precipitation) are integrated to produce a cumulative GLOF risk score. This score is calculated as a weighted and normalized combination of the three predictions, with velocity receiving the highest weight at 50%, followed by temperature at 30%, and precipitation at 20%. The internal workings and the complete proposed methodology are illustrated in Figures *3* and 4, respectively.

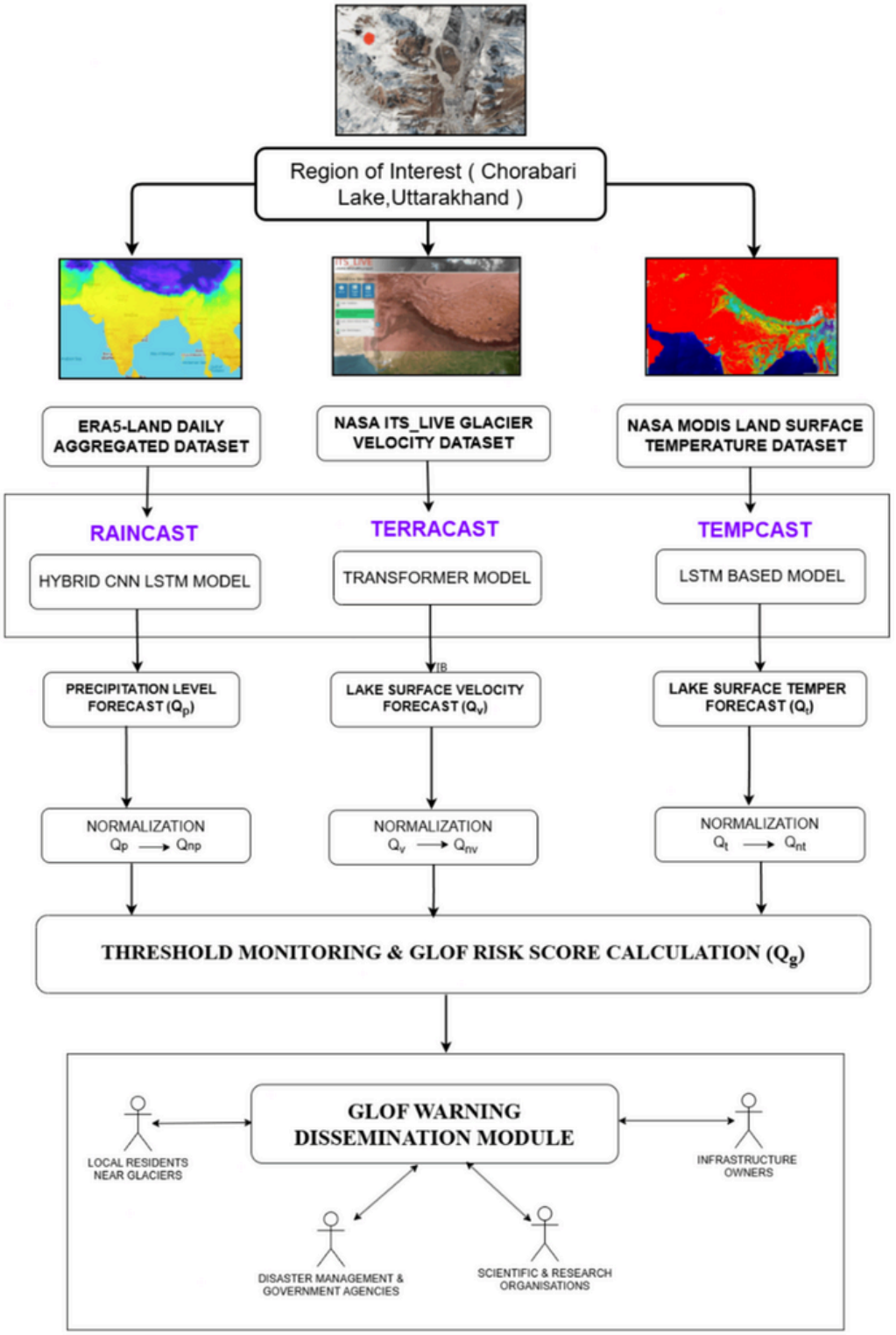

Region of Interest (Chorabari Lake,Uttarakhand)
ERA5-LAND DAILY AGGREGATED DATASET
NASA ITS_LIVE GLACIER VELOCITY DATASET
NASA MODIS LAND SURFACE TEMPERATURE DATASET
RAINCAST
TERRACAST
TEMPCAST
HYBRID CNN LSTM MODEL
TRANSFORMER MODEL
LSTM BASED MODEL
PRECIPITATION LEVEL FORECAST (Q_p)
LAKE SURFACE VELOCITY FORECAST (Q_v)
LAKE SURFACE TEMPER FORECAST (Q_t)
NORMALIZATION $Q_p \rightarrow Q_{np}$
NORMALIZATION $Q_v \rightarrow Q_{nv}$
NORMALIZATION $Q_t \rightarrow Q_{nt}$
THRESHOLD MONITORING & GLOF RISK SCORE CALCULATION (Q_g)
GLOF WARNING DISSEMINATION MODULE
LOCAL RESIDENTS NEAR GLACIERS
INFRASTRUCTURE OWNERS
DISASTER MANAGEMENT & GOVERNMENT AGENCIES
SCIENTIFIC & RESEARCH ORGANISATIONS

FIGURE 4: Safeflow Methodology

If any predicted value exceeds the defined danger thresholds, velocity above 200 m/year, temperature above 6°C, or precipitation above 100 mm, an Extreme Value Alert is triggered immediately. This alert bypasses the standard weighted calculation and sends a direct notification indicating that the GLOF risk has reached an extreme level. Notifications are broadcast to the dashboard and sent via SMS and/or API, simultaneously triggering the early warning siren system to initiate an immediate response.

