



Article

Cite this article: Haq MA, Azam MF, Vincent C (2021). Efficiency of artificial neural networks for glacier ice-thickness estimation: a case study in western Himalaya, India. *Journal of Glaciology* 67(264), 671–684. <https://doi.org/10.1017/jog.2021.19>

Received: 1 January 2020

Revised: 1 February 2021

Accepted: 2 February 2021

First published online: 25 March 2021

Key words:

Glacier modelling; glacier volume; ice thickness measurements; remote sensing; subglacial exploration geophysics

Author for correspondence:

Mohd Anul Haq, E-mail: m.anul@mu.edu.sa

Efficiency of artificial neural networks for glacier ice-thickness estimation: a case study in western Himalaya, India

Mohd Anul Haq¹ , Mohd Farooq Azam² and Christian Vincent³

¹Department of Computer Science, College of Computer and Information Sciences, Majmaah University, Al-Majma'ah, 11952, Saudi Arabia; ²Discipline of Civil Engineering, Indian Institute of Technology Indore, Simrol 453552, India and ³University of Grenoble Alpes, CNRS, IRD, IGE, F-38000 Grenoble, France

Abstract

Knowledge of glacier volume is crucial for ice flow modelling and predicting the impacts of climate change on glaciers. Rugged terrain, harsh weather conditions and logistic costs limit field-based ice thickness observations in the Himalaya. Remote-sensing applications, together with mathematical models, provide alternative techniques for glacier ice thickness and volume estimation. The objective of the present research is to assess the application of artificial neural network (ANN) modelling coupled with remote-sensing techniques to estimate ice thickness on individual glaciers with direct field measurements. We have developed two ANN models and estimated the ice thickness of Chhota Shigri Glacier (western Himalaya) on ten transverse cross sections and two longitudinal sections. The ANN model estimates agree well with ice thickness measurements from a ground-penetrating radar, available for five transverse cross sections on Chhota Shigri Glacier. The overall root mean square errors of the two ANN models are 24 and 13 m and the mean bias errors are ± 13 and ± 6 m, respectively, which are significantly lower than for other available models. The estimated mean ice thickness and volume for Chhota Shigri Glacier are 109 ± 17 m and 1.69 ± 0.26 km³, respectively.

1. Introduction

Himalayan glaciers have been showing wastage over the last few decades (Brun and others, 2017; Azam and others, 2018; Bolch and others, 2019); consequently, their dynamics are gradually adjusting (Azam and others, 2012; Dehecq and others, 2019) to the current mass distribution and hypsometry. Reliable information on glacier ice thickness and volume is necessary to understand the status of these glaciers, determine future fresh water availability in the region and assess future sea level rise potential. Field-based ground penetrating radar (GPR) surveys (Hubbard and Glasser, 2005), one of the best available methods to estimate the ice thickness, have been used on glaciers worldwide (Pettersson and others, 2003; Pattyn and others, 2009; Wagnon and others, 2013). Due to rough terrain, harsh climatic conditions and high expedition costs in the Himalaya, GPR surveys have been conducted for ice-thickness estimates only for a few selected glaciers: Khumbu (Iwata and others, 1980), Dokriani (Gergan and others, 1999), Chhota Shigri (Azam and others, 2012; Singh and others, 2012), Mera (Wagnon and others, 2013), Changri Nup (Vincent and others, 2016), Lirung (McCarthy and others, 2017) and Satopanth (Mishra and others, 2018) Glaciers.

1.1 Ice thickness and volume estimation

Glacier ice thickness and volume are generally estimated using mathematical models such as volume–area scaling (Meier and Bahr, 1996; Raper and Braithwaite, 2005; Frey and others, 2014), ice flux method (Haeberli and Hoelzle, 1995; Farinotti and others, 2009a, 2009b; Huss and Farinotti, 2012), surface velocity–slope relationship (Gantayat and others, 2014), inverse modelling of glacier ice flow (Raymond-Pralong and Gudmundsson, 2011; Mosbeux and others, 2016) and artificial neural networks (ANNs) (Clarke and others, 2009; Haq and others, 2014).

Due to its simplicity, volume–area scaling has been used extensively to estimate glacier ice thickness and volume (Singh, 2011; Frey and others, 2014; Bahr and others, 2015). Volume–area scaling is essentially a transposition because unknown parameters such as glacier depth are derived from glacier area (Lliboutry, 1987; Gantayat and others, 2014). This method has generally been applied at the regional scale (Radić and Hock, 2011; Bliss and others, 2013; Frey and others, 2014) and can be erroneous when applied to individual glaciers for which the scaling parameter have not been well-established (Agrawal and Tayal, 2013; Bahr and others, 2015). Another shortcoming of the volume–area scaling method is that it yields no useful information on subglacial topography, which is a necessary boundary condition for glacier dynamics models (Clarke and others, 2009). Paul and Linsbauer (2012) derived the ice thickness of glaciers based on a numerical model of perfect-plasticity assumption developed by Nye (1952). In this approach, the glacier volume estimate is based on the surface extent and average slope of the glacier. These models provide an estimate of ice thickness but are prone to errors if

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

cambridge.org/jog

their underlying assumptions are not fulfilled (Clarke and others, 2009). Farinotti and others (2017) applied 17 models to derive the ice thickness of 21 test glaciers around the world.

Ramsankaran and others (2018) applied the Glabtop-2 model with optimization of the shape factor to derive the volume of Chhota Shigri Glacier. The majority of methods used in Farinotti and others (2017), as well as GlabTop-2 in Ramsankaran and others (2018), are highly sensitive to the surface slope and shape factor (Rabatel and others, 2018). Rabatel and others (2018) discussed the uncertainties caused by the surface slope and demonstrated that they can be as high as 50% of the thickness calculated by ice flux methods. Vashishth and others (2017) applied different methods including glacier flow mechanics, thickness–area scaling and volume–area scaling to compute the ice thickness of Chhota Shigri Glacier. Singh and others (2012) used the thickness-area scaling method and GPR data, whereas Dobhal and others (1995) used residual Bouguer values to measure the thickness of Chhota Shigri Glacier. The availability of these studies on Chhota Shigri Glacier provides an opportunity to compare our results with a variety of models.

1.2 Artificial neural network

An ANN is a mathematical or computational model that is inspired by the operation of biological neural networks (Abraham, 2005; Eluyode and Akomolafe, 2013). ANNs have already been applied in cryospheric science to understand past and future glacier length variations (Steiner and others, 2005; Zumbühl and others, 2008), to predict snow cover in mountain ranges (Sauter and others, 2009, 2010; Mishra and others, 2014), to simulate meltwater runoff from glaciers (Caiping and Yongjian, 2009), to classify glacier/snow cover surfaces (Bishop and others, 1999; Doberva and Klein, 2009; Czychowska-Wisniewski and others, 2015), to detect glacier and ice-shelf margins (Baumhoer and others, 2019; Mohajerani and others, 2019) and to simulate glacier surface mass balances (Bolibar and others, 2020). An ANN model was used by Clarke and others (2009) to estimate the sub-glacial topography and glacier volume of the Mount Waddington icefield, British Columbia, Canada. They validated the ANN-estimated ice thickness with ice thickness values obtained from a numerical ice dynamics model. In the Himalaya, the use of ANNs to estimate glacier ice thickness and volume estimation is rare. Haq and others (2014) applied an ANN model to derive ice thickness and volume of Gangotri Glacier in the central Himalaya. However, the estimated ice thickness on Gangotri Glacier could not be validated due to the unavailability of ice thickness data from the field.

1.3 ANN principles

A neural network is a collection of interconnected simple processing elements called neurons. These processing elements are connected with coefficients or weights, which constitute the neural network structure and are organized in layers. The behaviour of a neural network is determined by its architecture, transfer functions and learning rule (Agatonovic-Kustrin and Beresford, 2000). Every connection of a neural network is assigned a weight that comes through training the ANN. Typically this weight represents the strength of the interconnections among neurons inside the neural network. The aim of weight initialization is to prevent layer activation outputs from exploding or vanishing during the course of a forward pass through a neural network. The weighted inputs are all summed up inside the artificial neuron. The neurons use transfer functions f such as *transig*, *logsig* and *purelin* to derive output from the sum of weighted functions and biases. The function *logsig* generates outputs between 0 and 1 as the

neuron's net input ranges from negative to positive infinity (Fortuna and others, 2002). Alternatively, multilayer networks may use the tan-sigmoid transfer function, *tansig*. Occasionally, the linear transfer function *purelin* is used in backpropagation networks (Demuth and Beale, 2000). If the final layer of a multilayer network has neurons and sigmoid functions, then the outputs of the network will have a small range.

