https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

The Cryosphere Discussions
Open Access
EGU

Controls of outbursts of moraine-dammed lakes in the greater Himalayan region

Melanie Fischer[1], Oliver Korup[1,2], Georg Veh[1], Ariane Walz[1]

[1]Institute of Environmental Science and Geography, University of Potsdam, Potsdam, 14476, Germany
[2]Institute of Geosciences, University of Potsdam, Potsdam, 14476, Germany

Correspondence to: Melanie Fischer (melaniefischer@uni-potsdam.de)

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020


The Cryosphere Discussions

Abstract

Glacial lakes in the Hindu-Kush Karakoram Himalaya Nyainqentanglha (HKKHN) have grown rapidly in number and area in past decades, and some dozens have drained in catastrophic glacial lake outburst floods (GLOFs). Estimating hazard from glacial lakes has largely relied on qualitative assessments and expert judgment, thus motivating a more systematic and quantitative appraisal. Before the backdrop of current climate-change projections and the potential of elevation-dependent warming, an objective and regionally consistent assessment is urgently needed. We use a comprehensive inventory of 3,390 moraine-dammed lakes and their documented outburst history in the past four decades to test whether elevation, lake area and its rate of change, glacier-mass balance, and monsoonality are useful inputs to a probabilistic classification model. We use these candidate predictors in four Bayesian multi-level logistic regression models to estimate the posterior susceptibility to GLOFs. We find that mostly larger lakes have been more prone to GLOFs in the past four decades, largely regardless of elevation band in which they occurred. We also find that including the regional average glacier-mass balance improves the model classification. In contrast, changes in lake area and monsoonality play ambiguous roles. Our study provides first quantitative evidence that GLOF susceptibility in the HKKHN scales with lake area, though less so with its dynamics. Our probabilistic prognoses offer some improvement with respect to a random classification based on average GLOF frequency. Yet they also reveal some major uncertainties that have remained largely unquantified previously and that challenge the applicability of single models. Ensembles of multiple models could be a viable alternative for more accurately classifying the susceptibility of moraine-dammed lakes to GLOFs.

1 Introduction

Glacial lake outburst floods (GLOFs) involve the sudden release and downstream propagation of water and sediment from naturally impounded meltwater lakes (Costa and Schuster, 1987; Emmer, 2017). About one third of the 25,000 glacial lakes in the Hindu-Kush Karakoram Himalaya Nyainqentanglha (HKKHN) are dammed by potentially unstable moraines (Maharjan et al., 2018). Some of this impounded meltwater can overtop or incise dams rapidly, with catastrophic consequences downstream (Costa and Schuster, 1987; Evans and Clague, 1994). High Mountain Asian countries are among the most affected by these abrupt floods, if considering both damage and fatalities (Carrivick and Tweed, 2016). For example, in June 2013, a GLOF from Chorabari Lake in the Indian state of Uttarakhand, caused >6,000 deaths in what is known as the “Kedarnath disaster” (Allen et al., 2016). The peak discharges of GLOFs can be orders of magnitude higher than those of seasonal floods. GLOFs can move large amounts of sediment, widen mountain channels, undermine hillslopes, and thus increase the hazard to local communities (Cenderelli and Wohl, 2003). Still, GLOFs in the HKKHN are rare and have occurred at an unchanged rate of about 1.3 per year in the past four decades (Veh et al., 2019). Ice avalanches and glacier calving are the most frequently reported triggers of GLOFs in the HKKHN. Most outbursts with known date (mostly June to October) might be also linked to

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020


The Cryosphere Discussions

high lake levels fed by monsoonal precipitation and summer ablation of glaciers (Richardson and Reynolds, 2000). The Kedarnath GLOF is the only case attributed to a rain-on-snow event early in the monsoon season (Allen et al., 2016). This particularly destructive GLOF underlines the need for understanding better how and why meltwater lakes can be susceptible to sudden outburst triggered by rainstorms, especially given projected impacts of atmospheric warming on the high-mountain cryosphere.

Current scenarios entail that atmospheric warming may change the susceptibility of HKKHN glacial lakes to sudden outburst floods: IPCC's most recent prognoses link the decay of low-lying glaciers and permafrost to commensurate increases in lake number and area because of rising air temperatures, more frequent rain-on-snow events at higher elevations, and changes in precipitation seasonality (Hock et al., 2019). Air surface temperature in the HKKHN rose by about 0.1 °C per decade from 1901 to 2014 (Krishnan et al., 2019), likely having reduced snowfall, altered permafrost distribution, and accelerated glacier melt at lower elevations (Hock et al., 2019). Ice loss in the Himalayas has significantly increased in the past four decades, from -0.22 ± 0.13 m w.e. y^{-1} (meters of water equivalent per year) between 1975 and 2000 to -0.43 ± 0.14 m w.e. y^{-1} between 2000 and 2016 (Maurer et al., 2019). Parts of this meltwater have been trapped in glacial lakes that have expanded by approximately 14.1% between 1990 and 2015 (Nie et al., 2017). The notion of elevation-dependent warming (EDW) posits that increases in air temperature are most pronounced at higher elevations (Hock et al., 2019; Pepin et al., 2015), and that EDW has affected cold temperature metrics, including the number of frost days and minima of near-surface air temperature in the HKKHN in the past decades (Krishnan et al., 2019; Palazzi et al., 2017). Essentially, all scenarios of atmospheric warming concern aspects of elevation, glacier-lake size and dynamics, and local climatic variability. Yet whether and how these aspects affect GLOF hazard still awaits more quantitative support.

Previous work on GLOF hazard in the region focused on identifying or classifying potentially unstable glacial lakes, including local case studies largely informed by fieldwork, dam-breach models (Koike and Takenaka, 2012; Somos-Valenzuela et al., 2012, 2014), and basin-wide assessments (Bolch et al., 2011; Mool et al., 2011; Rounce et al., 2016; Wang et al., 2011). GLOF hazard appraisals for the entire HKKHN, however, remain rare (Veh et al., 2020). Most basin-wide studies proposed qualitative to semi-quantitative decision schemes using selective lists of presumed GLOF predictors (Table 1; Rounce et al., 2016). Yet researchers have used subjective rules to choose these variables and associated thresholds, leading to diverging hazard estimates (Rounce et al., 2016). Expert knowledge has thus been essential in GLOF hazard appraisals, despite an increasing amount of freely available climatic, topographic, and glaciological data. Statistical models can help to estimate the occurrence probability of GLOFs, and thus reduce the inherent subjective bias (Emmer and Vilímek, 2013). For example, Wang et al. (2011) classified the outburst potential of moraine-dammed lakes on the southeastern Tibetan Plateau by applying a fuzzy consistent matrix method. They used as inputs the size of the parent glacier, the distance and slope between lake and glacier snout, and the mean steepness of the moraine dam and the glacier snout to come up with different nominal hazard categories.

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020


The Cryosphere Discussions

This and many similar qualitative ranking schemes are accessible to a broader audience and policy makers, but are difficult to compare and potentially oversimplify uncertainties.

One way to deal with these uncertainties in a more objective way involves a Bayesian approach. Here, we used this probabilistic reasoning utilising fully data-driven models. Specifically, we tested how well some of the more widely used diagnostics of GLOF susceptibility fare as predictors in a multi-level logistic regression that is informed more by data than by expert opinion. We checked whether this approach can identify glacial lakes in the HKKHN that had released GLOFs in the past four decades. We discuss what we can learn about how these historic GLOFs were linked to readily available measures of topography, monsoonality, and glaciological changes.

2 Study area, data, and methods

2.1 Study area and data

We studied glacial lakes of the Hindu-Kush Karakoram Himalaya Nyainqentanglha (HKKHN) region that we defined here as the Asian mountain ranges between 16° to 39°N and 61° to 105°E, i.e. from Afghanistan to Myanmar (Fig. 1; Bajracharya and Shrestha, 2011). Following the outlines of glacier regions in High Mountain Asia used in the Randolph Glacier Inventory (RGI, Pfeffer et al., 2014) with slight modifications, we subdivided our study area into the following seven mountain ranges: the Hindu Kush, the Karakoram, the Western Himalaya, the Central Himalaya, the Eastern Himalaya, the Nyainqentanglha, and the Hengduan Shan. Meltwater from the HKKHN's extensive snow and ice cover, often referred to as "Third Pole", feeds ten major river systems to provide water for some 1.3 billion people (Molden et al., 2014). There, glaciers have had an overall negative mass balance historically, having lost 150 ± 110 kg m^{-2} yr^{-1} on average from 2006 to 2015, with slightly, but exceptional, positive trends in the Karakoram and Western Himalaya (Hock et al., 2019). Since the 1970s, some Karakoram glaciers also accelerated in flow, whereas glaciers stalled elsewhere in the HKKHN (Dehecq et al., 2019). In the RCP8.5 scenario the HKKHN glaciers lose 64 ± 5% of their total mass until 2100 compared to 1995 to 2015 (Kraaijenbrink et al., 2017). How much of this melting of glaciers is due to EDW remains debated (Palazzi et al., 2017; Rangwala and Miller, 2012; Tudoroiu et al., 2016). Snowfall at lower elevations is also likely to decrease (Hock et al., 2019; Terzago et al., 2014), judging from snowfall and glacier-mass balances of past decades (Kapnick et al., 2014; King et al., 2019). Monsoon precipitation is likely to become more episodic and intensive (Palazzi et al., 2013).

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

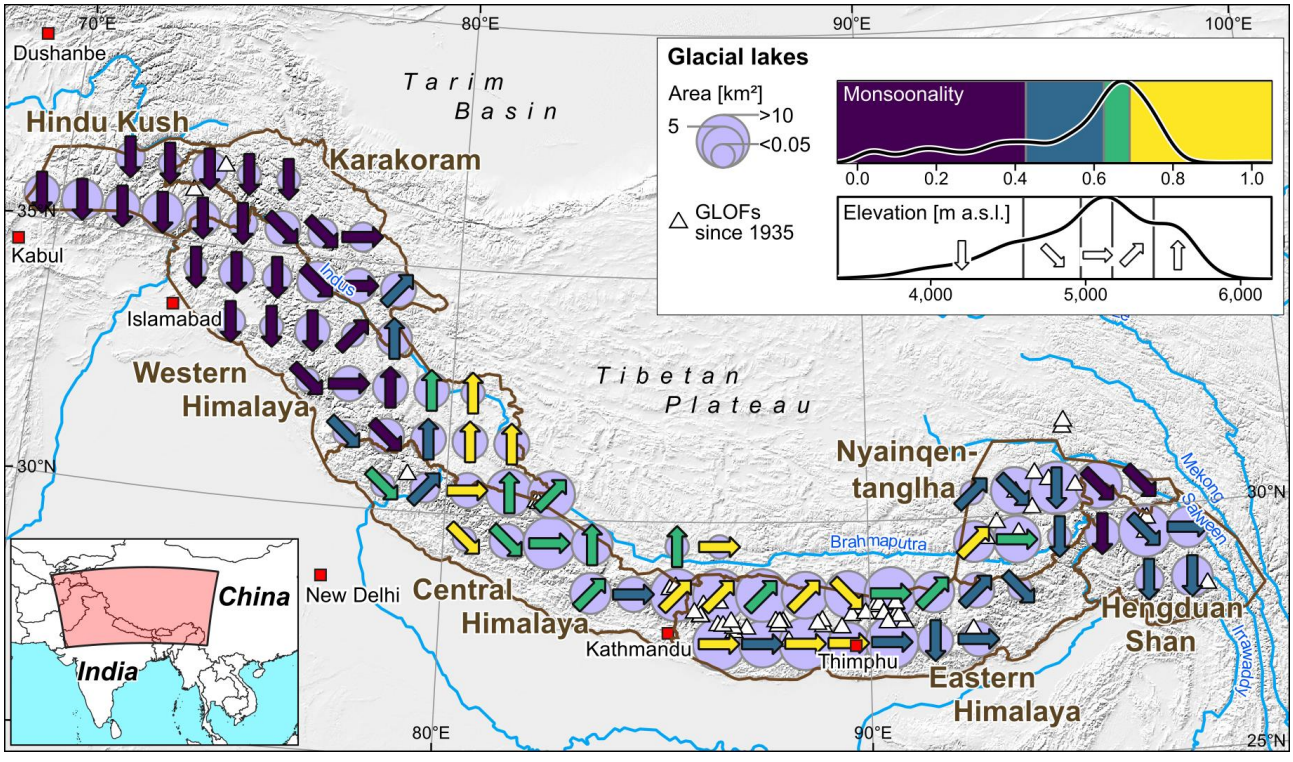


Figure 1: Overview map of the Hindu-Kush Karakoram Himalaya Nyainqentanglha (HKKHN) mountains showing distribution of moraine-dammed lakes (blue bubbles scaled by area), their elevation (expressed as quantiles coded by arrows; see inset for elevation distribution); and average monsoonality (colour coded; see inset for monsoonality distribution), defined here as the fraction of total annual precipitation falling in the summer months. Triangles indicate reported glacial lake outburst floods (GLOFs) in the study area since 1935 (Veh et al., 2019). The topographic map was created with Global 30 Arc-Second Elevation data (GTOPO30, https://doi.org/10.5066/F7DF6PQS).