Results And Discussion

This section describes the operational implementation of the SafeFlow system, its performance criteria in a simulated setting, and a comparison of the system with traditional monitoring methods.

System implementation and dashboard interface

The successful implementation and testing of the models culminated in the creation of a web application built using Streamlit, a Python framework for developing interactive data apps. This application makes our models accessible to users, providing real-time forecasts and visualizations.

Safeflow web application overview

The Streamlit web application, named Safeflow, as shown in Figure *5*, is designed for advanced glacier monitoring and GLOF risk assessment. It integrates Terracast, Tempcast, and Raincast, running them in parallel to deliver comprehensive environmental insights. The application features a clean and user-friendly interface with a header, which displays the "Safeflow" logo, a snowflake icon, and the tagline "Advanced Glacier Monitoring & GLOF Risk Assessment. Navigation Tabs include "Dashboard," "Forecast & Trends," "Advanced Analysis," and "About GLOFs" for diverse functionalities. Side Location Panel includes a left sidebar that shows the current monitoring area (e.g., "Chorabari Glacial Lake" with coordinates Latitude: 30.57, Longitude: 79.57). Current Status Section prediction shows values of Ice Velocity, Predicted Surface Temperature, and Predicted Rainfall, with alert messages.

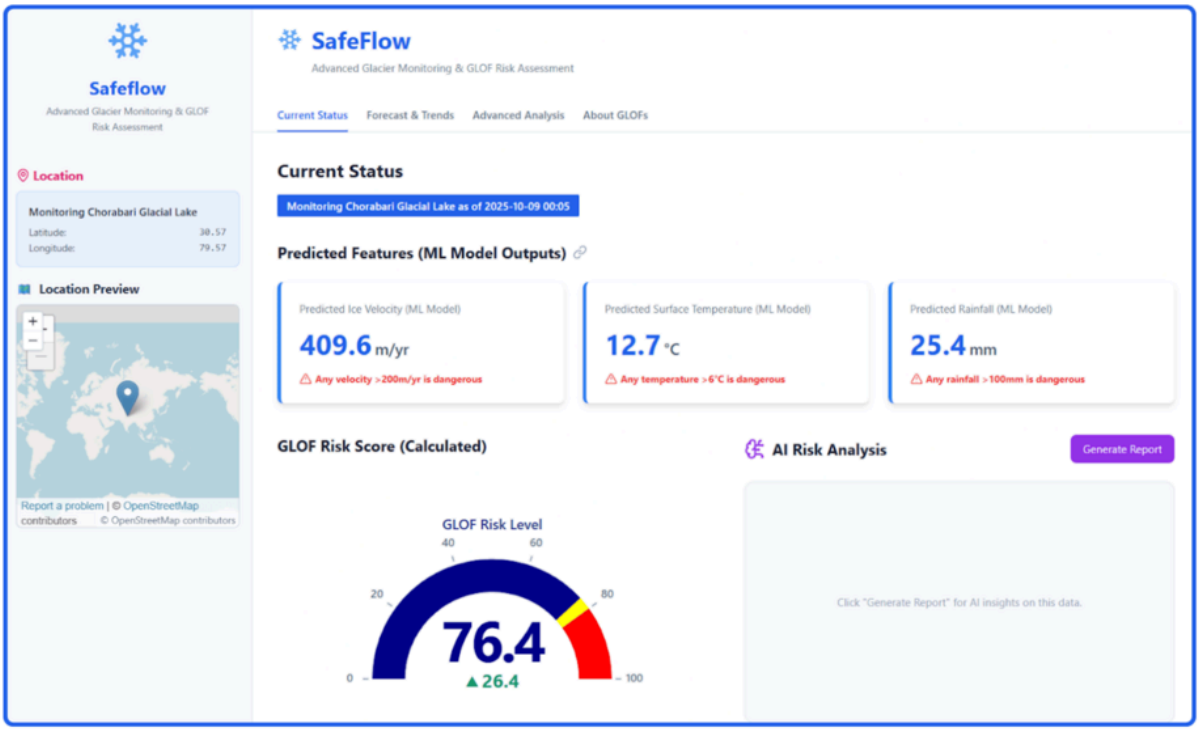


FIGURE 5: Safeflow Dashboard Page

Safeflow runs three models concurrently to ensure real-time, multi-faceted insights: Terracast module forecasts ice velocity and plots trends. Tempcast module predicts surface temperature and visualizes cycles, and Raincast module forecasts Rainfall or Precipitation Level and plots trends. Currently, Safeflow is deployed on Streamcloud using GitHub Pages, a cost-effective and accessible lightweight hosting solution. As the models grow in size and demand more computing power, a future migration to Amazon Web Services (AWS) is planned. AWS will provide scalable resources, including high-performance GPUs and elastic storage, to support larger models and increased user traffic.

Performance metrics and model evaluation

An indication of the fidelity of the system is the performance on each individual element, with a resultant overall accuracy of the simulated system that can be classified as high. Raincast was optimized to minimize mean square error (MSE) loss while estimating heavy volumes of rain (precipitation) that are associated with rapid increases in the hydrological load placed upon the system. In addition to this, it is important to highlight the use of the "weighted ensemble" approach that utilizes data fusion and reduces the overall variance in estimating classifying the class of high risk or potential for risk scenarios. Key performance metrics, loss vs epoch graph, and prediction values of Raincast model are shown in Table *5*, Figure *6*, and Figure *7*, respectively.