The training of ANN architecture can be performed using a defined set of rules, also known as optimizers, such as the Widrow–Hoff or gradient descent or backpropagation. The Widrow–Hoff learning rule minimizes the mean square error and moves the decision boundaries as far as it can from the training patterns (Demuth and Beale, 2000). The term 'backpropagation' refers to the way that the gradient is computed for nonlinear multilayer networks (Demuth and Beale, 2000). It is a generalization of the Widrow–Hoff learning rule to multiple-layer networks and nonlinear differentiable transfer functions. There are different types of backpropagation training algorithms such as Levenberg–Marquardt, conjugate gradient and resilient backpropagation (De Villiers and Barnard, 1993). The Levenberg–Marquardt training algorithm works on loss functions that take the form of a sum of squared errors. The loss function is composed of an error and regularization. The error evaluates how a neural network fits the dataset and the regularization is used to prevent overfitting based on the effective complexity of the neural network. The Levenberg–Marquardt algorithm computes the gradient vector and the Jacobian matrix. It is an algorithm that trains a neural network 10–100 times faster than the usual gradient descent backpropagation method (Demuth and Beale, 2000). The reason for choosing Levenberg–Marquardt optimization in the current study is that it is more powerful and faster than other conventional gradient descent techniques (Hagan and Menhaj, 1994; Kişi and Uncuoğlu, 2005).

The neural network architecture defines the network structure, including the number of hidden layers, the number of hidden nodes and the number of output nodes. The hidden layers provide the network with its ability to generalize. A network with one hidden layer can model any continuous function (Beale and Jackson, 1990). There are no a priori fixed criteria on the architecture of the ANN (Sheela and Deepa, 2013). A trial-and-error procedure and network growing technique are generally applied to decide the number of hidden layers and number of neurons in each hidden layer.

1.4 Objectives

In the present work, we assessed the efficiency of ANN modelling for the estimation of ice thickness and volume of Chhota Shigri Glacier (western Himalaya, India), where GPR ice-thickness measurements are also available (Azam and others, 2012). A major advantage of the proposed ANN model in this study is the limited input data requirements, which include only the glacier extent and a digital elevation model (DEM). As a first step, we developed an ANN model for ice thickness estimation on Chhota Shigri Glacier and then validated the output with field-derived ice thickness data from five GPR transverse cross sections. In the second step, a subset of field-derived observations of ice thickness was used as training data. The initial and trained models were then applied on five additional transverse cross sections and two longitudinal sections before deriving the final glacier volume.

2. Study area and data

Chhota Shigri Glacier (32.28° N, 77.58° E) is located in the Chandra Valley of the western Himalaya (Fig. 1). This glacier has a total length of 9 km with an area of 15.5 km² (Azam and

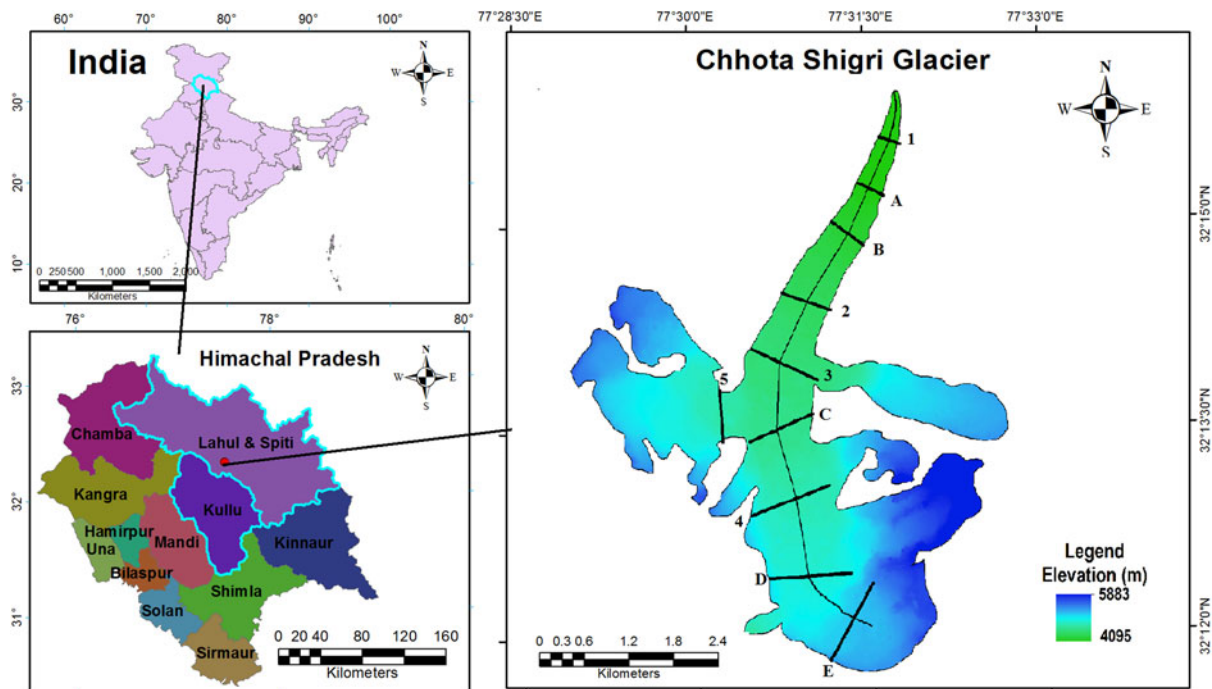


Fig. 1. Study area shown using a Cartosat-1 DEM, developed using a Cartosat-1 stereo pair (3 April 2007). Transverse cross sections 1–5 are those where GPR data and GCPs (Azam and others, 2012) are available. Transverse cross sections A–E are additional cross sections.

others, 2016). Chhota Shigri is one of the most documented glaciers in the Himalayan region with studies covering its mass balance, energy balance, ice thickness and dynamics (Wagon and others, 2007; Azam and others, 2012, 2014a, 2014b, 2019; Singh and others, 2012; Vincent and others, 2013; Frey and others, 2014; Vashishth and others, 2017; Ramsankaran and others, 2018; Mandal and others, 2020).

2.1 Digital elevation model

High-resolution topographic surface information is a prerequisite for accurate glacier volume estimates therefore resolution and quality of the DEM are crucial factors. Freely available DEMs such as Advanced Spaceborne Thermal Emission and Reflection Radiometer and the Shuttle Radar Topography Mission (SRTM) have a spatial resolution of 30 m. To satisfy the requirement of high-resolution topographic information in the present study, Cartosat-1 (2.5 m) stereopair imagery dated 3 April 2007 was used to generate the DEM at a spatial resolution and vertical accuracy of 10 m.

2.2 GPR observations

In October 2009, GPR measurements were conducted on Chhota Shigri Glacier to determine the ice thicknesses on five transverse cross sections between 4400 and 4900 m a.s.l. (cross sections 1–5 in Fig. 1) (Azam and others, 2012). The GPR thickness points were available at 10 m steps for each cross section. Details of the GPR measurements can be found in Azam and others (2012).

2.3 Differential global positioning system (DGPS) observations

Ground control points (GCPs) used in DEM generation and testing were collected in 2009 and 2010 at Chhota Shigri Glacier. Details of the 2009 survey, where 200 GCPs were collected during the GPR survey, are available in Azam and others (2012). Four additional GCPs were collected in 2010 using a Trimble R-7

DGPS operated in post-processing kinematics mode. The estimated accuracy of the 2009 and 2010 GCPs is <1 m.

2.4 Transverse and longitudinal slopes and elevations

Extending the GPR transects onto surrounding ice-free glacier slopes, we extracted sidewall slopes (SC) and elevations (EC) from the DEM. Longitudinal elevations (EL) and slopes (SL) were also extracted from the DEM (Fig. 2).

3. Methodology

3.1 DEM generation

A DEM was generated from a Cartosat-1 stereo pair collected on 3 April 2007. The DEM was generated using the Leica Photogrammetry Suite (LPS) and a subset of GCPs collected in 2009 and 2010 was used to refine the rational polynomial coefficient model. Tie points were generated both implicitly and explicitly for a better distribution of points over the glacier, and a first-order polynomial triangulation was used to refine the model.

The Automatic Terrain Extraction with Dense Point Matching (eATE) tool was used to place the subset of GCPs and generate tie points on the stereo pair. The GCPs were used as full control points (a full control point has x , y and z coordinates in LPS), while the tie points were generated using an implicit image correlation algorithm in LPS. A tie point is the image coordinate position of an object appearing on two or more images and its x , y and z coordinates are calculated during the aerial triangulation process. A total of 78 tie points were generated, manually and automatically. The triangulation was performed after adding GCPs and tie points. The DEM was extracted with a cell size of 10×10 m.