Guided by these projections, we selected several widely used diagnostics of GLOF potential (Table 1). We used lake elevation as a proxy for the standard lapse rate of tropospheric air temperature (Rolland, 2003; Yang and Smith, 1985). This elevation-dependent thermal gradient is also a major control on the distribution of alpine permafrost (Etzelmüller and Frauenfelder, 2009) and precipitation. Mean annual rainfall along the Himalayan front can exceed 4,000 mm at elevations some 4,000 m high, where c. 25% of all glacial lakes occur (Fig. 1; Bookhagen and Burbank, 2010). Lake elevation should also represent to first order topographic effects of EDW. For example, the stability of low-lying moraine dams may be compromised by the loss of permafrost and commensurate increases in permeability in the moraine barrier and adjacent valley slopes (Haeberli et al., 2017). Glacial lake area and its rate of change are another common diagnostic in GLOF studies (Allen et al., 2019; Bolch et al., 2011; Prakash and Nagarajan, 2017; see Table 1 for full list of references) that we considered here. Lake area is a proxy for lake volume (Huggel et al. 2002), and growing lakes increase the hydrostatic pressure acting on moraine dams, thus raising

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

the potential of failure (Rounce et al., 2016). Since 1990, lake areas have grown largest in the Central Himalayas (+23%), and lowest in the northwest Himalayas (+5.0%) (Nie et al., 2017), and many studies have emphasised the role of growing lakes on GLOF hazard (Bolch et al., 2011; Prakash and Nagarajan, 2017; Rounce et al., 2016) Yet to our best knowledge few, if any, studies offered tests of whether and how this change increased the susceptibility to sudden outburst. Similarly, glacier dynamics often find mention in GLOF studies, but are hardly quantified or used in quantitative models (Bolch et al., 2011; Ives et al., 2010). This motivated us to consider average changes in regional glacier-mass balances from 2000 to 2016 by Brun et al. (2017). Meteorological drivers entered previous qualitative GLOF hazard appraisals mostly as (the probability of) extreme precipitation events (Huggel et al., 2004; Prakash and Nagarajan, 2017). In the absence of suitable data we used a synoptic measure of monsoonality instead in terms of the annual proportion of summer precipitation. This proportion is highest in the southeast HKKHN, where it is linked to monsoonal low-pressure systems (Krishnan et al., 2019). Different precipitation regimes and climatic preconditions may influence mechanisms of moraine dam failure (Wang et al., 2012).

Table 1: Frequently used diagnostics of GLOF hazard in the HKKHN. Units and data sources refer to parameters used in this study.

Diagnostic groups	GLOF diagnostic parameters	Used in this study	Unit	Description	Data source	Reference
Lake characteristics and dynamics	Glacial lake elevation	✓	m asl		SRTM DEM	Mergili and Schneider, 2011
	Catchment area	✓	m²		SRTM DEM	Allen et al., 2019
	Glacial lake area	✓	m²		SRTM DEM	Aggarwal et al., 2016; Allen et al., 2019; Bolch et al., 2011; Ives et al., 2010; Mergili and Schneider, 2011; Prakash and Nagarajan, 2017; Wang et al., 2012
	Lake-area change (growth and shrinkage, absolute change)	✓	%		Wang et al., 2020	Aggarwal et al., 2016; Bolch et al., 2011; Ives et al., 2010; Mergili and Schneider, 2011; Prakash and Nagarajan, 2017; Rounce et al., 2016; Wang et al., 2012
Potential downstream impact	Lake volume	-				Aggarwal et al., 2016; Bolch et al., 2011; Kougkoulos et al., 2018; Mergili and Schneider, 2011
Moraine stability	Moraine-wall steepness	-				Allen et al., 2019; Bolch et al., 2011; Ives et al., 2010; Prakash and Nagarajan, 2017; Rounce et al., 2016; Wang et al., 2011; Worni et al., 2013
	Width-to-height ratio	-				Aggarwal et al., 2016; Bolch et al., 2011; Ives et al., 2010; Prakash and Nagarajan, 2017; Worni et al., 2013
	Lake freeboard	-				Bolch et al., 2011; Kougkoulos et al., 2018; Mergili and Schneider, 2011; Prakash and Nagarajan, 2017; Worni et al., 2013

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

	Existence of a buried ice core	-				Bolch et al., 2011; Ives et al., 2010; Rounce et al., 2016
	Dam type	-				Kougkoulos et al., 2018; Mergili and Schneider, 2011; Wang et al., 2012; Worni et al., 2013
Potential triggering mechanisms (geomorphic)	Seismic activity	-				Ives et al., 2010; Kougkoulos et al., 2018; Mergili and Schneider, 2011; Prakash and Nagarajan, 2017
	Distance from parent glacier snout	-				Aggarwal et al., 2016; Ives et al., 2010; Kougkoulos et al., 2018; Prakash and Nagarajan, 2017; Wang et al., 2011, 2012
	Steepness parent glacier snout	-				Bolch et al., 2011; Ives et al., 2010; Kougkoulos et al., 2018; Prakash and Nagarajan, 2017; Wang et al., 2011
	Regional or parent glacier-mass balance	✓	m w.e. (water equivalent) yr^{-1}		Brun et al., 2017	Bolch et al., 2011; Ives et al., 2010
	Mass movements (traces, trajectories, probabilities)	-				Allen et al., 2019; Bolch et al., 2011; Ives et al., 2010; Mergili and Schneider, 2011; Prakash and Nagarajan, 2017; Rounce et al., 2016; Worni et al., 2013
Potential triggering events (climatic)	Annual mean temperature	-	°C		CHELSA	Liu et al., 2014 (station data, Tibetan Plateau); Wang et al., 2008 (single station data)
	Temperature seasonality	-	-	Standard deviation of monthly mean temperature	CHELSA	Kougkoulos et al., 2018
	Wet-season temperature	-	°C	Mean temperature of wettest annual quarter	CHELSA	-
	Dry-season temperature	-	°C	Mean temperature of driest annual quarter	CHELSA	-
	Annual precipitation	-	mm		CHELSA	Wang et al., 2008, 2012 (station data)
	Precipitation seasonality	-	-	Coefficient of variation in monthly precipitation	CHELSA	Kougkoulos et al., 2018
	Summer precipitation	✓	mm	Precipitation of warmest annual quarter	CHELSA	-

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020


The Cryosphere Discussions

	Winter precipitation	-	mm	Precipitation of coldest annual quarter	CHELSA	-

We extracted information on these characteristics for glacial lakes recorded in two inventories. First, we used a database of 25,614 lakes manually mapped from Landsat imagery acquired in 2005 (± two years) (Maharjan et al. 2018), from which we extracted 7,284 lakes dammed mostly by lateral and end moraines. Second, we identified from an independent regional GLOF inventory (Veh et al. 2019) 31 lakes that had at least one outburst between 1981 and 2017. We focused on lakes >10,000 m² to ensure comparability between the two inventories, thus acquiring a final sample size of 3,390 lakes. Given the sparse network of weather stations in the HKKHN, we computed the monsoonality averaged for each lake from the 1-km resolution CHELSA data (Karger et al., 2017). We extracted topographic data from the void-free 30-m resolution SRTM (Shuttle Radar Topographic Mission of 2000) DEM, and use approximate lake-area changes for two intervals (1990 to 2005 and 2005 to 2018) by Wang et al. (2020). We discarded newer, higher resolved DEMs to minimise data gaps and artefacts. Overall, we considered six topographic, synoptic, and glaciological predictors (Fig. 2, Table 1). The interpolation method underlying the CHELSA data introduces correlation between climate (especially temperature) and elevation data so that we limited our models to those with poorly correlated predictors at the expense of possible other predictors such as mean annual temperature, annual precipitation totals, or their variability.

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

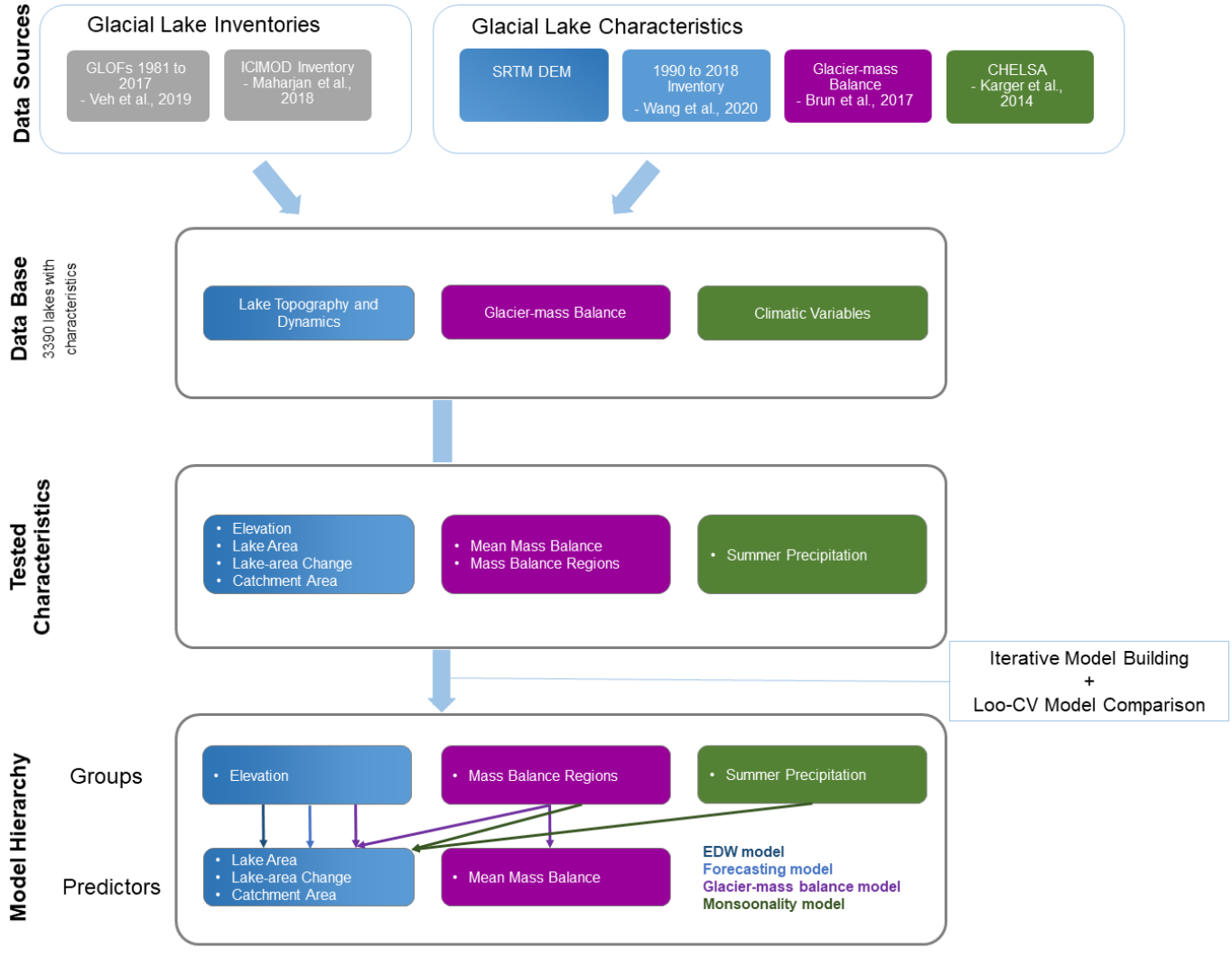


Figure 2: Data sources and workflow; EDW = elevation-dependent warming.

2.2 Bayesian multi-level logistic regression

145 We used logistic regression to learn the probability of whether a given lake in the HKKHN had a reported GLOF in the past four decades. This method was pioneered for moraine-dammed lakes in British Columbia (McKillop and Clague, 2007).