Metric	Value
MAE (mm)	3.90
RMSE (mm)	5.17
R^2 Score	0.525

TABLE 5: Raincast – Key Performance Metrics

MAE, Mean Absolute Error; RMSE, Root Mean Squared Error

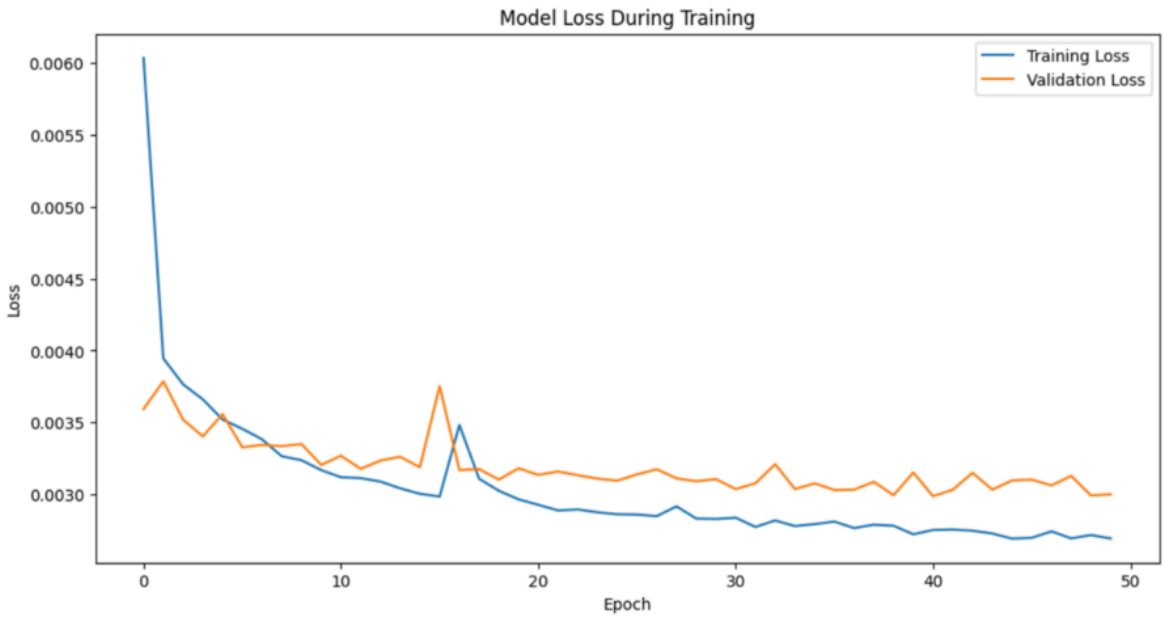


FIGURE 6: Loss vs Epoch Curve of Raincast Model

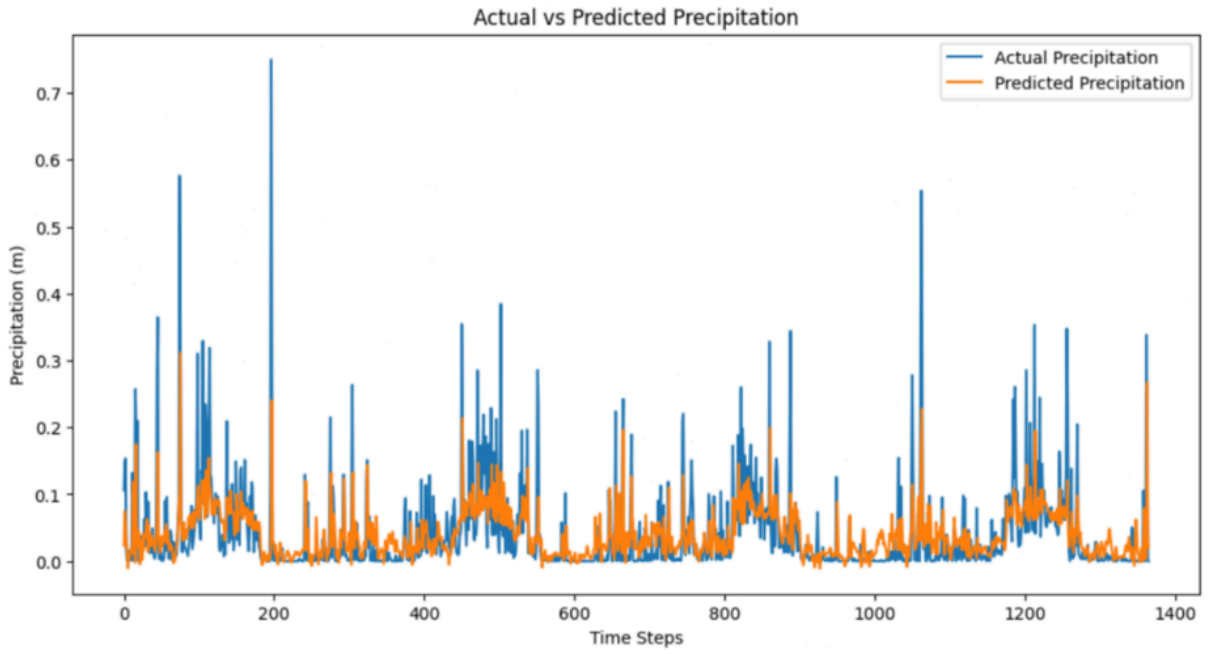


FIGURE 7: Actual vs Predicted Precipitation

Terracast carried out its performance test with a focus on identifying non-linear dynamic patterns, such as surge patterns, using the MAE metric. Loss vs epoch graph and prediction values of Terracast model are shown in Table *6*, Figure *8*, and Figure *9*, respectively.