A total of 160 DGPS points (October 2009) on five GPR cross sections (1–5) and four DGPS points (October 2010) were used to generate the DEM. A total of 40 points (eight points per profile) were reserved and used to validate the error of the DEM. While the Cartosat-1 DEM was generated from April 2007 imagery, the GCPs were collected in October 2009 and October 2010.

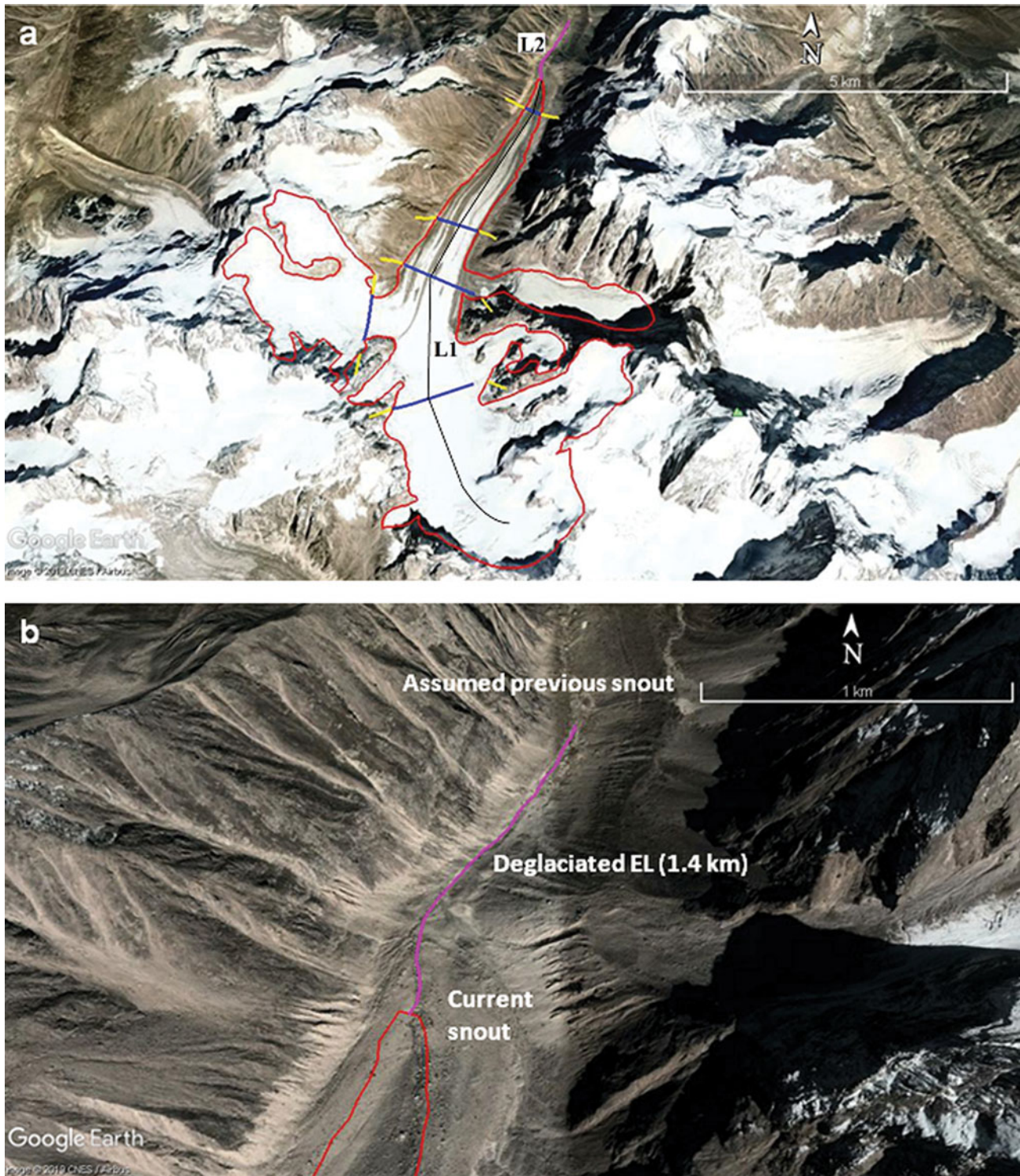


Fig. 2. (a) Outline of Chhota Shigri Glacier shown in red on Google Earth image. Blue lines indicate the five transverse sections that were manually digitized on GPR cross sections (1–5). Yellow lines indicate 300 m of de-glaciated valley sidewalls on either end of the glacier transects. The purple line indicates the extent between the current snout position and the assumed previous snout position (~1400 m from the current snout). (b) The extent between the current snout position and the assumed previous snout position is shown at a higher zoom level by zooming in on the Google Earth image shown in (a).

The resulting glacier DEM was edited based on break lines and sidewall slopes in LPS. The altitude values of the generated Cartosat DEM (3 April 2007) showed lower values than the global positioning system points (2009) for cross sections 1–4, however, cross section 5 showed a slightly higher altitude (Table S1).

3.2 ANN training

Two additional sources of information were used in this work, compared to previous ANN studies (Clarke and others, 2009;

Haq and others, 2014). The first is the additional geomorphic assumption of inclusion of sidewall slopes in the ANN model and the second is the availability of GPR data to assess and validate the ANN results. According to Clarke and others (2009), the areas that are currently ice-free were formerly ice-covered and are therefore subject to landscape erosion. Two geomorphic assumptions were used in the current investigation. First, there may be some relationship between the elevations of the longitudinal section of the de-glaciated frontal area and the glacier of the present time (Haq and others, 2014). Slopes and elevations along the

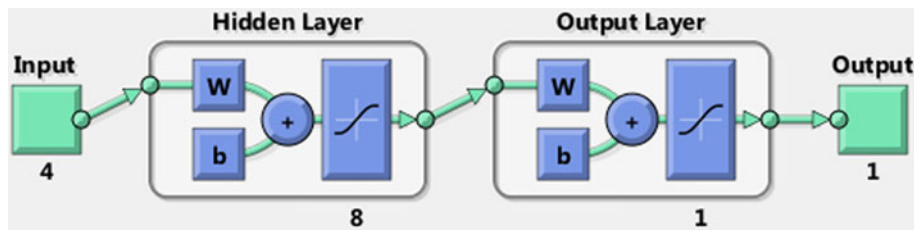


Fig. 3. ANN network architecture was adopted in this study for the execution of the ANN models. Here, w and b indicate weight and bias, respectively.

longitudinal profiles (SL and EL) were thus used as predictors in the ANN models. The second assumption is based on the extension of valley sidewall slopes beneath the current glacier surface. Valley slopes (SC) and elevations (EC) extended from transects 1–5 were also used as predictors in the ANN models. Clarke and others (2009) used bedrock geometries from deglaciated valleys to train their ANN model and improve dynamically modelled ice thickness.

3.3 ANN architecture

The performance of a neural network depends to a significant extent on how well it has been trained. Initially, a single layer linear network with the Widrow–Hoff least-square learning rule (also known as delta or least mean squared rule) was attempted. The training and testing showed unstable results on vector data (columns of input data). A multilayer feed-forward approach and standard Levenberg–Marquardt backpropagation training algorithm (Levenberg, 1944; Marquardt, 1963) were used in Matlab 2016b.

The Levenberg–Marquardt method is an effective optimization algorithm and is preferred over steepest descent and Gauss–Newton methods in a wide variety of problems (Wilson and Mantooth, 2013; Gavin, 2019). The Levenberg–Marquardt is a combination of the two optimization methods: the gradient descent method and the Gauss–Newton method. A major limitation of the Levenberg–Marquardt algorithm is its memory requirement. However, in the present study, the number of data points is low, therefore memory requirements were low.

In the ANN architecture, the design of the input and output layers is straightforward, however, the design of hidden layers is not. There is no straightforward way to determine good network architecture and the number of hidden layers. It varies with the number of training samples and the complexity of the problem. In the current study, several combinations of layers and transfer functions were tested on a network growing basis with a learning rate of 0.001. It typically starts with a small network and nodes are added until a suitably chosen measure stops decreasing (Gallant, 1986; Kwok and Yeung, 1995; Haq and others, 2014).

In this study, we use a three-layer architecture, in which the input layer had four neurons (4x), the first active layer had eight neurons with a tan-sigmoid (*transig*) transfer function (8T) and the last layer had a *transig* function (1T). The final selected ANN architecture notation was therefore 4x–8T–1T (Fig. 3). The assessment of different ANN architectures was based on a combined R value for training, validation and testing. This performance criterion was used to select the network architecture and the size of the training set.

3.4 ANN models for Chhota Shigri

Two ANN models were developed in the current study to estimate the ice thickness of Chhota Shigri Glacier. The first ANN model, which does not use observed ice depths in training, is abbreviated as ABT (ANN before training). The second ANN model includes observed ice depths in training and is abbreviated as AAT (ANN

after training). The ABT model was developed using four inputs such as slope and elevations extracted from the DEM on valley sidewalls and longitudinal profiles (EC, SC, EL, and SL). The target or dependent values were extrapolated using equation (S1); please see the Supplementary material.