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

The Cryosphere Discussions
Open Access
EGU

Logistic regression estimates a binary outcome y from the optimal linear combination of p weighted predictors $\mathbf{x} = \{x_1, \ldots, x_p\}$. The probability $y = P_{GLOF}$ that lake i had released a GLOF is expressed as:

$$y_i \sim \text{Bernoulli}(\mu_i) \quad (1)$$

$$\mu_i = \text{S}(\alpha_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \cdots + \beta_p x_{i,p}) \quad (2)$$

where

$$\text{S}(x) = \frac{1}{1+\exp(-x)} \quad (3)$$

Here α_0 is the intercept and $\boldsymbol{\beta} = \{\beta_1, \ldots, \beta_p\}^{\text{T}}$ are the p predictor weights (Gelman and Hill, 2007). The logit function $\text{S}^{-1}(x)$ describes the odds on a logarithmic scale (the log-odds ratio) such that a unit increase in predictor x_m raises the log-odds ratio by an amount of β_m, with all other predictors fixed. We used standardised data to ensure that the weights measure the relative contributions of their predictors to the classification, whereas the intercept expresses the base case for average predictor values. Our strategy was to explore commonly reported diagnostics of GLOFs as candidate predictors (Fig. 2, Table 1). We further acknowledged that data on moraine-dammed lakes in the HKKHN are structured, reflecting, for example, the variance in topography and synoptic regime such as the summer monsoon in the eastern HKKHN and westerlies in the western HKKHN. Different data sources, collection methods, and resolutions also add structure. This structure is routinely acknowledged, often raised as a caveat, but rarely treated, in GLOF studies. Ignoring such structure can lead to incorrect inference by bloating the statistical significance of irrelevant or inappropriate parameter estimates (Austin et al., 2003). To explicitly address this issue, we chose a multi-level logistic regression as a compromise between a single pooled model and individual models for each group in the data (Fig. 3; Gelman and Hill, 2007; Shor et al., 2007). s

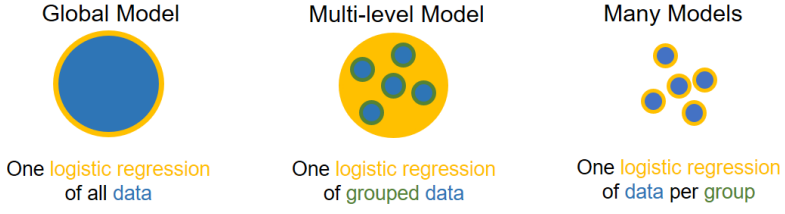


Figure 3: Schematic comparison of global vs. multi-level logistic regression models.

We recast Eq. (2) using a group index j:

$$\mu_i = \text{S}(\alpha_j + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \cdots + \beta_p x_{i,p}) \quad (4)$$

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

$$\alpha_j \sim \mathrm{N}(\mu_\alpha, \sigma_\alpha), \quad (5)$$

where μ_α is the mean, and σ_α is the standard deviation, of the group-level intercepts α_j that are learned from all data and inform each other via the model hierarchy. We used a Bayesian framework (Kruschke and Liddell, 2018) by combining the likelihood of observing the data with prior knowledge from previous GLOF studies (Fischer et al., 2020). We used the statistical programming language **R** with the package `brms`, which estimates joint posterior distributions using a Hamiltonian Monte Carlo algorithm and a No-U-Turn Sampler (NUTS) (Bürkner, 2017). We ran four chains of 1500 samples after 500 warm-up runs each, and checked for numerical divergences or other pathological issues. We only considered models with all values of $\hat{R}$ <1.01, a measure of numerical convergence of sampling chains, to avoid unbiased posterior distributions (Nalborczyk et al., 2019).

Unless stated otherwise, we used a weakly informative half Student-t distribution with three degrees of freedom and a scale parameter of 10 for the standard deviations of group-level effects (Table 2; Bürkner, 2017; Gelman, 2006). At the population level, we chose weakly informative priors for the intercept and coefficients for which we had no other prior knowledge. We encoded this lack of knowledge with a prior Cauchy distribution centred at zero and with scale 2.5, following the recommendation by Gelman et al. (2008). Rapidly growing moraine-dammed lakes are a widely used diagnostic of high GLOF potential (Aggarwal et al., 2016; Allen et al., 2019; Bolch et al., 2011; Ives et al., 2010; Mergili and Schneider, 2011; Prakash and Nagarajan, 2017). We encoded this notion in a prior Gaussian distribution with one unit mean and standard deviation, hence shifting more probability mass towards positive regression weights without excluding the possibility of negative weight estimates (Table 2).

Table 2: Prior distributions for group- and population-level effects.

Level	Model coefficient	Prior PDF
Group-level effects	Standard deviation σ of group model variables	$\sigma_\alpha \sim HalfStudentT(3,10)$
Population-level effects	Intercept	$\alpha_j \sim Cauchy(0,2.5)$
	Weight of predictors with weak prior knowledge	$\beta_p \sim Cauchy(0,2.5)$
	Weight of predictor lake area β_A	$\beta_A \sim Normal(1,2)$

We estimated the predictive performance of all models with leave-one-out (LOO) cross-validation as part of the `brms` package (Bürkner, 2017). LOO values like the expected log predictive density (ELPD) summarise the predictive error of Bayesian models, similar to the Akaike Information Criterion or related metrics of model selection (Vehtari et al., 2017). They are based on the log-likelihood of the posterior simulations of parameter values (Vehtari et al., 2017).

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020


The Cryosphere Discussions
Open Access
EGU

3 Results

Elevation-dependent warming model

Our first model addresses the notion of elevation-dependent warming (EDW) by considering lake elevation as a grouping structure in the data. The model further assumes that the GLOF history of a given lake is a function of its area A and net change ΔA. This dependence differs up to a constant, i.e. the varying model intercept, across elevation bands z that we define here in five quantile grouping levels (Fig. 1). The model intercept may vary across these elevation bands, whereas lake area (in 2005) and its net change remain fixed predictors. In essence, this varying-intercept model acknowledges that glacial lakes in the same elevation band might have a common susceptibility to GLOFs in the past four decades. The indicator variable ΔA records whether a given lake had a net growth or shrinkage between 1990 and 2018:

$$\mu_i = \mathrm{S}(\alpha_z + \beta_A A_i + \beta_{\Delta A} \Delta A_i) \quad (6)$$

$$\alpha_z \sim \mathrm{N}(\mu_z, \sigma_z), \quad (7)$$

where index z identifies the elevation band.

We obtain posterior estimates of $\beta_A = 0.79^{+0.27}/_{-0.27}$ and $\beta_{\Delta A} = 0.48^{+0.73}/_{-0.72}$ (95% highest density interval, HDI) that indicate that larger lakes are more likely classified as having had a GLOF, whereas net growth or shrinkage has ambivalent weight as its HDI includes zero (Fig. 4, Fig. 5, Table 3). On the population level, the low spread of intercepts ($\sigma_z = 0.29^{+0.68}/_{-0.28}$) estimated for each of the five elevation bands shows that elevation effects modulate the pooled model only minutely. These posterior effects are positive for the lower elevation bands, but negative for the higher elevation bands. Thus, the mean posterior probability of a GLOF history, P_{GLOF}, under this model increases slightly for lakes in lower elevations and with larger surface area in 2005. We also observe that P_{GLOF} <0.5 regardless of reported lake elevation, and that the associated uncertainties are highest for largest lakes.

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

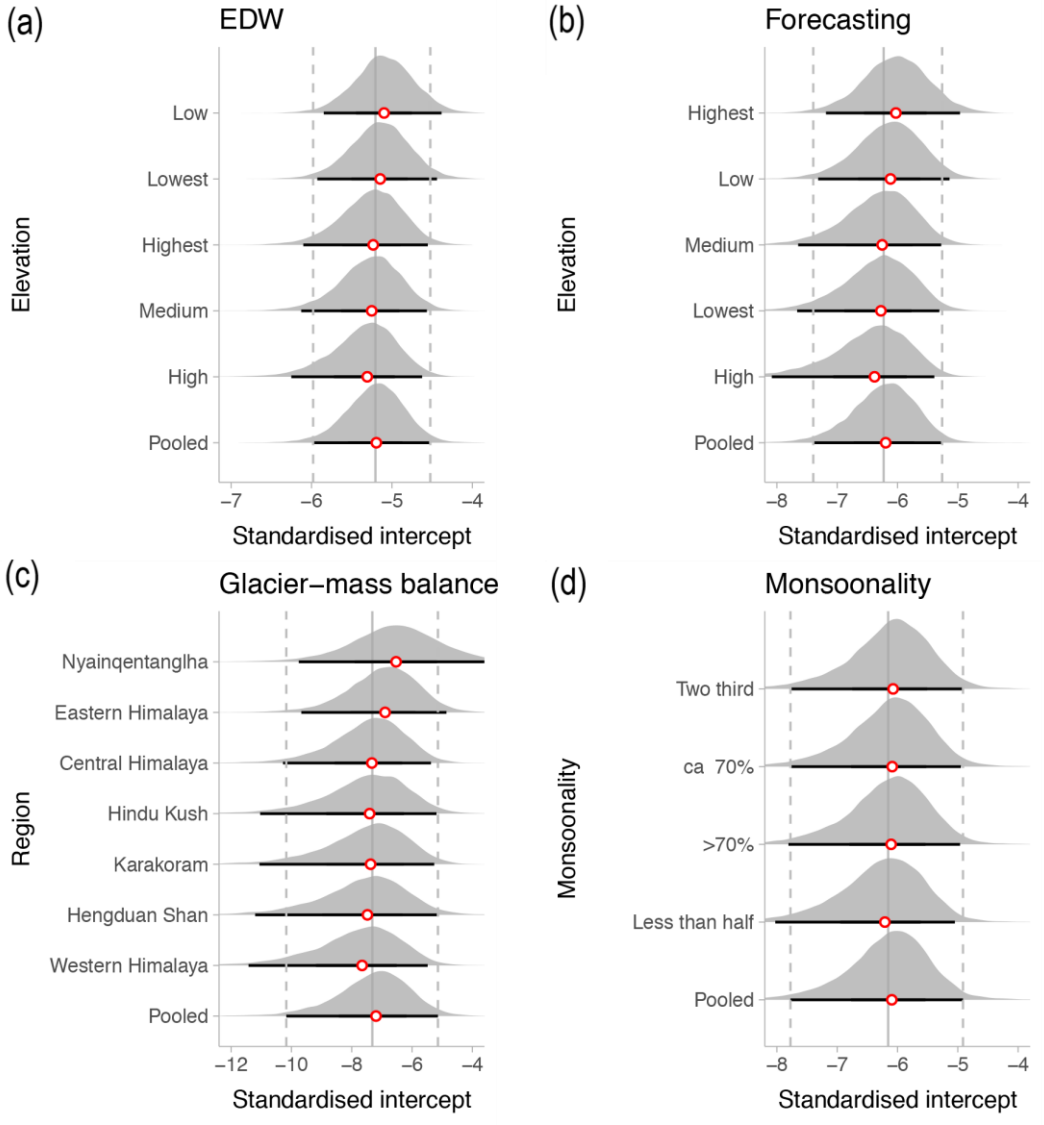


Figure 4: Posterior pooled and group-level intercepts for the four models considered; EDW = elevation-dependent warming; see Fig. 1 for a summary of the quantiles of elevation and monsoonality. Black horizontal lines delimit 95% HDI, red circles indicate posterior medians. Vertical continuous (dashed) grey lines are posterior means (95% HDI) of the pooled intercept of each model. Intercepts are standardised and thus refer to lakes with average predictor values.