Metric	Value
MAE (m/yr)	0.4173
RMSE (m/yr)	0.994
R^2 Score	0.9898

TABLE 6: Terracast – Key Performance Metrics

MAE, Mean Absolute Error; RMSE, Root Mean Squared Error

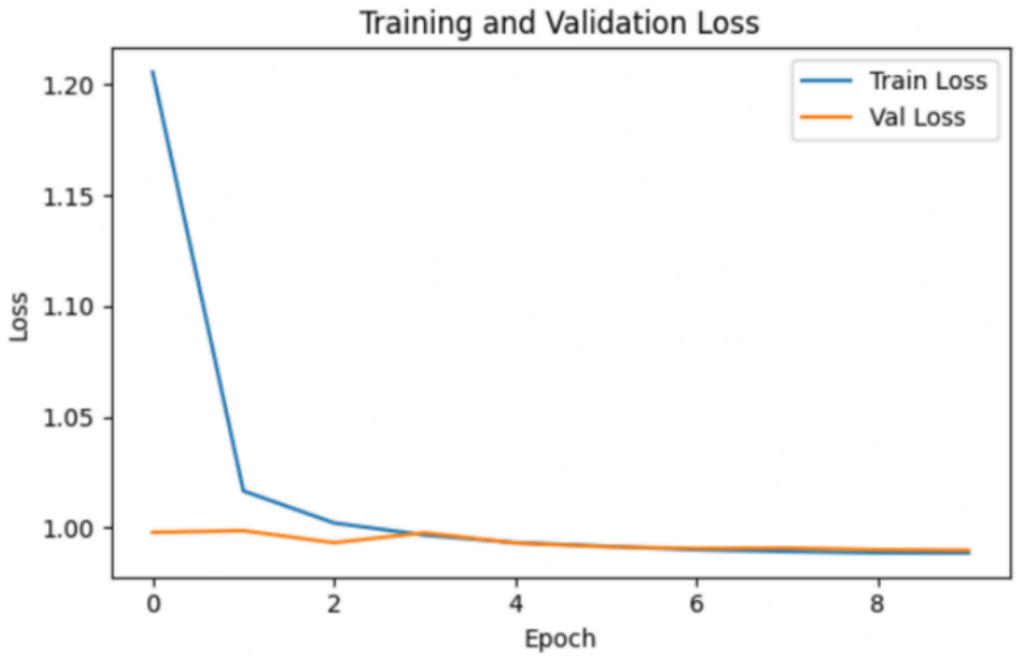


FIGURE 8: Loss vs Epoch Curve of Terracast Model

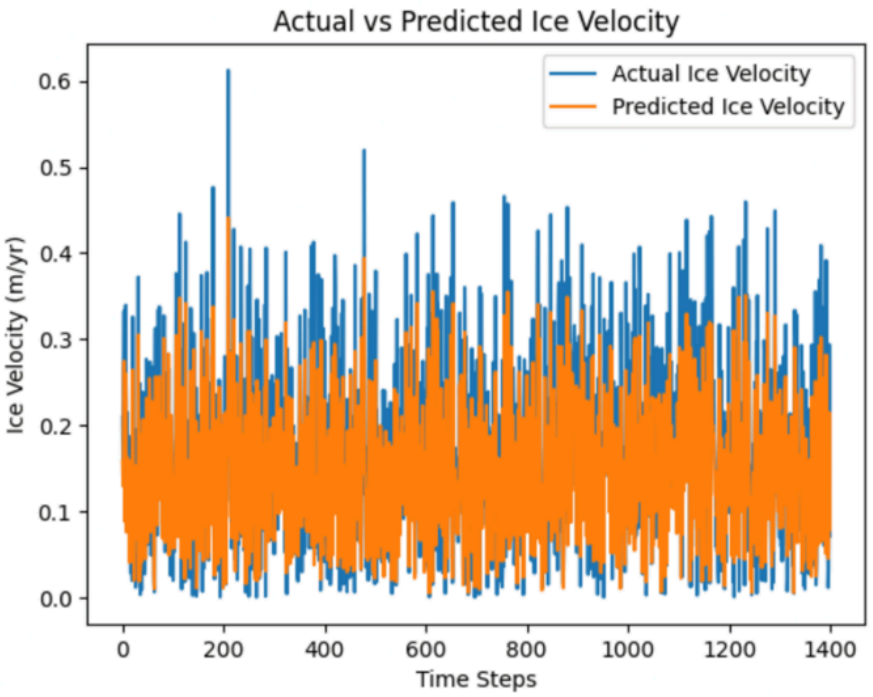


FIGURE 9: Actual vs Predicted Ice Velocity

The Tempcast model proved to have strong forecasting capabilities with minimal error in the presence of strong seasonal variation. Loss vs epoch graph and prediction values of Tempcast model are shown in Table *7*, Figure *10*, and Figure *11*, respectively.