3.5 ANN assessment

ANN models were assessed using a k -fold cross-validation technique, multiple R values, and by comparing mean and maximum ice thicknesses. The evaluation of the developed ANN model was attempted based on k -fold cross-validation. This k -fold cross-validation technique was applied to split the dataset for better generalization and to avoid the overfitting problem and provide a stable estimate of the model error (Little and others, 2017). The first fold is kept for testing and the model is trained on $k - 1$ folds. The process is repeated 10 times and each time different fold or a different group of data points are used for validation. Data points for training, testing and validation sets were taken randomly from all the available transects in both models and these points change depending on each fold.

3.5.1 Assessment statistics

The comparisons of observed and ANN estimated ice thicknesses along transverse cross sections were based on various statistics including multiple R -value of training, testing and, validation; max ice thickness and mean ice thickness.

3.5.2 Ice thickness interpolation

The ANN models produced estimates of bedrock depths at ten transverse cross sections, which were then converted into ice thicknesses by differencing with the surface elevation. We used kriging to interpolate the spatially irregular ice-thickness data over the whole glacier to obtain gridded estimates of ice thickness. Kriging is a stochastic interpolation method that offers an assessment of prediction errors. At the 10 m resolution of the DEM, five transverse cross sections (1–5), five additional transverse cross sections (A–E), and a longitudinal section (L1) give a total number of 1449 points for the interpolation. A workflow diagram (Fig. 4) further explains the two ANN models, AAT and ABT, adopted in this study.

3.6 Error analysis

ANN modelled ice thickness data are continuous since they are based on the DEM. GPR-based ice depth observations are discontinuous, and data are often missing at the end of transects where the terrain is rough and GPR signals are weak due to steep valley walls (Azam and others, 2012). Comparisons of observed and modelled ice thicknesses were made with datasets where (a) only the observed values were considered and (b) missing ice thickness observations were interpolated from nearby observations. Errors in modelled ice thickness are based on root mean squared error (RMSE) and the mean bias error. There are four sources of error: (1) error in GPR thickness, (2) mean average bias of ABT and AAT, (3) error in kriging interpolation for

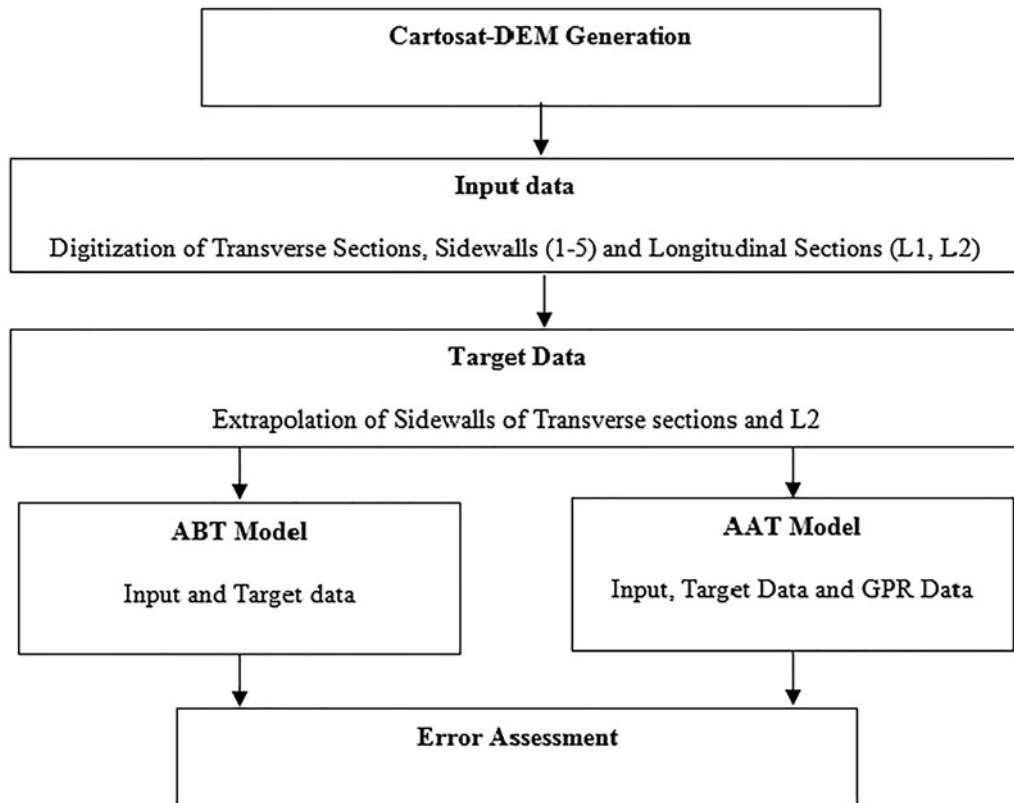


Fig. 4. A workflow diagram explaining the two ANN models, AAT and ABT, adopted in this study.

ABT and for AAT, and (4) error in DEM. To calculate the total errors in the resulting ice thickness estimation and total volume calculation, quadratic addition was applied based on the above-mentioned four error sources (Vanlooy and others, 2014).

4. Results

4.1 ANN models

Out of 25 training runs conducted with the ABT model, three were rejected due to low combined *R*-values. The performances computed from the average of all folds for different networks are given in Table 1. Among the 22 remaining runs, the best correlation was 0.90 with a mean bias error for ice thickness of 13 m and RMSE of 24 m (Table 1 and 2). Out of 25 training runs conducted with the AAT model, which includes 20 GPR measurements as inputs, four runs were rejected (Table 2). Among the 25 runs, the best correlation was 0.93, with a mean bias error of 6 m and a RMSE of 13 m. The assimilation of GPR data for ANN network training decreased the RMSE by 11 m (24–13 m) and the mean thickness error of the five transverse sections by 7 m (13–6 m). Levenberg–Marquardt algorithm outperformed the gradient descent method in the current study.

4.2 Estimated ice thickness and volume

The performance of the ABT and AAT models is shown using scatter plots (Figs 5 and 6, respectively). These plots represent model performance based on the *R*-value for training, validation, testing and overall (a combination of training, validation, and testing), respectively. The solid line represents the best fit linear regression between the outputs and targets. The targets are the data (measured thicknesses) and the outputs are the results of the ANN model. The ABT model gives an overall *R* of 0.90 whereas the AAT model gives an overall *R* of 0.93. Performance

Table 1. The performance of different ABT architectures is based on the combined *R*-value for training, validation and testing

Architecture	Combined <i>R</i> value
4x-8T-1L	0.77
4x-8T-1S	0.67
4x-8S-1T	0.82
4x-8S-1S (rejected run)	0.19
4x-8S-1L	0.53
4x-8T-1T	0.90
4x-8T-8T-1T	0.81
4x-8T-8T-1S	0.80
4x-8T-8S-1S	0.73
4x-8S-8S-1S	0.82
4x-8S-8T-1T	0.80
4x-8S-8S-1T	0.71
4x-8L-8S-1T (rejected run)	0.22
4x-8S-8L-1T	0.56
4x-8S-8S-1L	0.72
4x-8T-8S-1L	0.64
4x-8L-8T-1S	0.75
4x-8S-8T-1S	0.84
4x-8S-8T-1L	0.67
4x-8T-8T-1L	0.81
4x-8T-8S-1T	0.69
4x-8L-8T-1T (rejected run)	0.23

Table 2. Performance of ANN modelling with and without GPR training

Performance criterion	ABT architecture (4x-8T-1T)	AAT architecture (4x-8T-1T)
Total runs	25	25
Accepted runs	22	21
Rejected runs	3	4
Mean training correlation coefficient (<i>r</i>)	0.90	0.94
Multiple <i>R</i>	0.90	0.93
RMSE bed elevation (m)	24	13
Mean bias error for ice thickness (m)	13	6