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

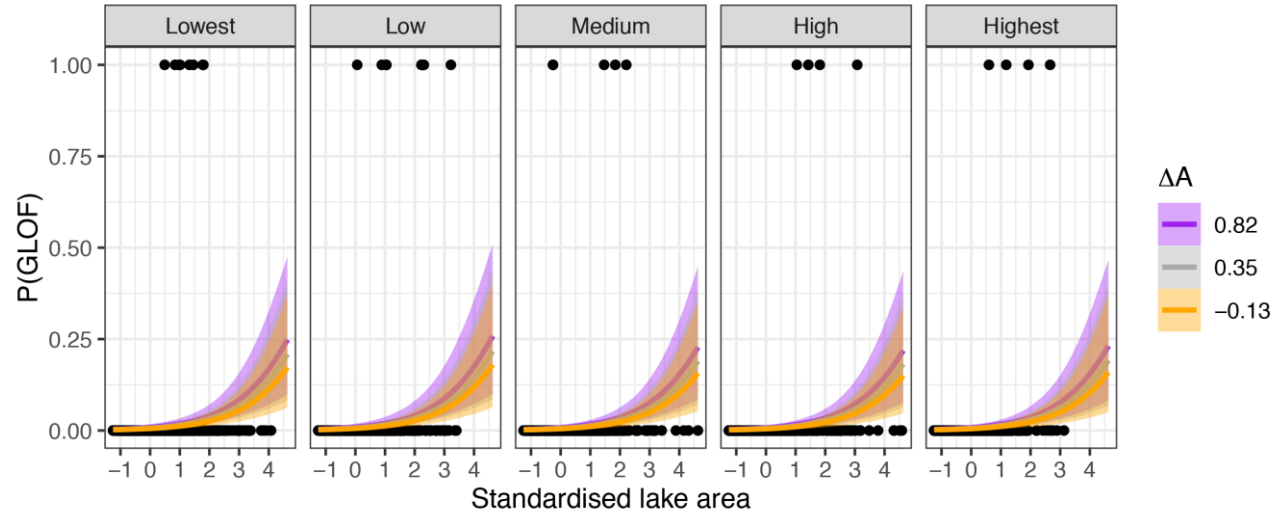


Figure 5: Elevation-dependent warming model: posterior probabilities P_{GLOF} as a function of standardised lake area (in 2005) and the sign of standardised lake-area change ΔA (i.e. net growth or shrinkage), grouped by quantiles of elevation (defined in Fig. 1). Black dots are lake data with (no) reported GLOF records. Thick coloured lines are mean fits, and colour shades encompass the associated 95% HDIs.

Forecasting model

Our second model refines our approach by including only relative changes in lake area before the reported GLOFs happened. We can use this model to fore- or hindcast historic GLOFs in our inventory. Here we use lake area A (in 2005) and its relative change A^* from 1990 to 2005 as predictors of eleven GLOFs that occurred between 2005 and 2018 across the five elevation bands. We assume that larger and deeper lakes are more robust to relative size changes and thus also include a multiplicative interaction term between lake area and its change:

$$\mu_i = \mathrm{S}(\alpha_z + \beta_A A_i + \beta_{A^*} A_i^* + \beta_{A\times A^*} A_i \times A_i^*) \quad (8)$$

We find that lake area has a credible positive posterior weight of $\beta_A = 0.86^{+0.44}/_{-0.43}$, hence greater lakes are more likely to having had a GLOF between 2005 and 2018. The weight of relative lake-area change in the 15 years before is ambiguous ($\beta_{*A} = -0.04^{+0.76}/_{-0.67}$) and so is the interaction ($\beta_{A\times A^*} = -0.16^{+0.41}/_{-0.51}$). On average, however, relative increases in lake area between 1990 and 2005 slightly decrease P_{GLOF}. Unlike in the elevation-dependent warming model, the effects of elevation bands are less clear, while the uncertainties are more pronounced and highest for larger and shrinking lakes (Fig. 4, Fig. 6).

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

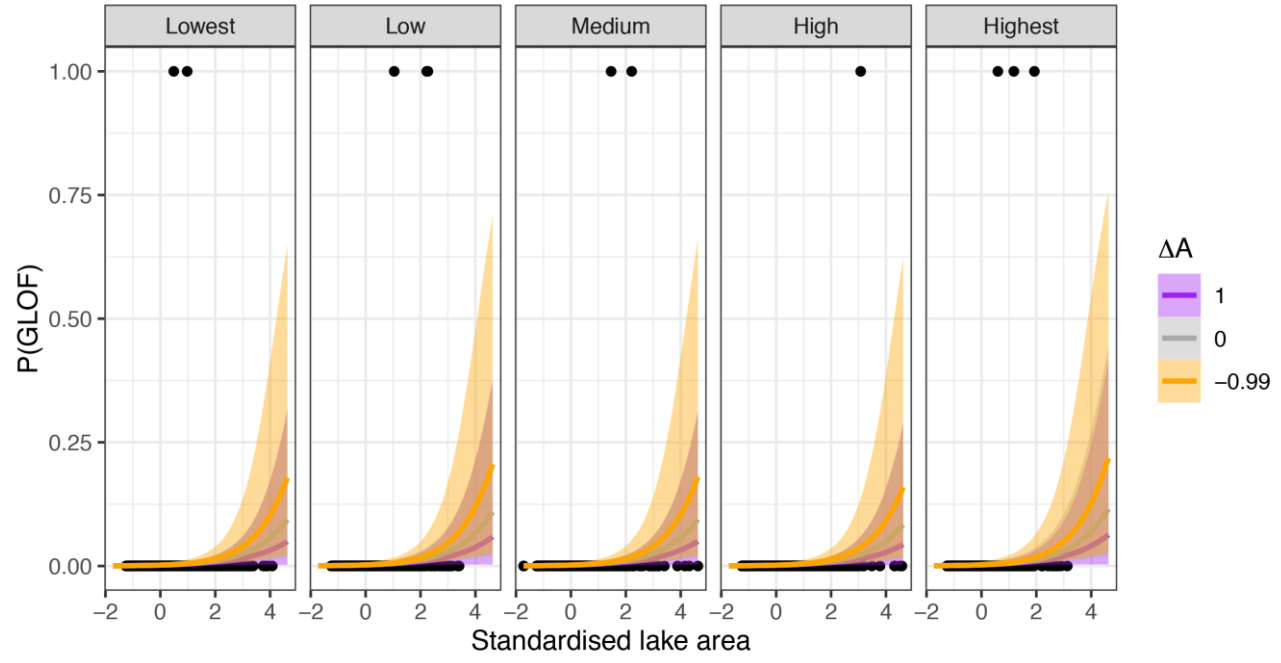


Figure 6: Forecasting model: posterior probabilities P_{GLOF} as a function of standardised lake area (in 2005) and standardised lake-area change ΔA between 1990 and 2005, grouped by quantiles of elevation (defined in Fig. 1). Black dots are lake data with (no) reported GLOF records for the interval 2005 to 2018. Thick coloured lines are mean fits, and colour shades encompass the associated 95% HDIs.

Glacier-mass balance model

Besides elevation, our third model considers the average historic glacier-mass balances across the HKKHN. The model assumes that mean ice losses Δm add a distinctly regional structure to the susceptibility to GLOFs in the past four decades, given that accelerated glacier melt may raise GLOF potential (Emmer, 2017; Richardson and Reynolds, 2000). We use the seven RGI regions as defined by Brun et al. (2017) as group-levels r and their average glacier-mass balance as a group-level predictor Δm_r. Our pooled predictors are the relative change of lake area A^* from 2005 to 2018 (to ensure a comparable time interval) and the catchment area C upstream of each lake. We replace lake area by its upstream catchment area, which is less prone to change, but well correlated to lake area.

$$\mu_i = \mathrm{S}(\alpha_z + \alpha_r + \beta_{A^*} A_i^* + \beta_C C_i), \quad (9)$$

$$\alpha_r \sim \mathrm{N}(\mu_r + \gamma_r \Delta m_r, \sigma_r). \quad (10)$$

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

This model returns a positive weight for catchment area ($\beta_C = 0.85^{+0.50}/_{-0.50}$) and a negative weight for relative lake-area changes ($\beta_{A^*} = -0.69^{+0.64}/_{-0.61}$), whereas the effect of the mean glacier-mass balance remains inconclusive ($\gamma_r = -2.98^{+4.87}/_{-6.70}$). On the basis of higher standard deviations, we learn that effects of glaciological regions vary more than those of elevation bands ($\sigma_r = 0.81^{+1.60}/_{-0.78}$ and $\sigma_z = 0.48^{+1.19}/_{-0.47}$). This is also reflected in the posterior distributions across the glacier-mass balance regions (Fig. 4) as well as the calculated group-level effects. This model has the highest values of P_{GLOF} for average lakes in the Nyainqentanglha Mountains and the Eastern Himalaya (Fig. 4). In contrast to the forecasting model, we observe that increases in lake area now credibly depress P_{GLOF} (Fig. 7).

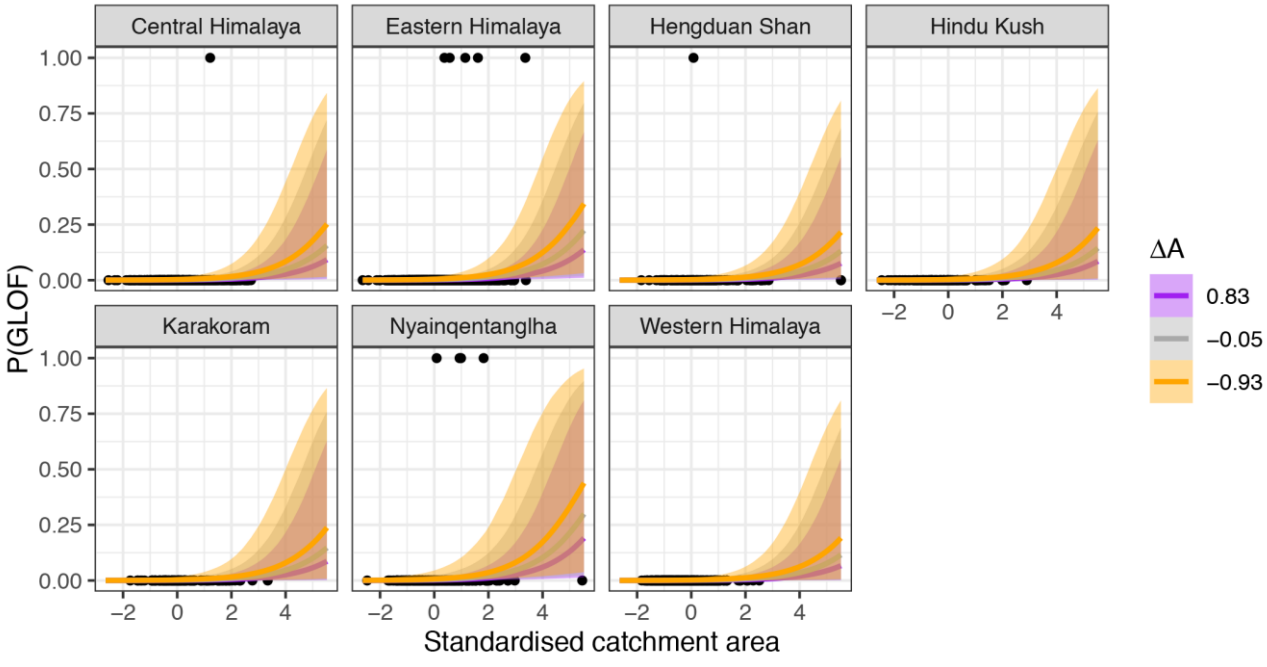


Figure 7: Glacier-mass balance model: posterior probabilities P_{GLOF} as a function of standardised catchment area and standardised lake-area change ΔA between 2005 and 2018, grouped by regions of average glacier-mass balance (see Fig. 1). Black dots are lake data with (no) reported GLOF records for the interval 2005 to 2018. Thick coloured lines are mean fits, and colour shades encompass the associated 95% HDIs.

Monsoonality model

Our last model explores a synoptic influence on GLOF susceptibility by grouping the data by the summer proportion of mean annual precipitation and thus by approximate monsoonal contribution. We defined five monsoonality levels based on quantiles of the annual proportions of summer precipitation (Fig. 1). We use relative lake-area change A^* between 1990 and 2018, and catchment area C as population-level predictors, as well as the additional grouping by regional glacier-mass balance:

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

The Cryosphere Discussions
Open Access
EGU

$$\mu_i = \mathrm{S}(\alpha_M + \alpha_r + \beta_{A^*} A_i^* + \beta_C C_i), \quad (11)$$

$$\alpha_M \sim \mathrm{N}(\mu_M, \sigma_M), \quad (12)$$

where index M identifies the monsoonality group. We find that larger catchment areas ($\beta_C = 0.82^{+0.46}/_{-0.48}$) and lakes with relative shrinkage ($\beta_{A^*} = -0.63^{+0.59}/_{-0.59}$) credibly raise P_{GLOF} (Fig. 4, Fig. 8). Higher standard deviations show that regional effects vary more for the mean glacial-mass balance than for monsoonality ($\sigma_r = 0.79^{+1.59}/_{-0.76}$ and $\sigma_M = 0.40^{+1.04}/_{-0.39}$), although both hardly change the pooled model trend.