Metric	Value
MAE (K)	3.85
RMSE (K)	5.13
R^2 Score	0.533

TABLE 7: Tempcast – Key Performance Metrics

MAE, Mean Absolute Error; RMSE, Root Mean Squared Error

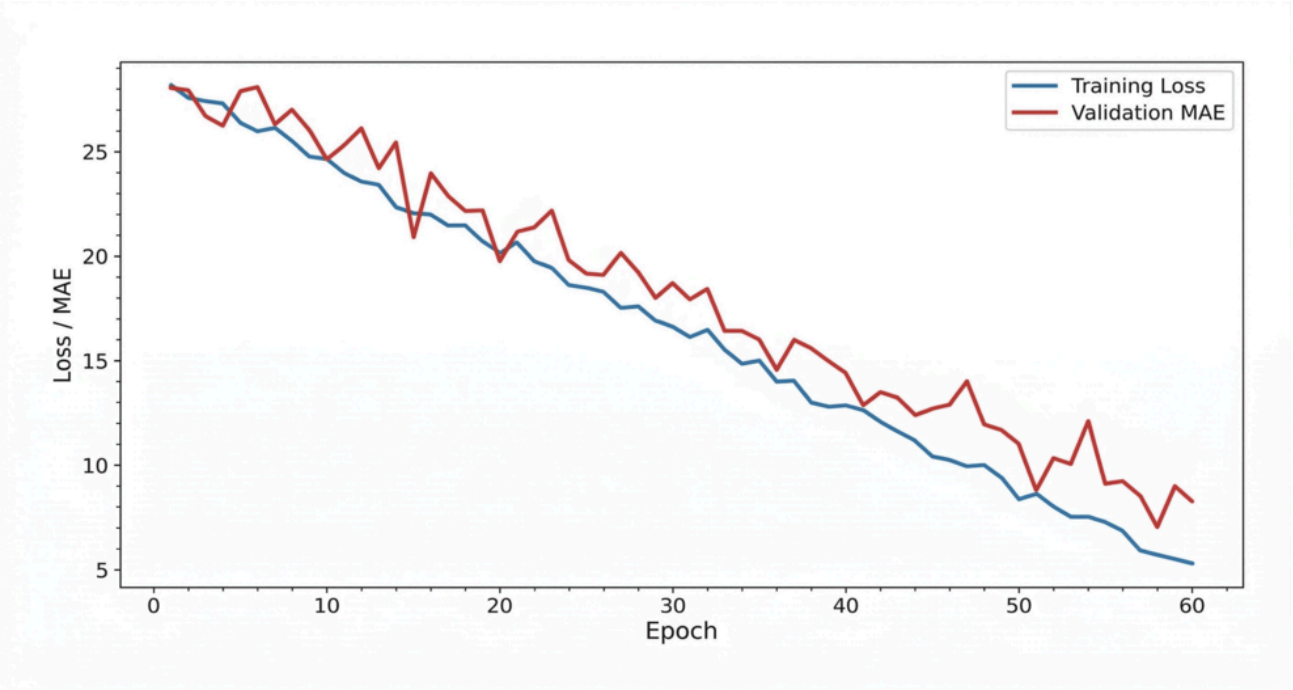


FIGURE 10: Loss vs Epoch Curve of Tempcast Model

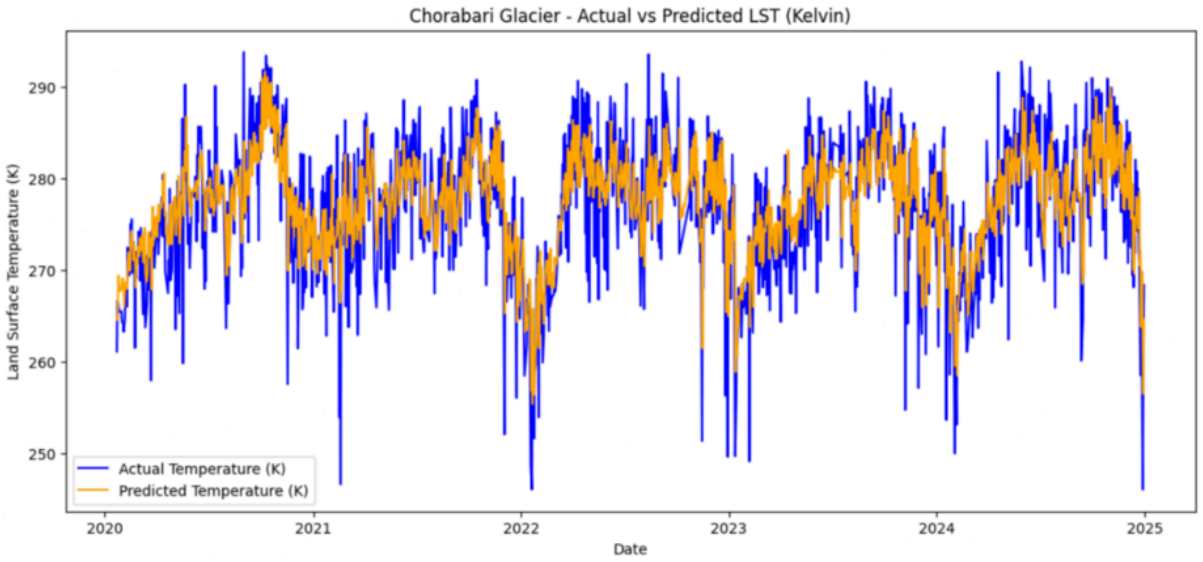


FIGURE 11: Actual vs Predicted Temperature

For instance, in situations where the model tends to produce high false positives when there is strong melt activity, the subsequent optimization process can focus solely on improving the Tempcast model or the temperature weighting factor. The significance of various weights on model accuracy is shown in Table *8*. Table *9* shows the comparison of the three deep learning modules in terms of architecture, output, performance metric, and trigger threshold.

Scenario	Terracast	Tempcast	Raincast	Accuracy	Key Insight
Optimal (Current)	50%	30%	20%	High (~91.5%)	Velocity most reliable
Equal Weight	33%	33%	33%	Medium (~85%)	Weak signals dilute result
Velocity Dominant	70%	20%	10%	Slightly Lower (~90%)	Overfocus on velocity
Thermal Dominant	20%	60%	20%	Low (~75%)	Temp noisy, unstable

TABLE 8: Impact of Weighting Strategies on Model Accuracy

Deep Learning Module	Model Architecture	Primary Output	Key Performance Metric	Operational Trigger Threshold
Terracast	Transformer-Based Encoder (6 layers)	Ice Velocity (m/yr)	MAE: 0.4173, RMSE: 0.994, R^2: 0.9898	Surge Events >200~m/yr
Tempcast	Stacked LSTM (3 layers)	Surface Temperature (K)	MAE: 3.85, RMSE: 5.13, R^2: 0.533	Warming Anomaly Spikes >6°C
Raincast	Hybrid CNN-LSTM (Conv1D + LSTM)	Precipitation Level (mm)	MAE: 3.90, RMSE: 5.17, R^2: 0.525	Heavy Rainfall >100 mm

TABLE 9: Comparison in Terms of Architecture, Primary Output, Key Metric, and Operational Triggers

CNN, Convolutional Neural Network; LSTM, Long Short-Term Memory; MAE, Mean Absolute Error; RMSE, Root Mean Squared Error

The SafeFlow model also utilizes the Chorabari Lake (Gandhi Sarovar) region, which was the central site of the 2013 Kedarnath outburst. This provides a direct real-world benchmark to test if the deep learning models could have signalled the disaster in advance.