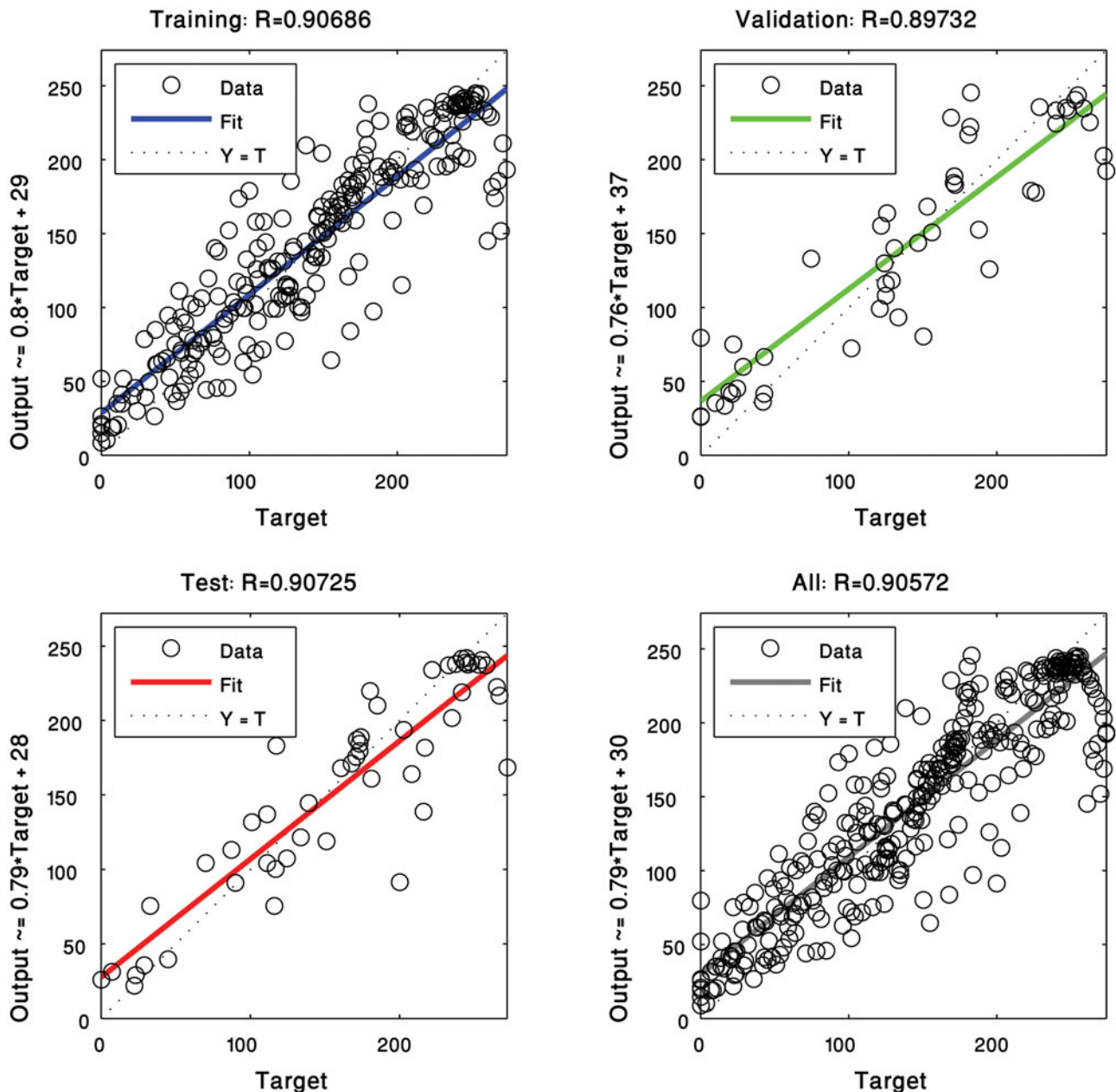


Fig. 5. Performance of ABT for Chhota Shigri Glacier ice thickness (in metres). The four plots display the outputs of the neural network model with respect to the targets for training, validation, and test sets of the neural network model.

statistics were computed from the average of all folds for the 4x-8T-1T set. The performance of the training, validation, and test sets are similar because of the L1 regularization used. Most ice thickness errors fall between ± 10 m (Fig. 7). The spread of errors ranges between ± 29 m and ± 0.83 m for maximum and minimum values, respectively.

The modelled ice-thickness estimates were 104 and 109 m for ABT and AAT, respectively. The volume of Chhota Shigri Glacier was estimated to be 1.61 and 1.69 km³ for ABT and AAT models, respectively.

4.3 Ice thickness profiles

Interpolated ice thicknesses for Chhota Shigri Glacier were obtained from the selected ABT and AAT models, and differences between the two approaches were calculated (Fig. 8). On the boundaries of the glacier, the ice thickness ranges from 0 to 15 m, whereas the mid area of the glacier shows ice-thickness values

ranging from 175 to 300 m. The altitudinal range between 4550 and 5000 m shows maximum ice thickness ranging from 200 to 300 m. The mean ice thickness of GPR points was 159 m whereas the mean thickness of the AAT model was estimated to be 152 m. The mean elevation of the DGPS survey (October 2009), mean altitude from the Cartosat-1 DEM, and the maximum and mean ice depth of GPR, ABT and AAT are shown in Table S1.

The ice thicknesses (bedrock topography) obtained from ABT, AAT, GlabTop (Frey and others, 2014) and Glabtop-2 (Ramsankaran and others, 2018) models were compared with GPR thickness measurements for all five transverse sections (Fig. 9). The maximum ice thickness was observed at cross section 4 for both the AAT and ABT models but the minimum ice thickness given by the ABT model was slightly higher than the GPR value (Table S1). The mean surface elevation, maximum ice depth and mean ice depth for additional cross sections A-E are given in Table S2. A maximum ice thickness of 290 was estimated at cross section D using the AAT model. The ANN modelled

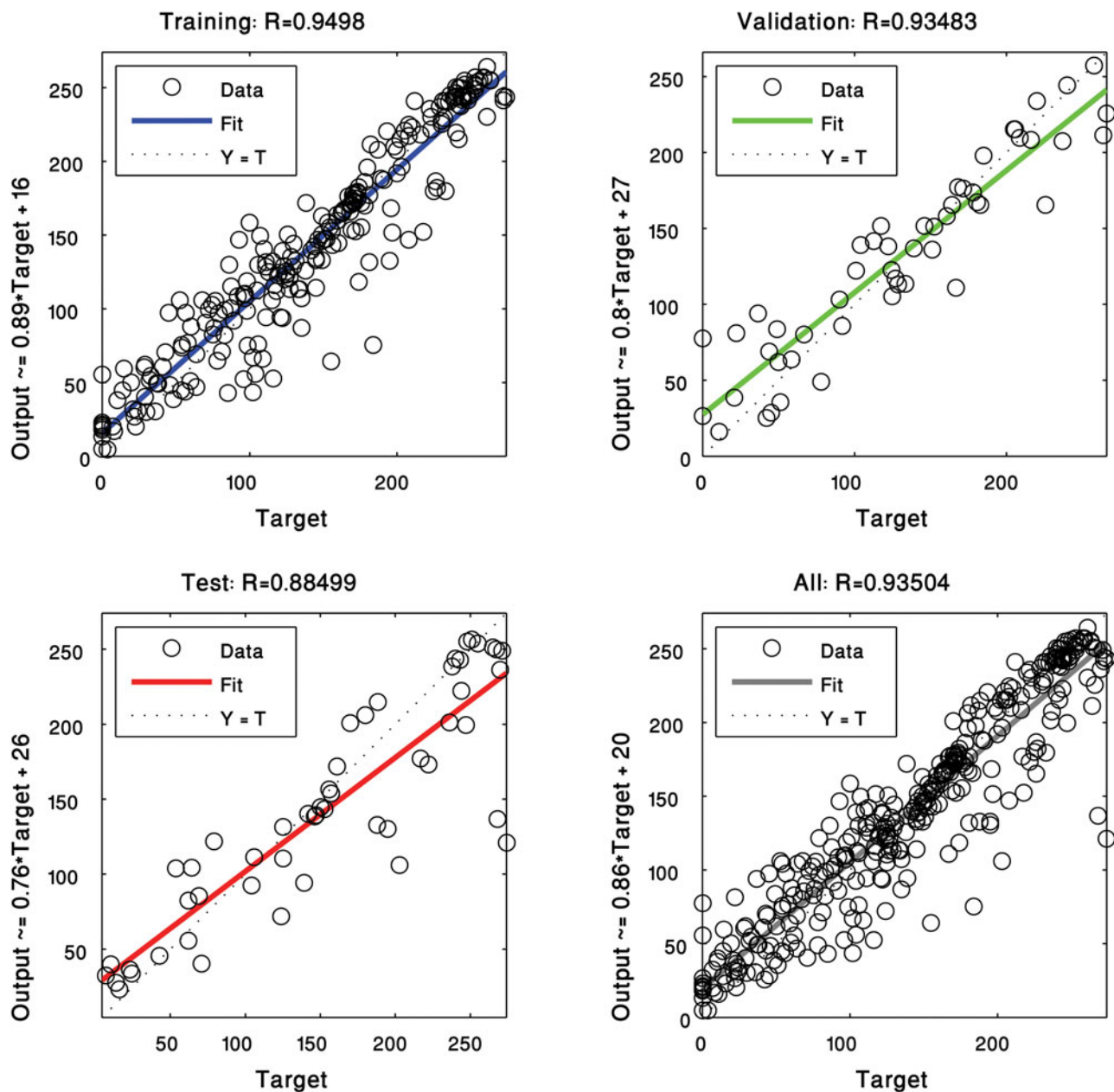


Fig. 6. Performance of AAT for Chhota Shigri Glacier ice thickness (in metres). The four plots display the outputs of the neural network model with respect to the targets for training, validation, and test sets of the neural network model.

ice-thickness distribution for sections A–E is plotted in Figure 10. The elevation vs distance along the longitudinal cross section (L1) of Chhota Shigri surface and ice thickness vs distance obtained using the ANN (AAT) model are shown in Figure 11.