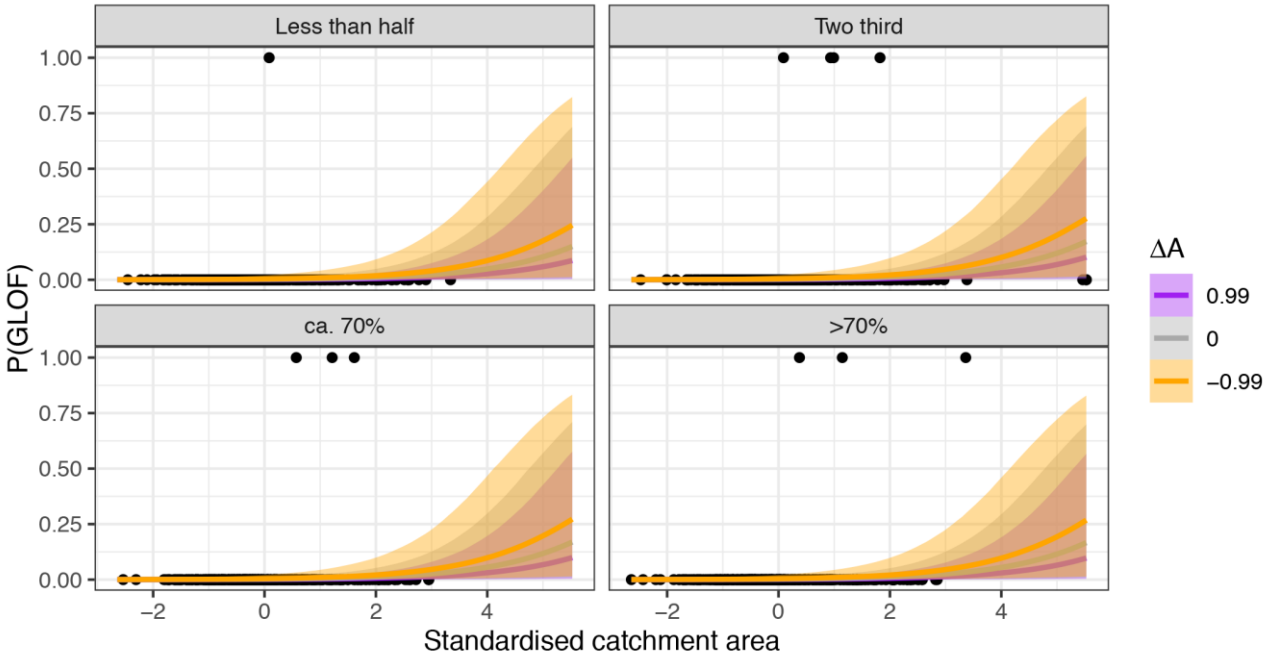


Figure 8: Monsoonality model: posterior probabilities P_{GLOF} as a function of standardised catchment area and standardised lake-area change ΔA between 1990 and 2018, grouped by quantiles of the annual proportion of precipitation falling during summer (defined in Fig. 1). Black dots are lake data with (no) reported GLOF records for the interval 1990 to 2018. Thick coloured lines are mean fits, and colour shades encompass the associated 95% HDIs.

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

Table 3: Summary of the results of our four models.

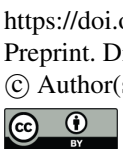
Model	Model parameter	Estimate	Estimation error	Lower 95% CI boundary	Upper 95% CI boundary
Elevation-dependent warming model	α_z	-5.22	0.36	-5.96	-4.56
	β_A	0.79	0.14	0.52	1.06
	$\beta_{\Delta A\ (1990\ to\ 2018)}$	0.49	0.38	-0.28	1.24
	σ_z	0.28	0.27	0.01	0.99
Forecasting model	α_z	-6.23	0.54	-7.39	-5.26
	β_A	0.87	0.22	0.44	1.31
	$\beta_{A^*\ (1990\ to\ 2005)}$	-0.04	0.38	-0.71	0.73
	β_{AxA^*}	-0.16	0.24	-0.67	0.26
	σ_z	0.43	0.41	0.01	1.49
Glacier-mass balance model	$\alpha_{z,r}$	-7.31	1.26	-10.15	-5.19
	$\beta_{A^*\ (2005\ to\ 2018)}$	-0.69	0.32	-1.31	-0.06
	β_C	0.85	0.26	0.35	1.36
	γ_r	-2.90	2.80	-9.27	1.80
	σ_z	0.47	0.44	0.01	1.61
	σ_r	0.83	0.66	0.03	2.47
Monsoonality model	$\alpha_{M,r}$	-6.14	0.70	-7.70	-4.91
	$\beta_{A^*\ (1990\ to\ 2018)}$	-0.63	0.31	-1.23	-0.02
	β_C	0.82	0.24	0.34	1.28
	σ_M	0.40	0.42	0.01	1.49
	σ_r	0.78	0.62	0.03	2.31

Model performance and validation

We estimate the performance of our models in terms of the posterior improvement of our prior chance of finding a lake with known outburst in the past four decades in our inventory by pure chance. We compare the posterior predictive mean P_{GLOF} with a mean prior probability that we estimate from the ~1% proportion of lakes with known GLOFs in our training data. We measure what we have learned from each model in terms of the log-odds ratio that readily translates into probabilities using Eq. (3). A positive (negative) log-odds ratio means that we obtain a higher (lower) posterior probability of attributing a historic GLOF to a given lake compared to a random draw. Based on this metric, all models have higher true positive than true negative rates. For a prior probability informed by the historic frequency of GLOFs, the models have at least about 80% true positives, and at least 70% true negatives on average (Fig. 9, Table 4).

The values of the LOO cross-validation of the predictive capabilities show that the EDW model formally has the least favourable, i.e. higher, values for both LOO metrics (Table 4). This is potentially due to the different true positives counts in

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020



The Cryosphere Discussions
Open Access
EGU

the training data sets. However, the range of estimated ELPD values between the remaining three models is small (ΔELPD = 1.9).

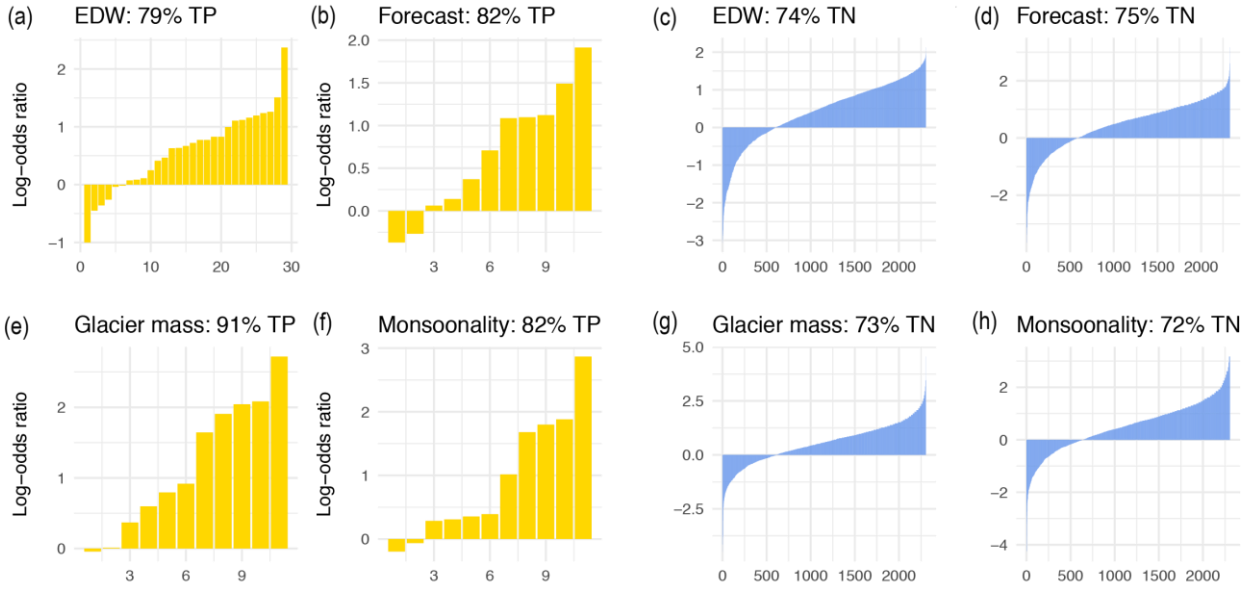


Figure 9: Average posterior log-odds ratios for true positives TP (negatives, TN), i.e. lakes with (without) a GLOF in the past four decades (on the x axis) for the four different models. The log-odds ratios describe here the ratio of the mean posterior over the mean prior probability of classifying a given lake as having had a GLOF. We estimate the mean prior probability from the relative frequency of GLOFs in the datasets; EDW = elevation-dependent warming model.

Table 4: Overview of model validation measures for the predictive capabilities of our models.

Model	Prior vs. posterior knowledge: X% true positives / X% true negatives correctly identified	ELPD	LOOIC
Elevation-dependent warming model	79% / 74%	-144.2	288.3
Forecasting model	82% / 75%	-66.5	132.9
Glacier-mass balance model	91% / 73%	-64.6	129.1
Monsoonality model	82% / 72%	-65.6	131.2

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020


The Cryosphere Discussions

4 Discussion

4.1 Topographic and climatic predictors of GLOFs

We used Bayesian multi-level logistic regression to test whether several widely advocated diagnostics of GLOFs are credible predictors of at least one outburst in the past four decades. All four models that we considered identify **lake area** and **catchment area** as predictors with weights that credibly differ from zero with 95% probability. Our model results quantitatively support qualitative notions of several basin-wide studies in the HKKHN (Bolch et al., 2011; Ives et al., 2010; Mergili and Schneider, 2011) and elsewhere (McKillop and Clague, 2007), which proposed that larger moraine-dammed lakes have a higher potential for releasing GLOFs.

We also found that **changes in lake area** have partly inconclusive influences in the models. Two exceptions are the negative weight of lake-area changes β_{A*} in the glacier-mass balance model and in the monsoonality model, regardless of the differing intervals that these changes were determined for (Table 3). While this result formally indicates that shrinking lakes are more likely to be classified as having had a historic GLOF, the period over which these lake-area changes are valid (2005 to 2018) overlaps with the timing of eleven recorded GLOFs (Eq. 9). In other words, the lake shrinkage could be a direct consequence of these GLOFs instead of vice versa. Nonetheless, our results indicate that lake-area changes, either absolute or directional, are somewhat inconclusive in informing us whether a given lake has a recent GLOF history. This result contradicts the assumptions made in many previous studies that assumed that rapidly growing lakes are the most prone to sudden outburst (Aggarwal et al., 2016; Bolch et al., 2011; Ives et al., 2010; Mergili and Schneider, 2011; Prakash and Nagarajan, 2017; Rounce et al., 2016; Wang et al., 2012). One advantage of the Bayesian approach, however, is that we can express the role of lake-area changes in GLOF susceptibility by choosing different highest density intervals. For example, if we adopted a narrower (80%) HDI for ΔA, we could be 80% certain that net lake-area growth increased P_{GLOF} under the elevation-dependent warming model (Eq. 6). In the forecasting model, however, the influence of lake-area change remains negligible even for <50% HDIs.

The role of **elevation** in GLOF predictions is also less pronounced than that of lake or catchment area, at least as a group level. The weights of the elevation-dependent warming model indicate that lower (higher) lakes are slightly more (less) likely to have had a historic GLOF (Fig. 4), but hardly warrant any better model performance compared to the pooled (or elevation-independent) model. In the forecasting model, however, the contributions of lake elevation to P_{GLOF} are devoid of any systematic pattern and likely reflect several, potentially combined, drivers (Fig. 4). This model was trained on fewer GLOFs and thus suffers from greater uncertainties in terms of the 95% HDIs. Clearly, the role of elevation may need more future investigation. In terms of elevation bands, it hardly seems to aid GLOF detection with the models used here. Similarly, Emmer et al. (2016) reported that lake elevation was hardly affecting GLOF hazard in the Cordillera Blanca, Peru.