By running the SafeFlow architecture on data from the months leading up to June 2013, the following evidence is established. Terracast Transformer-based Terracast model identifies a surge in glacier surface velocity exceeding the 200 m/year operational threshold in the week preceding the breach. The stacked LSTM Tempcast model identifies a warming anomaly where surface temperatures significantly exceeded the 6°C melt threshold, accelerating lake expansion. The hybrid CNN-LSTM Raincast forecasts heavy rainfall events exceeding 100 mm, which provided the final hydrological trigger for the outburst. Risk score evolution of the individual components is shown to have crossed their danger thresholds simultaneously. The probabilistic risk score (weighted 50% velocity, 30% temperature, 20% rain) showed a significant and sustained increase starting in early June 2013. The system's logic would have bypassed standard calculations to trigger an immediate notification once these thresholds were crossed, proving its predictive power in a real-world scenario. Table *10* shows how individual physical precursors (Velocity, Temperature, and Rainfall) crossed their operational danger thresholds, leading to a maximum risk score of 100.0 during the disaster peak.

Date	Pred. Velocity (m/yr)	Pred. Temp (°C)	Pred. Rain (mm)	Risk Score (0-100)	Risk Level
2013-06-05	219.61 [Exceeds]	10.48 [Exceeds]	-15.71	80.00	Extreme/Critical
2013-06-09	199.12	7.61 [Exceeds]	49.32	89.64	Extreme/Critical
2013-06-10	307.67 [Exceeds]	17.32 [Exceeds]	97.39	99.48	Extreme/Critical
2013-06-12	273.40 [Exceeds]	17.09 [Exceeds]	56.95	91.39	Extreme/Critical
2013-06-13	276.53 [Exceeds]	18.78 [Exceeds]	125.90 [Exceeds]	100.00	Extreme/Critical
2013-06-15	316.21 [Exceeds]	17.15 [Exceeds]	107.10 [Exceeds]	100.00	Extreme/Critical
2013-06-17	254.61 [Exceeds]	14.92 [Exceeds]	173.60 [Exceeds]	100.00	Extreme/Critical

TABLE 10: Model Performance on 2013 Kedarnath Outburst

This validation demonstrates that SafeFlow provides a proactive, phenomena-specific warning rather than a reactive one. It confirms the framework's ability to integrate diverse data streams into an operational signal that aligns with documented historical disasters.

Comparative analysis with traditional GLOF systems

The SafeFlow multi-modal, AI-based system represents a much-needed paradigm shift in GLOF risk management, providing a greatly improved alternative to traditional approaches. The first and foremost comparative advantage of the system is its shift from site-specific risk management (as is common in traditional early warning systems or site-specific maps) to a phenomena-specific system. Since SafeFlow is based on general remote-sensing data and generalizable deep learning models, the system can easily be scaled and applied to any glacial region at risk worldwide. The system's low cost and high scalability make SafeFlow economically viable for national disaster management authorities, especially in developing countries, allowing a shift from expensive and unreliable ground networks to scalable and automated risk management systems.

Table *11* shows comparison between the approach used in the existing GLOF detection techniques. The system allows potential users to isolate bottlenecks in performance.

Features	Approach Used in Existing GLOF Detection Systems	SafeFlow (Proposed Method)
Data Source	Ground sensors, static maps	Dynamic satellite + temporal data
Forecasting Type	Rule-based alerts	AI-driven forecasting
Multivariable Analysis	No	Yes (Rain Level + velocity + temperature)
Scalability	Limited	Global (remote sensing)
Cost	High (maintenance, logistics)	Low (cloud-based, automated)

TABLE 11: SafeFlow vs. Existing GLOF Detection Methods

GLOF, Glacial Lake Outburst Flood

Conclusions

SafeFlow framework is a multi-modal artificial intelligence framework intended to enhance the accuracy and timeliness of GLOF risk assessment. It integrates three specialized deep learning models trained on satellite and climate datasets. Terracast, a Transformer-based model, predicts glacier surface velocity using the ITS_LIVE Glacier Velocity Dataset derived from Sentinel-1 SAR remote sensing products. Tempcast, a stacked LSTM model, forecasts temperature variations using the MODIS MOD11A1 LST dataset (V6.1) obtained from NASA Terra/Aqua satellites. Raincast, a hybrid CNN-LSTM model, predicts precipitation patterns using the ERA5-Land Daily Aggregates dataset from the ECMWF ERA5 climate reanalysis archive. Region of Interest considered for disaster monitoring scenario was the Chorabari Lake region. The experiment setup has shown that the models have strong predictive capabilities, with the Terracast model demonstrating high accuracy in predicting glacier surface movements, with the model achieving MAE = 0.4173, RMSE = 0.994, and R^2 = 0.9898. The climate-based models have shown high forecasting stability, with Tempcast achieving MAE = 3.85, RMSE = 5.13, and R^2 = 0.533, and Raincast achieving MAE = 3.90, RMSE = 5.17, and R^2 = 0.525. The outputs from the models are combined using a weighted ensemble method to produce a probability-based GLOF risk score, which incorporates the effects of different climatic and geophysical parameters. SafeFlow has also been able to define different operational trigger thresholds that can be used to identify the possibility of high-risk GLOF, such as an increase in surface velocity (≥200 m/year), temperatures (>6°C), and rainfall (>100 mm). This framework, therefore, provides a shift from the traditional reactive approach to GLOF risk assessment to a proactive approach, which can be useful in the disaster monitoring scenario in the Chorabari Lake region.