Modelled ice thicknesses show varying levels of agreement with the observations at all five cross sections (Fig. 9). The AAT model showed lesser deviations towards the west side while the deviations were greater on the east side when compared to the GPR cross sections. The ABT model also showed greater deviations for all sections on the east side. The reason behind this shift may be the topographic pattern of the eastern sidewall of the glacier, given that the transverse section also included a portion of the glacier sidewall (300 m) in the input and training datasets. Another reason may be related to the interpolated values of GPR data at the sides due to the complexities of data collection.

To understand the relationship of ABT and AAT data points with GPR data points, the Pearson R was used. The values of R for ABT and AAT are given in Table S3. AAT performed better

than ABT for cross sections 1–5, as suggested by the R values. ABT showed the highest correlation for cross section 3 whereas AAT showed the best correlation for cross section 2. The differences are probably due to the fact that AAT used partial GPR information of four points per profile randomly whereas ABT used only extrapolated values as targets. For both cases, the lowest correlation was observed for cross section 1. Large differences between observed and modelled ice thicknesses at transverse section 1 (Fig. 12), particularly at greater ice thicknesses, may be due to variations in surface slope, which is used as a predictor in the ANNs.

4.4 Error analysis

The overall RMSE of the DEM was estimated to be 10 m. While the Cartosat-1 DEM was generated from April 2007 imagery, the GCPs were collected in October 2009 and October 2010. The mean bias error for transverse cross section altitudes between

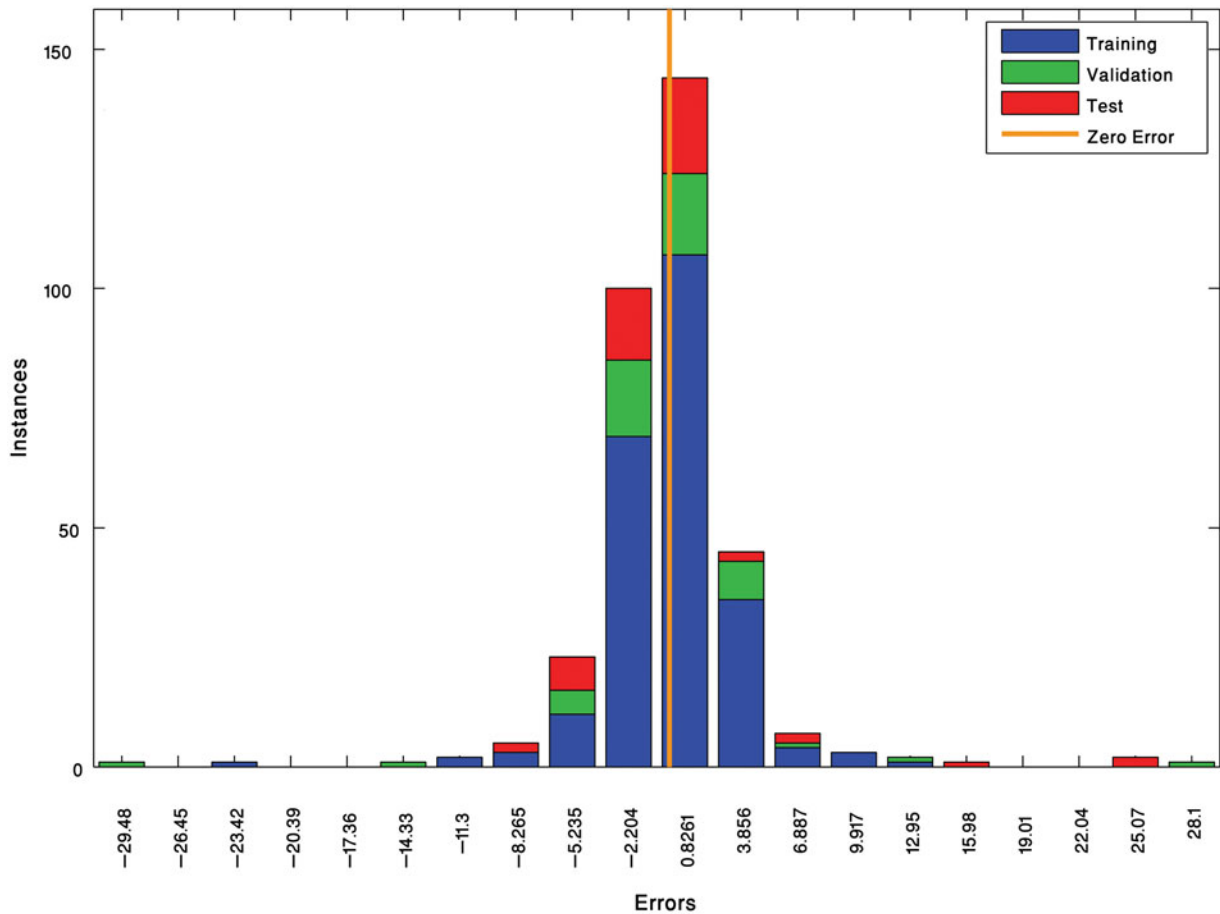


Fig. 7. Error histogram of the ABT model with 20 bins showing the error distribution (error histograms show training, validation and testing data as blue, green and red bars, respectively).

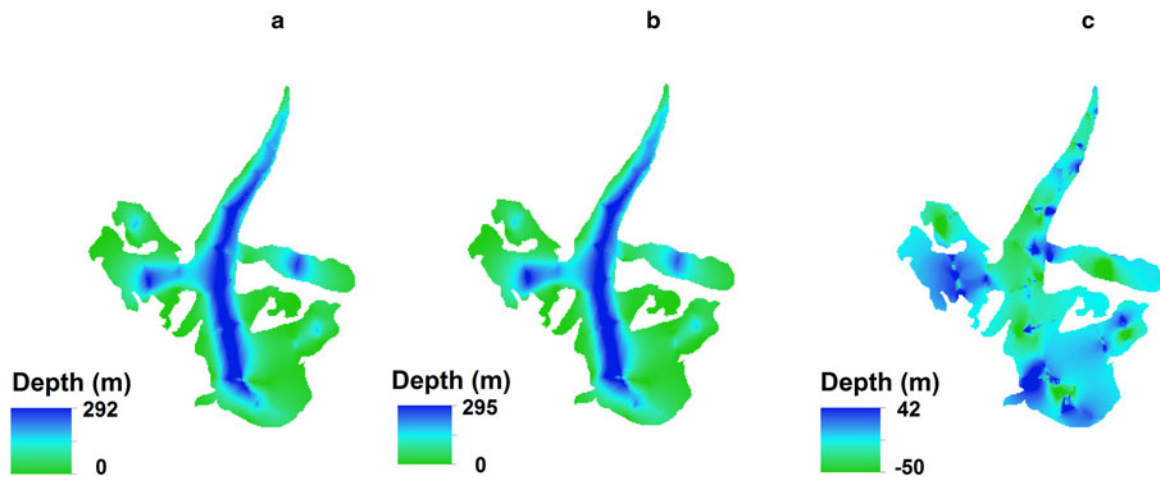


Fig. 8. Interpolated ice-thicknesses of Chhota Shigri Glacier, obtained from the ABT model (a) and the AAT model (b). The difference between the two approaches is shown in (c).

DEM and GCPs was 4.6 m (Table S1). Based on the lower RMSE and mean bias errors, the temporal difference between DEM and GCPs can be ignored.

Comparisons of modelled and observed ice thicknesses that include interpolated GPR data yield RMSEs of 27 and 16 m for ABT and AAT models, respectively. Without the interpolated data, the RMSE decreased to 24 and 13 m for ABT and AAT models, respectively.