Judging from the regionally averaged **glacier-mass balances**, our models predict the highest GLOF probabilities in the Nyainqentanglha Mountains and the Eastern Himalaya, which have had the highest historic GLOF counts (Fig. 1). The timing

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

and seasonality of snowfall affects how glaciers respond to rising air temperatures. Observed frequencies and predicted probabilities of historic GLOFs are lowest for several glaciers with positive mass balance in the Karakoram and Western Himalayas (Fig. 1, Fig. 10). Most moraine-dammed lakes in the HKKHN, however, are fed by glaciers with negative mass balances that likely help to elevate GLOF potential through increased meltwater input and glacier-tongue calving rates (Emmer, 2017; Richardson and Reynolds, 2000). More than 70% of all lakes that burst out in the past four decades were in contact to their parent glaciers (Veh et al., 2019). Given that the regional glacier-mass balance is linked to synoptic precipitation patterns (Kapnick et al., 2014; King et al., 2019; Krishnan et al., 2019), our glacier-mass balance model highlights that the regional ice loss outweighs the role of monsoonality in terms of higher changes to the group-level intercepts for comparable mean P_{GLOF} and associated uncertainties (Fig. 4, Fig. 7, Fig. 8).

Our results offer insights into the links between historic GLOFs and the **synoptic precipitation patterns**. Richardson and Reynolds (2000) presumed that seasonal floods and GLOFs are both caused by high monsoonal precipitation and summer ablation. In contrast, our results indicate that the fraction of summer precipitation changes the predictive probabilities of historic GLOFs only marginally, at least at the group level, so that deviations from a pooled model for the HKKHN are minute. In essence, our results underline the need for exploring more the interactions of both precipitation and temperature as potential GLOF triggers. It may well be that seasonal timing of heavy precipitation events and type (rain or snow) at a given lake may be more meaningful to GLOF susceptibility than annual totals or averages. Whether our finding that glacier-mass balances driven by superimposed synoptic regimes credibly influence regional GLOF susceptibility in the HKKHN is applicable to other regions, for example the Cordillera Blanca in the South American Andes (Emmer et al., 2016; Emmer and Vilímek, 2014; Iturrizaga, 2011), also needs further investigation.

4.2 Model Assessment

We consider our quantitative and data-driven approach as complementary to existing qualitative and basin-wide GLOF hazard appraisals. Our models cannot replace field observations that deliver local details on GLOF-disposing factors such as moraine or adjacent rock-slope stability, presence of ice cores, glacier calving rates, or surges. Our selection of predictors is a compromise between widely used diagnostics of GLOFs and their availability as data covering the entire HKKHN. To this end, we used lake (or catchment) area and lake-area changes as predictors, and elevation, regional glacier-mass balance, and monsoonality as group levels of past GLOF activity of several thousand moraine-dammed lakes in the HKKHN. Among the many possible combinations of predictors and group levels we focused on those few combinations with minimal correlation among the input variables. We minimised the potential for misclassification by using a purely remote-sensing-based inventory of GLOFs, which reduces reporting bias for GLOFs too small to be noticed or happening in unpopulated areas: more destructive GLOFs are recorded more often than smaller GLOFs in remote areas (Veh et al., 2018, 2019). We are thus confident that we trained our models on lakes with a confirmed GLOF history at the expense of discarding known outbursts predating

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

the onset of Landsat satellite coverage in 1981. We acknowledge that climate products such as precipitation can have large biases because of orographic effects or climate circulation patterns and interpolation using topography (Karger et al., 2017; Mukul et al., 2017). Cross-validation of CHELSA precipitation estimates with station data has a global mean coefficient of determination R^2 of 0.77, with regional variations between 0.53 and 0.90 (Karger et al., 2017). By accounting for orographic wind effects, CHELSA products outperform previous global datasets such as the WorldClim (Hijmans et al., 2005), especially in the rugged HKKHN topography. We stress that we therefore used all climatic data as aggregated group-level variables to avoid spurious model results. At the level of individual lakes, we thus resorted only to size, elevation, and upstream catchment area as more robust predictors.

Due to strong imbalance in our training data, we opted for prior vs. posterior log-odd comparison instead of commonly applied Receiver Operating Characteristics (ROC) in estimating the predictive capabilities of our models (Saito and Rehmsmeier, 2015). In our models, only few posterior estimates of P_{GLOF} are >0.5 and they, thus, offer very conservative estimates of a GLOF history (Fig. 10). All models have wide 95% HDIs that attest a high level of uncertainty. This observation may be sobering, but nevertheless documents objectively the minimum amount of accuracy that these simple models afford for objectively detecting historic outbursts.

The low fraction of lakes with a GLOF history (~1%) curtails a traditional logistic regression model and favours instead a Bayesian multi-level approach that can handle imbalanced training data and collinear predictors (Gelman and Hill, 2007; Hille Ris Lambers et al., 2006; Shor et al., 2007). We prefer the straight-forward interpretation of posterior regression weights to random forest classifiers, neural networks or support vector machines (Caniani et al., 2008; Falah et al., 2019; Kalantar et al., 2018; Taalab et al., 2018). While these methods may perform better, they disclose little about the relationship between model inputs and outputs (Blöthe et al., 2019; Dinov, 2018); much of their higher accuracy is also linked to the overwhelming number of true negatives. Yet so far, multi-criteria decision analysis or decision-making trees have been the method of choice in GLOF hazard assessments, both in High Mountain Asia (Bolch et al., 2011; Prakash and Nagarajan, 2017; Rounce et al., 2016; Wang et al., 2012) and elsewhere (Emmer et al., 2016; Emmer and Vilímek, 2014; Huggel et al., 2002; Kougkoulos et al., 2018). While these methods strongly rely on expert judgement (Allen et al., 2019), a Bayesian logistic regression encodes any prior knowledge or constraints explicitly and reproducibly as probability distributions. Still, inconsiderate or inappropriate prior choices can introduce bias (Van Dongen, 2006; Kruschke and Liddell, 2018). Therefore, we carefully considered our choice of weakly informative priors for predictors with limited prior knowledge, following the guidelines concerning regression models by Gelman (2006) and Gelman et al. (2008). We also cross-checked our results when applying varying prior choices and found negligible differences in the resulting posterior distributions.

To summarise, our simple classification models hardly support the notion that elevation or changes in lake area are straightforward predictors of a GLOF history, at least for the moraine-dammed lakes that we studied in the HKKHN. Lake size and regional differences in glacier-mass balance are items that future studies of GLOF susceptibility may wish to consider

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

further. The performance of these models is moderate to good if compared to a random classification, yet associated with high uncertainties in terms of wide highest density intervals. We underline that these uncertainties have rarely been addressed, let alone quantified, in previous work. One way forward may be to create ensembles of such models to improve their predictive capability instead of relying on any single model.

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

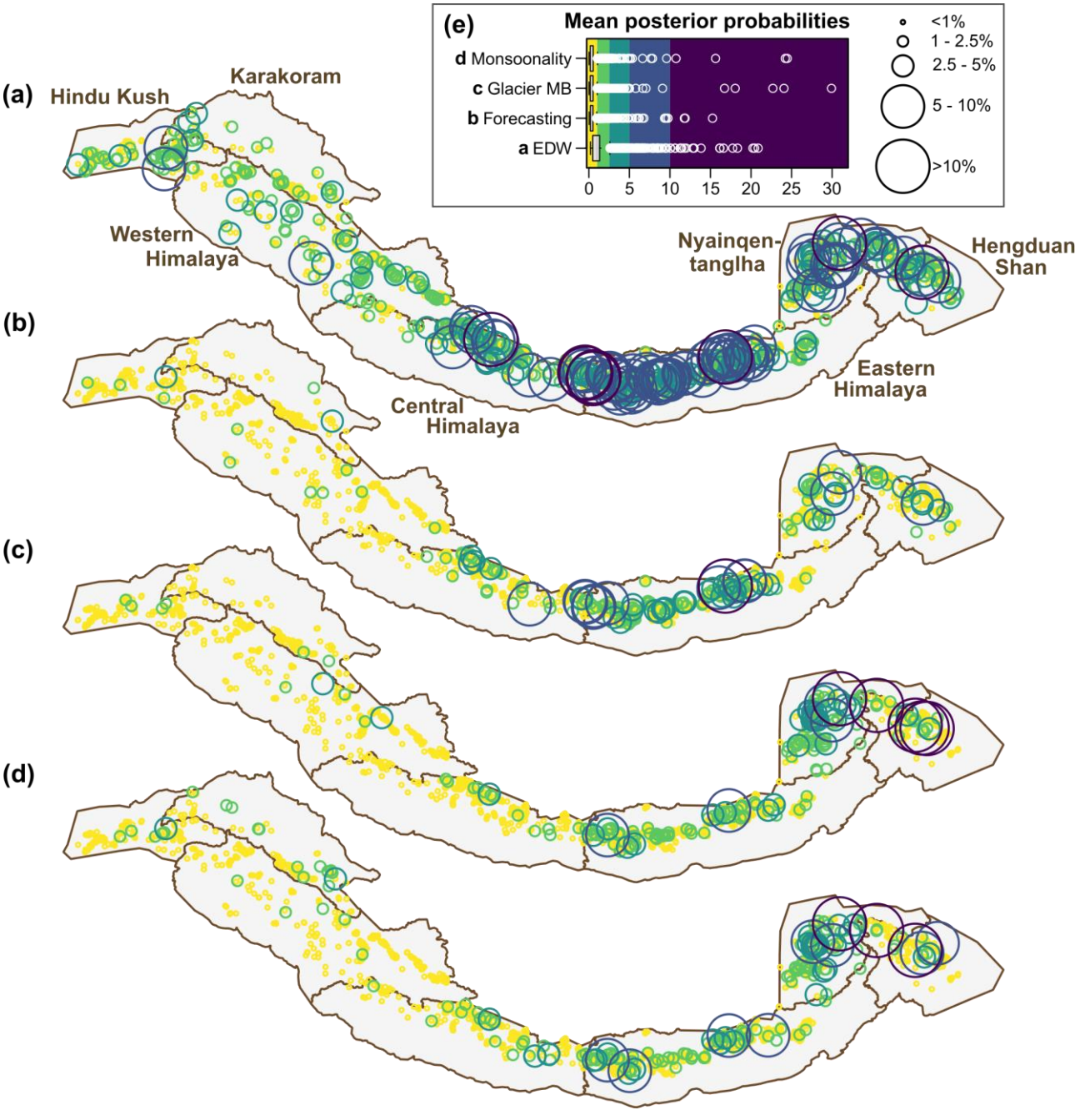


Figure 10: Mean posterior probabilities of HKKHN glacial lakes for having a GLOF history (P_{GLOF}) in the past four decades as estimated in the (a) elevation-dependent warming model, (b) forecasting model, (c) glacier-mass balance model, and (d) monsoonality model. Size and colours of bubbles are scaled by posterior probabilities (e).

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

5 Conclusions

We quantitatively investigated the susceptibility of moraine-dammed lakes to GLOFs in major mountain regions of High Asia. We used a systematically compiled and comprehensive inventory of moraine-dammed lakes with documented GLOFs in the past four decades to test how elevation, lake area and its rate of change, glacier-mass balance, and monsoonality perform as predictors and group levels in a Bayesian multi-level logistic regression. Our results show that larger lakes in larger catchments have been more prone to sudden outburst floods, as have those lakes in regions with pronounced negative glacier-mass balance. While elevation-dependent warming (EDW) may control a number of processes conducive to GLOFs, grouping our classification by elevation bands adds little to a pooled model for the entire HKKHN. Historic changes in lake area, both in absolute and relative values, have an ambiguous role in these models. We observed that shrinking lakes favour the classification as GLOF-prone, although this may arise from overlapping measurement intervals such that the reduction in lake size arises from outburst rather than vice versa. In any case, the widely adapted notion that (rapid) lake growth may be a diagnostic of impending outburst remains poorly supported by our model results. Our Bayesian approach allows explicit probabilistic prognoses of the role of these widely cited controls on GLOF susceptibility, but also attests to previously hardly quantified uncertainties, especially for the larger lakes in our study area. While individual models offer some improvement with respect to a random classification based on average GLOF frequency, we recommend considering ensemble models for obtaining more accurate and flexible predictions of outbursts from moraine-dammed lakes.