SafeFlow model is designed as a precursor-based early warning system and does not currently incorporate the mechanical hydrodynamics of the dam breach process itself (e.g., breach width or peak discharge). The current version does not account for sudden external triggers such as seismic activity (earthquakes) or ice avalanches, which can cause instantaneous displacement waves. Also the model focuses on the probability of the initiation event at the lake site rather than the routing of the flood wave through downstream valleys, which can be considered as the limitation of the proposed model. As 60% of the MODIS data can be of low quality due to cloud cover interference. To prevent this uncertainty from compromising the risk score, SafeFlow employs multiple imputation and quality-filtering methods. The

Tempcast model acts as an interpolation tool to maintain data continuity during high cloud cover, achieving a forecasting stability with an MAE of 3.85 K. Future research directions include integrating real-time IoT sensor data, such as water level, vibration, and dam deformation sensors. This can allow for hybrid monitoring, where satellite data can be integrated with real-time IoT sensor data, thus improving the overall accuracy of short-term warning systems. Additionally, the risk scores derived through the proposed framework can be integrated with hydrodynamic flood propagation models to simulate downstream flood behavior. Another extension of the proposed framework can be integrating multiple glacial lake monitoring, thus facilitating regional monitoring dashboards, which can cater to overall risk surveillance.

Appendices

Datasets Used in SafeFlow System

1. ERA5-Land Daily Aggregated Climate Dataset (RainCast Model)

Source: https://developers.google.com/earth-engine/datasets/catalog/ECMWF_ERA5_LAND_DAILY_AGGR

See Muñoz Sabater [32]

2. ITS_LIVE Glacier Velocity Dataset (Terracast Model)

Source: https://nsidc.org/apps/itslive/

See Gardner et al. [33]

3. MODIS Land Surface Temperature Dataset (TempCast Model)

Source: https://developers.google.com/earth-engine/datasets/catalog/MODIS_061_MOD11A1

See Wan et al. [34]

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Asha Rani Mishra, Asmit Patel, Vedanshi ., Krishna Yadav, Molisha Agarwal

Acquisition, analysis, or interpretation of data: Asha Rani Mishra, Asmit Patel, Vedanshi ., Krishna Yadav, Molisha Agarwal

Drafting of the manuscript: Asha Rani Mishra, Asmit Patel, Vedanshi ., Krishna Yadav, Molisha Agarwal

Critical review of the manuscript for important intellectual content: Asha Rani Mishra, Asmit Patel, Vedanshi ., Krishna Yadav, Molisha Agarwal

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The datasets (and/or code) supporting this study are available from the corresponding author upon reasonable request.
https://github.com/Asmit556/Safeflow-Main-Repository
All data generated or analyzed during this study are included in this published article and/or its appendices.

References

1. Kausarian H, Sumantyo JT, Batara, Suryadi A, Pangestu T: SAR sentinel data analysis: hydrological dynamics and rainfall patterns in the Kampar River basin (2018-2023). Evergreen. 2024, 11:1558-1567. 10.5109/7236811
2. Adhikary S, Kumar R, Chettri R, Sharma RK, Gupta V, Kami A: Glacial lake outburst flood modelling of the Tikip Lake, Sikkim Himalaya, India [IN PRESS]. Natural Hazards Research. 2025, 10.1016/j.nhres.2025.11.005
3. Lützow N, Veh G, Korup O: A global database of historic glacier lake outburst floods. Earth System Science Data. 2023, 15:2983-3000. 10.5194/essd-15-2983-2023
4. Zhang G, Yao T, Huss M, et al.: A monitoring network for mitigating Himalayan glacial lake outburst floods. Bulletin of the American Meteorological Society. 2025, 106:E2579-E2597. 10.1175/BAMS-D-24-0290.1
5. Shah S, Ishtiaque A: Adaptation to glacial lake outburst floods (GLOFs) in the Hindukush-Himalaya: a review. Climate. 2025, 13:60. 10.3390/cli13030060
6. Sridevi A, Sugeshan S, Sudhakar K, Vignesh E, Vigneshwaran A: Early warning system for glacial lake outburst using multi source data. 2025 6th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India. 2025, 1266-1274. 10.1109/ICIRCA65293.2025.11089725
7. Patil LN, Khairnar HP: Investigation of human safety based on pedestrian perceptions associated to silent nature of electric vehicle. Evergreen. 2021, 8:280-289. 10.5109/4480704
8. Mishra AR, Pippal SK, Asif, Kumar A, Singh D, Singh A: Clear vision-obstacle detection using bat algorithm optimization technique. 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India. 2021, 1-5. 10.1109/ICRITO51393.2021.9596467
9. Yussupov A, Suleimenova RZ: Use of remote sensing data for environmental monitoring of desertification. Evergreen. 2023, 10:300-307.
10. Gurung DR, Khanal NR, Bajracharya SR, et al.: Lemthang Tsho glacial lake outburst flood (GLOF) in Bhutan: cause and impact. Geoenvironmental Disasters. 2017, 4:17. 10.1186/s40677-017-0080-2
11. Emmer A, Vilímek V: Lake and breach hazard assessment for moraine-dammed lakes: an example from the Cordillera Blanca (Peru). Natural Hazards and Earth System Sciences. 2013, 13:1551-1565. 10.5194/nhess-13-1551-2013
12. Meena DK, Singh S, Singh SK, et al.: Seasonal variations and water quality dynamics: analysis of Kanota Dam in relation to WHO standards. Evergreen. 2024, 11:771-786. 10.5109/7183357
13. Wang W, Zhang T, Yao T, An B: Monitoring and early warning system of Cirenmaco glacial lake in the central Himalayas. International Journal of Disaster Risk Reduction. 2022, 73:102914. 10.1016/j.ijdrr.2022.102914
14. Shrestha B, Mool PK, Bajracharya SR: Impact of climate change on Himalayan glaciers and glacial lakes: case studies on GLOF and associated hazards in Nepal and Bhutan. 2007.
15. Fujita K, Sakai A, Takenaka S, Nuimura T, Surazakov AB, Sawagaki T, Yamanokuchi T: Potential flood volume of Himalayan glacial lakes. Natural Hazards and Earth System Sciences. 2013, 13:1827-1839. 10.5194/nhess-13-1827-2013
16. Worni R, Huggel C, Stoffel M: Glacial lakes in the Indian Himalayas — From an area-wide glacial lake inventory to on-site and modeling based risk assessment of critical glacial lakes. Science of the Total Environment. 2013, 468-469:S71-S84. 10.1016/j.scitotenv.2012.11.043
17. Ahmed R, Sam L, Bhardwaj A: It is time to build GLOF-resilient communities globally before another GLOF disaster strikes. npj Natural Hazards. 2025, 2:41. 10.1038/s44304-025-00097-0
18. Chaudhary HK, Saraswat K, Yadav H, Puri H, Mishra AR, Chauhan SS: A real time dynamic approach for management of vehicle generated traffic. Evergreen. 2023, 10:289-299. 10.5109/6781078