The errors in GPR thickness (± 15 m), mean average bias of ABT (± 13 m) and AAT (± 6 m), error in kriging interpolation

(± 3.5 m for ABT and ± 3 m for AAT) and error in DEM (± 5 m) were calculated as quadratic addition (Vanlooy and others, 2014). The total error for glacier ice thickness estimation using quadratic addition was ± 21 m (ABT) and ± 17 m (AAT). Therefore, mean ice-thickness estimates with errors were 104 ± 21 and 109 ± 17 m for ABT and AAT, respectively. The total error in ice thickness was converted to a total error in volume by multiplying by the area of the glacier (15.5 km^2). The result has a total uncertainty in the ice volume of $\pm 0.33 \text{ km}^3$ (ABT) and $\pm 0.26 \text{ km}^3$ (AAT). Therefore, the volume of the Chhota

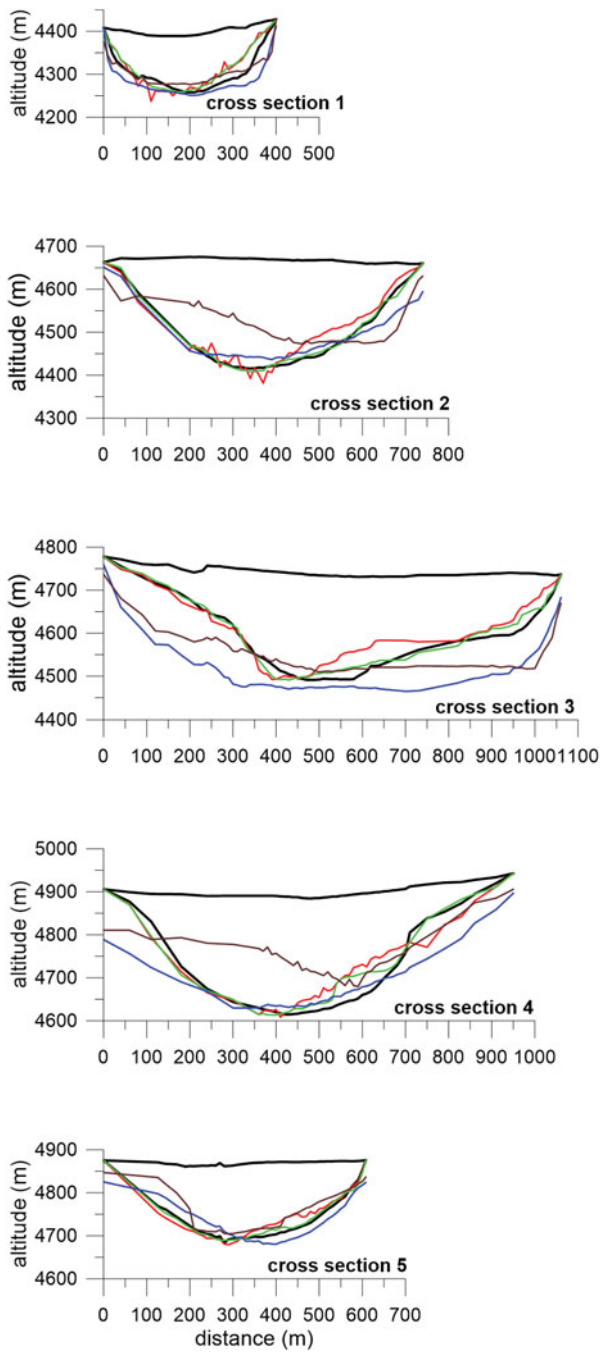


Fig. 9. Comparison of observed and modelled ice thicknesses at five cross sections using the AAT model (green), ABT model (red), GPR measurements (black), GlabTop2 model of Ramsankaran and others (2018) (blue), and the model of Frey and others (2014) (brown).

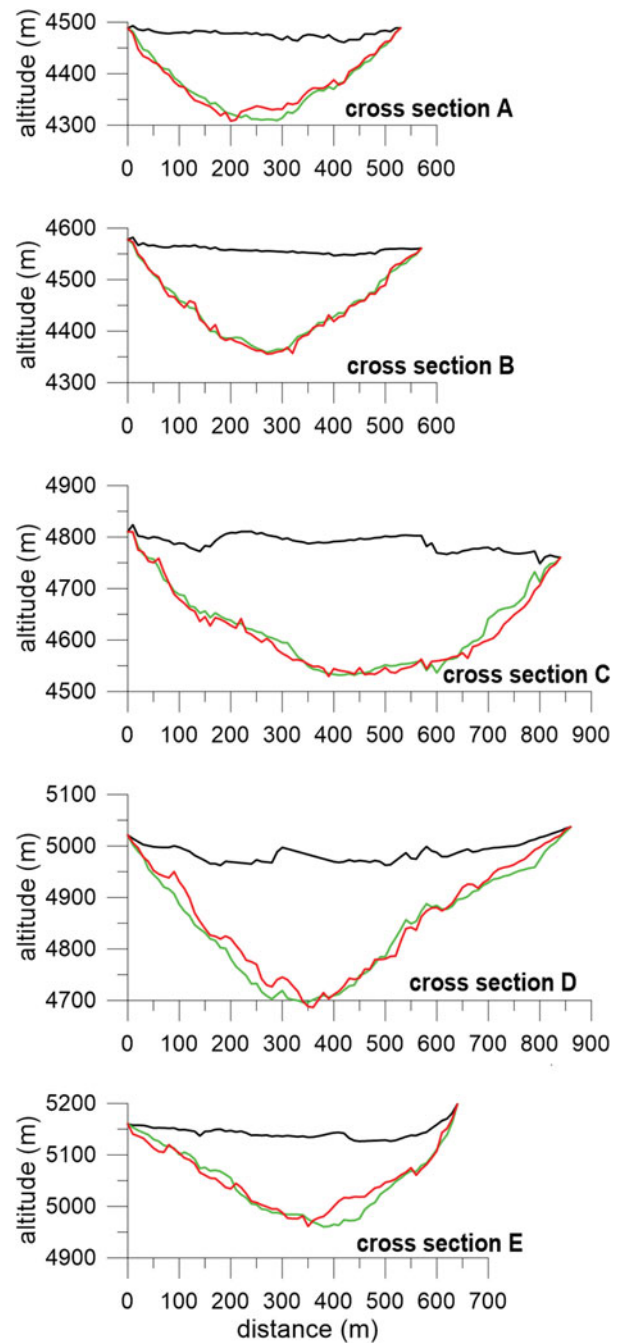


Fig. 10. Modelled ice thicknesses at transects A-E (green for AAT and red for ABT), with surface elevations from the DGPS survey (black).

Shigri Glacier is estimated to be 1.61 ± 0.33 and 1.69 ± 0.26 km³ for ABT and AAT models, respectively. The minor differences in the ice-thickness uncertainty between the GPR estimates (± 15 m) reported by Azam and others (2012) and the present modelling estimates (± 21 and ± 17 m) indicate that the ANN model performance is acceptable.

5. Discussion

Field-based GPR surveys are one of the best available methods for ice-thickness measurements. However, due to rough terrain, harsh climatic conditions and high expedition costs, GPR surveys for ice-thickness estimates have been conducted on very few glaciers in the Himalaya. The motivation for this study is to estimate ice

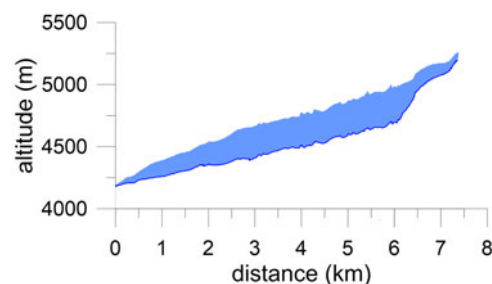


Fig. 11. Curves of elevation vs distance along the longitudinal cross section (L1) of Chhota Shigri surface (in light blue) and ice thickness vs distance obtained using the ANN (AAT) model (dark blue) for the section L1. The glacier is shown as a light blue fill.