Data and code availability

This study is based on freely available data. Shuttle Radar Topography Mission (SRTM) data are available from the US Geological Survey (https://www.earthexplorer.usgs.gov). We derived climatic variables from the CHELSA Bioclim data set (https://chelsa-climate.org/bioclim/) described by Karger et al. (2017) and regional glacier-mass balances from Brun et al. (2017). We extracted glacial lake information from inventories published by Maharjan et al. (2018), Veh et al. (2019), and Wang et al. (2020). We processed our data with free **R** statistical software (https://cran.r-project.org/), including the `brms` package by Bürkner (2017) (https://CRAN.R-project.org/package=brms). The model code to this article by Fischer et al. (2020) is published in a GitHub repository and available online at: https://doi.org/10.5281/zenodo.4161577.

Author contributions

This study was conceptualised by all authors. While formal analysis and methodology were conducted by MF and OK, data curation was mainly carried out by GV. Visualisations of data and results, including maps, were prepared by GV, OK and MF. MF prepared the original manuscript; OK, GV, and AW reviewed and edited the writing.

460 **Competing interests**

The authors declare that they have no conflict of interest.

Acknowledgements

This research was funded by the Deutsche Forschungsgemeinschaft (DFG) via the graduate research training group NatRiskChange (GRK 2043/1) at the University of Potsdam (https://www.natriskchange.de).

465 **References**

Aggarwal, A., Jain, S. K., Lohani, A. K. and Jain, N.: Glacial lake outburst flood risk assessment using combined approaches of remote sensing, GIS and dam break modelling, Geomatics, Nat. Hazards Risk, 7(1), 18–36, doi:10.1080/19475705.2013.862573, 2016.

Allen, S. K., Rastner, P., Arora, M., Huggel, C. and Stoffel, M.: Lake outburst and debris flow disaster at Kedarnath, June
470 2013: hydrometeorological triggering and topographic predisposition, Landslides, 13(6), 1479–1491, doi:10.1007/s10346-015-0584-3, 2016.

Allen, S. K., Zhang, G., Wang, W., Yao, T. and Bolch, T.: Potentially dangerous glacial lakes across the Tibetan Plateau revealed using a large-scale automated assessment approach, Sci. Bull., (April), doi:10.1016/j.scib.2019.03.011, 2019.

Austin, P. C., Tu, J. V. and Alter, D. A.: Comparing hierarchical modeling with traditional logistic regression analysis among
475 patients hospitalized with acute myocardial infarction: Should we be analyzing cardiovascular outcomes data differently?, Am. Heart J., 145(1), 27–35, doi:10.1067/mhj.2003.23, 2003.

Bajracharya, S. R. and Shrestha, B.: The Status of Glaciers in the Hindu Kush–Himalayan Region, Kathmandu. [online] Available from: http://dx.doi.org/10.1659/mrd.mm113, 2011.

Blöthe, J. H., Rosenwinkel, S., Höser, T. and Korup, O.: Rock-glacier dams in High Asia, Earth Surf. Process. Landforms,
480 44(3), 808–824, doi:10.1002/esp.4532, 2019.

Bolch, T., Peters, J., Yegorov, A., Pradhan, B., Buchroithner, M. and Blagoveshchensky, V.: Identification of potentially dangerous glacial lakes in the northern Tien Shan, Nat. Hazards, 59(3), 1691–1714, doi:10.1007/s11069-011-9860-2, 2011.

Bookhagen, B. and Burbank, D. W.: Toward a complete Himalayan hydrological budget: Spatiotemporal distribution of snowmelt and rainfall and their impact on river discharge, J. Geophys. Res. Earth Surf., 115(3), 1–25,
485 doi:10.1029/2009JF001426, 2010.

Brun, F., Berthier, E., Wagnon, P., Kääb, A. and Treichler, D.: A spatially resolved estimate of High Mountain Asia glacier mass balances from 2000 to 2016, Nat. Geosci., 10(9), 668–673, doi:10.1038/ngeo2999, 2017.

Bürkner, P.-C.: brms: An R package for Bayesian multilevel models using Stan, J. Stat. Softw., 80(1), 1–28, 2017.

Caniani, D., Pascale, S., Sdao, F. and Sole, A.: Neural networks and landslide susceptibility: A case study of the urban area of Potenza, Nat. Hazards, 45(1), 55–72, doi:10.1007/s11069-007-9169-3, 2008.

Carrivick, J. L. and Tweed, F. S.: A global assessment of the societal impacts of glacier outburst floods, Glob. Planet. Change, 144, 1–16, doi:10.1016/j.gloplacha.2016.07.001, 2016.

Cenderelli, D. A. and Wohl, E. E.: Flow hydraulics and geomorphic effects of glacial-lake outburst floods in the Mount Everest region, Nepal, Earth Surf. Process. Landforms, 28(4), 385–407, doi:10.1002/esp.448, 2003.

Costa, J. E. and Schuster, R. L.: The Formation and Failure of Natural Dams, Vancouver., 1987.

Dehecq, A., Gourmelen, N., Gardner, A. S., Brun, F., Goldberg, D., Nienow, P. W., Berthier, E., Vincent, C., Wagnon, P. and Trouvé, E.: Twenty-first century glacier slowdown driven by mass loss in High Mountain Asia, Nat. Geosci., 12(1), 22–27, doi:10.1038/s41561-018-0271-9, 2019.

Dinov, I. D.: Data science and predictive analytics: Biomedical and health applications using R, Springer., 2018.

Van Dongen, S.: Prior specification in Bayesian statistics: Three cautionary tales, J. Theor. Biol., 242(1), 90–100, doi:https://doi.org/10.1016/j.jtbi.2006.02.002, 2006.

Emmer, A.: Glacier Retreat and Glacial Lake Outburst Floods (GLOFs), in Oxford Research Encyclopedia of Natural Hazard Science, pp. 1–37, Oxford University Press USA., 2017.

Emmer, A. and Vilímek, V.: Review article: Lake and breach hazard assessment for moraine-dammed lakes: An example from the Cordillera Blanca, Nat. Hazards Earth Syst. Sci., 13(6), 1551–1565, doi:10.5194/nhess-13-1551-2013, 2013.

Emmer, A. and Vilímek, V.: New method for assessing the susceptibility of glacial lakes to outburst floods in the Cordillera Blanca, Peru, Hydrol. Earth Syst. Sci., 18(9), 3461–3479, doi:10.5194/hess-18-3461-2014, 2014.

Emmer, A., Klimeš, J., Mergili, M., Vilímek, V. and Cochachin, A.: 882 lakes of the Cordillera Blanca: An inventory, classification, evolution and assessment of susceptibility to outburst floods, Catena, 147, 269–279, doi:10.1016/j.catena.2016.07.032, 2016.

Etzelmüller, B. and Frauenfelder, R.: Factors controlling the distribution of mountain permafrost in the northern hemisphere and their influence on sediment transfer, Arctic, Antarct. Alp. Res., 41(1), 48–58, doi:10.1657/1523-0430-41.1.48, 2009.

Evans, S. G. and Clague, J. J.: Recent climatic change and catastrophic geomorphic processes in mountain environments, Geomorphology, 10(1–4), 107–128, doi:10.1016/0169-555X(94)90011-6, 1994.

Falah, F., Rahmati, O., Rostami, M., Ahmadisharaf, E., Daliakopoulos, I. N. and Pourghasemi, H. R.: 14 - Artificial Neural Networks for Flood Susceptibility Mapping in Data-Scarce Urban Areas, in Spatial Modeling in GIS and R for Earth and Environmental Sciences, edited by H. R. Pourghasemi and C. B. T.-S. M. in G. I. S. and R. for E. and E. S. Gokceoglu, pp. 323–336, Elsevier., 2019.

Fischer, M., Korup, O., Veh, G. and Walz, A.: GLOFsusceptibility: First release of the GLOF susceptibility model (Version

v.1.0). Zenodo. doi:10.5281/ZENODO.4161577, 2020.

Gelman, A.: Prior distributions for variance parameters in hierarchical models, Bayesian Anal., 1(3), 515–533, doi:10.1002/cjs.5550340302, 2006.

Gelman, A. and Hill, J.: Data Analysis using Regression and Multilevel/Hierarchical Models, Cambridge University Press, New York., 2007.

Gelman, A., Jakulin, A., Pittau, M. G. and Su, Y. S.: A weakly informative default prior distribution for logistic and other regression models, Ann. Appl. Stat., 2(4), 1360–1383, doi:10.1214/08-AOAS191, 2008.

Haeberli, W., Schaub, Y. and Huggel, C.: Increasing risks related to landslides from degrading permafrost into new lakes in de-glaciating mountain ranges, Geomorphology, 293, 405–417, doi:https://doi.org/10.1016/j.geomorph.2016.02.009, 2017.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. and Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas, Int. J. Climatol., 25(15), 1965–1978, doi:10.1002/joc.1276, 2005.

Hille Ris Lambers, J., Aukema, B., Diez, J., Evans, M. and Latimer, A.: Effects of global change on inflorescence production : a Bayesian hierarchical analysis, in Hierarchical Modelling for the Environmental Sciences - Statistical Methods and Applications, edited by J. S. Clark and A. E. Gelfand, pp. 59–76, Oxford University Press North Carolina, Cary., 2006.

Hock, R., Rasul, G., Adler, C., Cáceres, B., Gruber, S., Hirabayashi, Y., Jackson, M., Kääb, A., Kang, S., Kutuzov, S., Milner, A., Molau, U., Morin, S., Orlove, B. and Steltzer, H. I.: Chapter 2: High Mountain Areas, IPCC Spec. Rep. Ocean Cryosph. a Chang. Clim., 131–202, 2019.

Huggel, C., Kääb, A., Haeberli, W., Teysseire, P. and Paul, F.: Remote sensing based assessment of hazards from glacier lake outbursts: a case study in the Swiss Alps, Can. Geotech. J., 39(2), 316–330, doi:10.1139/t01-099, 2002.

Huggel, C., Haeberli, W., Kääb, A., Bieri, D. and Richardson, S.: An assessment procedure for glacial hazards in the Swiss Alps, Can. Geotech. J., 41(6), 1068–1083, doi:10.1139/T04-053, 2004.

Iturrizaga, L.: Glacier Lake Outburst Floods, in Encyclopedia of Snow, Ice and Glaciers, edited by V. P. Singh, P. Singh, and U. K. Haritashya, pp. 381–399, Springer Netherlands, Dodrecht., 2011.

Ives, J. D., Shrestha, R. B. and Mool, P. K.: Formation of Glacial Lakes in the Hindu Kush-Himalayas and GLOF Risk Assessment, International Centre for Integrated Mountain Development (ICIMOD), Kathmandu., 2010.

Kalantar, B., Pradhan, B., Naghibi, S. A., Motevalli, A. and Mansor, S.: Assessment of the effects of training data selection on the landslide susceptibility mapping : a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN), Geomatics, Nat. Hazards Risk, 9(1), 49–69, doi:10.1080/19475705.2017.1407368, 2018.

Kapnick, S. B., Delworth, T. L., Ashfaq, M., Malyshev, S. and Milly, P. C. D.: Snowfall less sensitive to warming in Karakoram than in Himalayas due to a unique seasonal cycle, Nat. Geosci., 7(11), 834–840, doi:10.1038/ngeo2269, 2014.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P. and Kessler, M.: Climatologies at high resolution for the earth's land surface areas, Sci. Data, 4, 1–20, doi:10.1038/sdata.2017.122,

2017.

King, O., Bhattacharya, A., Bhambri, R. and Bolch, T.: Glacial lakes exacerbate Himalayan glacier mass loss, Sci. Rep., 9(1), 1–9, doi:10.1038/s41598-019-53733-x, 2019.

Koike, T. and Takenaka, S.: Scenario Analysis on Risks of Glacial Lake Outburst Floods on the Mangde Chhu River, Bhutan, Glob. Environ. Res., 16, 41–49, 2012.

Kougkoulos, I., Cook, S. J., Jomelli, V., Clarke, L., Symeonakis, E., Dortch, J. M., Edwards, L. A. and Merad, M.: Use of multi-criteria decision analysis to identify potentially dangerous glacial lakes, Sci. Total Environ., 621, 1453–1466, doi:10.1016/j.scitotenv.2017.10.083, 2018.

Kraaijenbrink, P. D. A., Bierkens, M. F. P., Lutz, A. F. and Immerzeel, W. W.: Impact of a global temperature rise of 1.5 degrees Celsius on Asia's glaciers, Nature, 549(7671), 257–260, doi:10.1038/nature23878, 2017.