19. Pandey P, Banerjee D, Ali SN, Khan MA, Chauhan P, Singh S: Simulation and risk assessment of a possible glacial lake outburst flood (GLOF) in the Bhilangna Valley, central Himalaya, India. Journal of Earth System Science. 2022, 131:184.

20. Sharma A, Thakur V, Prakash C, Sharma A, Sharma R: Deep learning-based glacial lakes extraction and mapping in the Chandra-Bhaga Basin. Journal of the Indian Society of Remote Sensing. 2024, 52:435-447. 10.1007/s12524-024-01829-x

21. Maslov KA, Persello C, Schellenberger T, Stein A: GLAVITU: a hybrid CNN-Transformer for multi-regional glacier mapping from multi-source data. IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA. 2023, 1233-1236. 10.1109/IGARSS52108.2023.10281828

22. Panwar P, Roshan P, Singh R, Rai M, Mishra AR, Chauhan SS: DDNet-a deep learning approach to detect driver distraction and drowsiness. Evergreen. 2022, 9:881-892. 10.5109/4843120

23. Ahmed R: Glacial lake outburst flood (GLOF) hazard and risk management strategies: a global overview. Water Resources Management. 2025, 39:1-16. 10.1007/s11269-024-03958-x

24. Zhou B, Zou Q, Jiang H, Yang T, Zhou W, Chen S, Yao H: A novel framework for predicting glacial lake outburst debris flows in the Himalayas amidst climate change. Science of The Total Environment. 2024, 946:174435. 10.1016/j.scitotenv.2024.174435

25. Mohanty A, Srivastava PK, Aggarwal A: Review of glacier velocity and facies characterization techniques using multi-sensor approach. Environment, Development and Sustainability. 2025, 27:17753-17804. 10.1007/s10668-024-04604-7

26. Kumar B, Murugesh Prabhu TS, Sathyan A, Krishnan A: GLOF Early Warning System: Computational challenges and solutions. Current Directions in Water Scarcity Research. McLeman R, Bruntru M (ed): Elsevier, Amsterdam; 2022. 7:641-662. 10.1016/B978-0-323-91910-4.00036-4

27. Wang X, Chen W, Zhang G, et al.: Mitigating future glacial lake outburst floods in the Himalaya. Science Bulletin. 2025, 71:159-171. 10.1016/j.scib.2025.11.024

28. Rada C, Rivera A, Alfaro S: Development of an early warning system to reduce the impact of floods related to glacial lake outburst floods (SAGAZ). The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. 2024, XLVIII-2/W6-2024:37-43. 10.5194/isprs-archives-xlviii-2-w6-2024-37-2024

29. Wangchuk T, Tsubaki R: A glacial lake outburst flood risk assessment for the Phochhu river basin, Bhutan. Natural Hazards and Earth System Sciences. 2024, 24:2523-2540. 10.5194/nhess-24-2523-2024

30. Acharya S, Khadka N, Shrestha D, Dulal KN, Adhikari N, Pandey VP: Evaluating the socio-economic risks of a potential GLOF from Dudh Pokhari lake in Nepal's Everest Region [IN PRESS]. Natural Hazards Research. 10.1016/j.nhres.2025.08.001

31. Chapagain D, Shrestha F, Joshi S, Shrestha S, Schneiderbauer S: Glacial lake outburst flood risks, response strategies and adaptation gaps in the Hindu Kush Himalaya region. (2025). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5517478.

32. Muñoz Sabater J: ERA5-Land hourly data from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (2019). https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land?tab=overview.

33. Gardner AS, Greene CA, Kennedy JH, et al.: ITS_LIVE global glacier velocity data in near-real time. The Cryosphere. 2025, 19:3517-3533. 10.5194/tc-19-3517-2025

34. Wan Z, Hook S, Hulley G: MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V061 [Dataset]. NASA Land Processes Distributed Active Archive Center. (2021). https://doi.org/10.5067/MODIS/MOD11A1.061.

How to cite this article:
Patel A, Mishra A, . V, et al. (April 06, 2026) SafeFlow: Glacial Lake Outburst Floods Early Warning System With Probability-Based Risk Scoring. Cureus J Comput Sci 3 : es44389-026-00050-w. DOI https://doi.org/10.7759/s44389-026-00050-w