Krishnan, R., Shrestha, A. B., Ren, G., Rajbhandari, R., Saeed, S., Sanjay, J., Syed, M. A., Vellore, R., Xu, Y., You, Q. and Ren, Y.: Unravelling Climate Change in the Hindu Kush Himalaya: Rapid Warming in the Mountains and Increasing Extremes, in The Hindu Kush Himalaya Assessment: Mountains, Climate Change, Sustainability and People, edited by P. Wester, A. Mishra, A. Mukherji, and A. B. Shrestha, pp. 57–97, Springer International Publishing, Cham., 2019.

Kruschke, J. K. and Liddell, T. M.: Bayesian data analysis for newcomers, Psychon. Bull. Rev., 25(1), 155–177, doi:10.3758/s13423-017-1272-1, 2018.

Liu, J. J., Cheng, Z. L. and Su, P. C.: The relationship between air temperature fluctuation and Glacial Lake Outburst Floods in Tibet, China, Quat. Int., 321, 78–87, doi:10.1016/j.quaint.2013.11.023, 2014.

Maharjan, S. B., Mool, P. K., Lizong, W., Xiao, G., Shrestha, F., Shrestha, R. B., Khanal, N. R., Bajracharya, S. R., Joshi, S., Shai, S. and Baral, P.: The Status of Glacial Lakes in the Hindu Kush Himalaya, International Centre for Integrated Mountain Development (ICIMOD), Kathmandu., 2018.

Maurer, J. M., Schaefer, J. M., Rupper, S. and Corley, A.: Acceleration of ice loss across the Himalayas over the past 40 years, Sci. Adv., 5(6), doi:10.1126/sciadv.aav7266, 2019.

McKillop, R. J. and Clague, J. J.: Statistical, remote sensing-based approach for estimating the probability of catastrophic drainage from moraine-dammed lakes in southwestern British Columbia, Glob. Planet. Change, 56(1–2), 153–171, doi:10.1016/j.gloplacha.2006.07.004, 2007.

Mergili, M. and Schneider, J. F.: Regional-scale analysis of lake outburst hazards in the southwestern Pamir, Tajikistan, based on remote sensing and GIS, Nat. Hazards Earth Syst. Sci., 11(5), 1447–1462, doi:10.5194/nhess-11-1447-2011, 2011.

Molden, D. J., Vaidya, R. A., Shrestha, A. B., Rasul, G. and Shrestha, M. S.: Water infrastructure for the Hindu Kush Himalayas, Int. J. Water Resour. Dev., 30(1), 60–77, doi:10.1080/07900627.2013.859044, 2014.

Mool, P. K., Maskey, P. R., Koirala, A., Joshi, S. P., Wu, L., Shrestha, A. B., Eriksson, M., Gurung, B., Pokharel, B., Khanal, N. R., Panthi, S., Adhikari, T., Kayastha, R. B., Ghimire, P., Thapa, R., Shrestha, B., Shrestha, S. and Shrestha, R. B.: Glacial

https://doi.org/10.5194/tc-2020-327
Preprint. Discussion started: 16 December 2020

Lakes and Glacial Lake Outburst Floods in Nepal, International Centre for Integrated Mountain Development (ICIMOD), Kathmandu., 2011.

Mukul, M., Srivastava, V., Jade, S. and Mukul, M.: Uncertainties in the Shuttle Radar Topography Mission (SRTM) Heights: Insights from the Indian Himalaya and Peninsula, Sci. Rep., 7(February), 1–10, doi:10.1038/srep41672, 2017.

Nalborczyk, L., Batailler, C., Loevenbruck, H., Vilain, A. and Bürkner, P. C.: An introduction to bayesian multilevel models using brms: A case study of gender effects on vowel variability in standard Indonesian, J. Speech, Lang. Hear. Res., 62(5), 1225–1242, doi:10.1044/2018_JSLHR-S-18-0006, 2019.

Nie, Y., Sheng, Y., Liu, Q., Liu, L., Liu, S., Zhang, Y. and Song, C.: A regional-scale assessment of Himalayan glacial lake changes using satellite observations from 1990 to 2015, Remote Sens. Environ., 189, 1–13, doi:10.1016/j.rse.2016.11.008, 2017.

Palazzi, E., Von Hardenberg, J. and Provenzale, A.: Precipitation in the Hindu-Kush Karakoram Himalaya: Observations and future scenarios, J. Geophys. Res. Atmos., 118(1), 85–100, doi:10.1029/2012JD018697, 2013.

Palazzi, E., Filippi, L. and von Hardenberg, J.: Insights into elevation-dependent warming in the Tibetan Plateau-Himalayas from CMIP5 model simulations, Clim. Dyn., 48(11–12), 3991–4008, doi:10.1007/s00382-016-3316-z, 2017.

Pepin, N., Bradley, R. S., Diaz, H. F., Baraer, M., Caceres, E. B., Forsythe, N., Fowler, H., Greenwood, G., Hashmi, M. Z., Liu, X. D., Miller, J. R., Ning, L., Ohmura, A., Palazzi, E., Rangwala, I., Schöner, W., Severskiy, I., Shahgedanova, M., Wang, M. B., Williamson, S. N. and Yang, D. Q.: Elevation-dependent warming in mountain regions of the world, Nat. Clim. Chang., 5(5), 424–430, doi:10.1038/nclimate2563, 2015.

Pfeffer, W. T., Arendt, A. A., Bliss, A., Bolch, T., Cogley, J. G., Gardner, A. S., Hagen, J. O., Hock, R., Kaser, G., Kienholz, C., Miles, E. S., Moholdt, G., Mölg, N., Paul, F., Radić, V., Rastner, P., Raup, B. H., Rich, J., Sharp, M. J., Andreassen, L. M., Bajracharya, S., Barrand, N. E., Beedle, M. J., Berthier, E., Bhambri, R., Brown, I., Burgess, D. O., Burgess, E. W., Cawkwell, F., Chinn, T., Copland, L., Cullen, N. J., Davies, B., De Angelis, H., Fountain, A. G., Frey, H., Giffen, B. A., Glasser, N. F., Gurney, S. D., Hagg, W., Hall, D. K., Haritashya, U. K., Hartmann, G., Herreid, S., Howat, I., Jiskoot, H., Khromova, T. E., Klein, A., Kohler, J., König, M., Kriegel, D., Kutuzov, S., Lavrentiev, I., Le Bris, R., Li, X., Manley, W. F., Mayer, C., Menounos, B., Mercer, A., Mool, P., Negrete, A., Nosenko, G., Nuth, C., Osmonov, A., Pettersson, R., Racoviteanu, A., Ranzi, R., Sarikaya, M. A., Schneider, C., Sigurdsson, O., Sirguey, P., Stokes, C. R., Wheate, R., Wolken, G. J., Wu, L. Z. and Wyatt, F. R.: The randolph glacier inventory: A globally complete inventory of glaciers, J. Glaciol., 60(221), 537–552, doi:10.3189/2014JoG13J176, 2014.

Prakash, C. and Nagarajan, R.: Outburst susceptibility assessment of moraine-dammed lakes in Western Himalaya using an analytic hierarchy process, Earth Surf. Process. Landforms, 42(14), 2306–2321, doi:10.1002/esp.4185, 2017.

Rangwala, I. and Miller, J. R.: Climate change in mountains: A review of elevation-dependent warming and its possible causes, Clim. Change, 114(3–4), 527–547, doi:10.1007/s10584-012-0419-3, 2012.

Richardson, S. D. and Reynolds, J. M.: An overview of glacial hazards in the Himalayas, in Quaternary International, vol. 65–66, pp. 31–47., 2000.

Rolland, C.: Spatial and seasonal variations of air temperature lapse rates in alpine regions, J. Clim., 16(7), 1032–1046, doi:10.1175/1520-0442(2003)016<1032:SASVOA>2.0.CO;2, 2003.

Rounce, D. R., McKinney, D. C., Lala, J. M., Byers, A. C. and Watson, C. S.: A new remote hazard and risk assessment framework for glacial lakes in the Nepal Himalaya, Hydrol. Earth Syst. Sci., 20(9), 3455–3475, doi:10.5194/hess-20-3455-2016, 2016.

Saito, T. and Rehmsmeier, M.: The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets, PLoS One, 10(3), e0118432 [online] Available from: https://doi.org/10.1371/journal.pone.0118432, 2015.

Shor, B., Bafumi, J., Keele, L. and Park, D.: A Bayesian multilevel modeling approach to time-series cross-sectional data, Polit. Anal., 15(2), 165–181, doi:10.1093/pan/mpm006, 2007.

Somos-Valenzuela, M. A., McKinney, D. C., Byers, A. C., Voss, K., Moss, J. and McKinney, J. C.: Ground Penetrating Radar Survey for Risk Reduction at Imja Lake, Nepal, Austin. [online] Available from: http://hdl.handle.net/2152/19751, 2012.

Somos-Valenzuela, M. A., McKinney, D. C., Byers, A. C., Rounce, D. R., Portocarrero, C. and Lamsal, D.: Assessing downstream flood impacts due to a potential GLOF from Imja Lake in Nepal, Hydrol. Earth Syst. Sci. Discuss., 11(11), 13019–13053, doi:10.5194/hessd-11-13019-2014, 2014.

Taalab, K., Cheng, T. and Zhang, Y.: Mapping landslide susceptibility and types using Random Forest, Big Earth Data, 2(2), 159–178, doi:10.1080/20964471.2018.1472392, 2018.

Terzago, S., von Hardenberg, J., Palazzi, E. and Provenzale, A.: Snowpack Changes in the Hindu Kush–Karakoram–Himalaya from CMIP5 Global Climate Models, J. Hydrometeorol., 15(6), 2293–2313, doi:10.1175/JHM-D-13-0196.1, 2014.

Tudoroiu, M., Eccel, E., Gioli, B., Gianelle, D., Schume, H., Genesio, L. and Miglietta, F.: Negative elevation-dependent warming trend in the Eastern Alps, Environ. Res. Lett., 11(4), doi:10.1088/1748-9326/11/4/044021, 2016.

Veh, G., Korup, O., Roessner, S. and Walz, A.: Detecting Himalayan glacial lake outburst floods from Landsat time series, Remote Sens. Environ., 207(February), 84–97, doi:10.1016/j.rse.2017.12.025, 2018.

Veh, G., Korup, O., Specht, S., Roessner, S. and Walz, A.: Unchanged frequency of moraine-dammed glacial lake outburst floods in the Himalaya, Nat. Clim. Chang., 2000, 1–5, doi:10.1038/s41558-019-0437-5, 2019.

Veh, G., Korup, O. and Walz, A.: Hazard from Himalayan glacier lake outburst floods, Proc. Natl. Acad. Sci. U. S. A., 117(2), 907–912, doi:10.1073/pnas.1914898117, 2020.

Vehtari, A., Gelman, A. and Gabry, J.: Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC, Stat. Comput., 27(5), 1413–1432, doi:10.1007/s11222-016-9696-4, 2017.

Wang, W., Yao, T., Gao, Y., Yang, X. and Kattel, D. B.: A First-order Method to Identify Potentially Dangerous Glacial Lakes

in a Region of the Southeastern Tibetan Plateau, Mt. Res. Dev., 31(2), 122–130, doi:10.1659/MRD-JOURNAL-D-10-00059.1, 2011.

Wang, X., Liu, S., Guo, W. and Xu, J.: Assessment and simulation of glacier lake outburst floods for Longbasaba and Pida lakes, China, Mt. Res. Dev., 28(3–4), 310–317, doi:10.1659/mrd.0894, 2008.

Wang, X., Liu, S., Ding, Y., Guo, W., Jiang, Z., Lin, J. and Han, Y.: An approach for estimating the breach probabilities of moraine-dammed lakes in the Chinese Himalayas using remote-sensing data, Nat. Hazards Earth Syst. Sci., 12(10), 3109–3122, doi:10.5194/nhess-12-3109-2012, 2012.

Wang, X., Guo, X., Yang, C., Liu, Q., Wei, J., Zhang, Y., Liu, S., Zhang, Y., Jiang, Z. and Tang, Z.: Glacial lake inventory of High Mountain Asia (1990–2018) derived from Landsat images, Earth Syst. Sci. Data Discuss., (January), 1–23, doi:10.5194/essd-2019-212, 2020.

Worni, R., Huggel, C. and 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, Sci. Total Environ., 468–469, S71–S84, doi:10.1016/j.scitotenv.2012.11.043, 2013.

Yang, S.-K. and Smith, G. L.: Further Study on Atmospheric Lapse Rate Regimes, J. Atmos. Sci., 42(9), 961–966 [online] Available from: http://repositorio.unan.edu.ni/2986/1/5624.pdf, 1985.