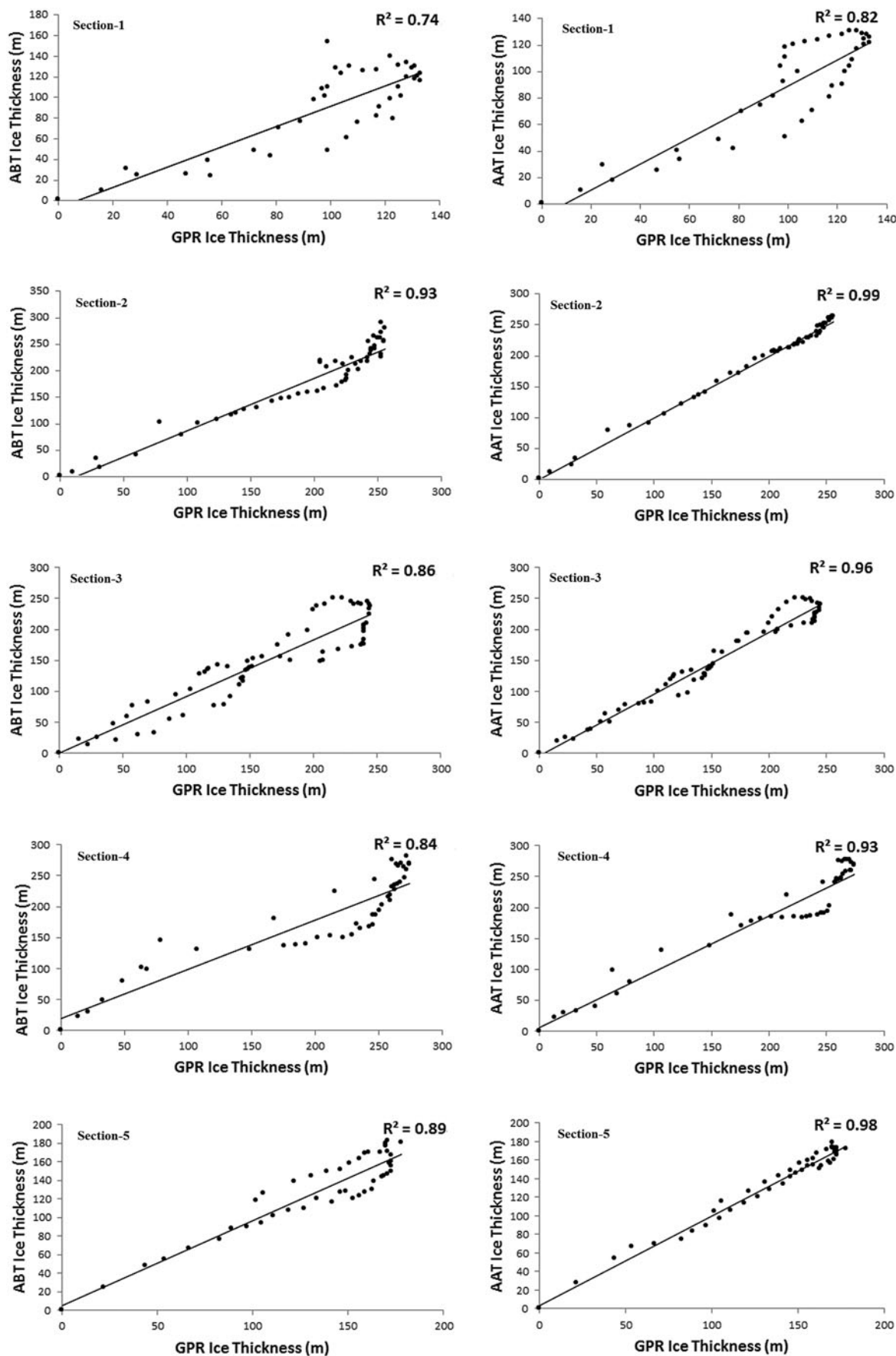


Fig. 12. Left: scatter plots of ice thicknesses obtained using the ABT model vs GPR ice thickness measurements for cross sections (1–5). Right: scatter plots of ice thicknesses obtained using the AAT model vs GPR ice thickness measurements for cross sections (1–5).

Table 3. The volume of Chhota Shigri Glacier is derived from different methods and studies

Methods	Volume (km ³)	Reference
ANN (ABT)	1.61 ± 0.33	Current investigation
ANN (AAT)	1.69 ± 0.26	Current investigation
GlabTop-2	1.74 ± 0.25	Ramsankaran and others (2018)
Glacier flow mechanics	0.55	Vashishth and others (2017)
Thickness-area scaling	1.40	Vashishth and others (2017)
Volume-area scaling	7.15	Vashishth and others (2017)
Thickness-area scaling and GPR data	1.20	Singh and others (2012)
Residual Bouguer values	0.64	Dobhal and others (1995)

thickness for individual glaciers with or without GPR measurements. In the absence of validation data, the performance of the ANN model was difficult to assess in the case of Gangotri glacier (Haq and others, 2014). Haq and others (2014) used the slope of sidewalls that extended 500 m out from the glacier to train an ANN, but the sidewalls were treated as a single entity.

5.1 Limitations

ANNs do not make use of glacier physics, so the modelling requires careful selection of parameters. One of the important issues faced during the present investigation was the selection of an optimal number of transverse sections. Ten transverse cross sections were used in this study, five of which were coincident with GPR measurements of ice thickness. More cross sections could potentially be added, but model stability and complex topography at higher elevations (>5200 m a.s.l.) limit the value of having additional cross sections.

Major limitations of ANN models include their black box nature, the possibility of overfitting, and the empirical nature of model development. Another shortcoming of the ANN approach is that it is computationally intensive and hence difficult to apply over large regions. Studies such as Frey and others (2014) and Farinotti and others (2020) applied their approaches on a regional scale. Regional-scale models require normalization of parameters such as shape factor to account for varying glacier shapes and topography. The present approach, calibrated for one glacier with in situ observations, is more likely to perform well on similar glaciers than a model trained at a regional scale.

5.2 Comparison with other studies

In this section, we compare our results with those of previously published studies on Chhota Shigri Glacier (Table 3). Our model has two outputs: ABT and AAT. One of the goals of the current study is to show the capability of ANN models to simulate ice thickness without calibration. Other studies, such as Frey and others (2014) and Ramsankaran and others (2018), used GPR data from Chhota Shigri Glacier to train the GlabTop and the improved GlabTop2 models (Table 3, Table S4). The outputs from these models were compared with those of our trained model (AAT) to show the strength of ANN modelling. In the present study, the mean ice thickness for the entire Chhota Shigri Glacier is estimated to be 109 m, whereas Ramsankaran and others (2018) estimated it to be 121 m. The ice thickness of five transverse cross sections (1–5) is plotted in Figure 7 along with the transverse cross sections generated by Ramsankaran and others (2018) and Frey and others (2014) using the 10 m DEM from Ramsankaran and others (2018). AAT and ABT showed a better match with the GPR data (Tables S4 and S5) compared

to previous ice thickness estimates by Frey and others (2014) and Ramsankaran and others (2018). The modelled average ice-thickness errors ranging from 6 to 21 m for ABT and from 1 to 13 m for AAT were lower than the mean bias error for ice-thickness ranging from 9 to 22 m estimated by Ramsankaran and others (2018) at different transverse sections. The mean bias error for ice-thickness was also found to be lower than estimated by Frey and others (2014) (Table S4). The high errors in Frey and others (2014) were probably due to the relatively low-resolution DEM (SRTM 90 m) they used and the value of the shape factor (Ramsankaran and others, 2018). To address the spatial resolution issue in Frey and others (2014), we used a 10 m DEM (TandemX) developed by Ramsankaran and others (2018) and re-estimated the ice thicknesses. It was observed that the mean bias error for ice-thickness for transverse cross sections was between –1 and 53 m. Additionally, the performance for each transverse cross section based on the RMSE is given in Table S5. Figure 7 shows that GlabTop results can have 100–150 m differences with GPR measurements at some locations. The ice volume of 1.69 ± 0.26 km³ for Chhota Shigri Glacier derived using the AAT model in the present study showed good agreement with the volume of 1.74 ± 0.25 km³ derived by Ramsankaran and others (2018).

Vashisht and others (2017) used glacier flow mechanics, thickness-area scaling and volume-area scaling methods and estimated the volume of Chhota Shigri Glacier as 0.55, 1.40 and 7.15 km³, respectively. Interestingly, the estimated volume using the volume–area scaling method was 13 times higher than the volume estimated from glacier flow mechanics. Singh and others (2012) used the thickness-area scaling method and GPR data and computed a volume of 1.20 km³. Dobhal and others (1995) used residual Bouguer values to compute a volume of 0.64 km³.

5.3 Outlook

Worthwhile directions for future study would be to investigate the potential of deep learning networks such as convolution neural networks (CNNs) with more number of parameters such as surface velocities (e.g. Dehecq and others, 2019) and mass balance (e.g. Wu and others, 2019) for more number of glaciers. The more the model is constrained with input predictors, the less dependency on data it will have. Another future scope of the present research will be to apply the ABT model on additional datasets for testing.

6. Conclusion

In this work, two ANN models were developed to estimate the glacier ice thickness at 10 transverse cross sections and two longitudinal cross sections on Chhota Shigri Glacier in western Himalaya, India. The model results were compared with the GPR-derived ice thicknesses on five transverse cross sections located between 4400 and 4900 m a.s.l. on Chhota Shigri Glacier. The results of the current investigation indicate that the ice thickness and volume of a glacier can be estimated fairly well using ANN models. Furthermore, an ANN model alone, without integrating ice thickness observations, was shown to provide a good estimate of the glacier ice-thickness distribution. The ABT estimated mean ice thickness and volume for Chhota Shigri Glacier are 104 ± 21 m and 1.61 ± 0.33 km³, respectively, while the AAT estimated mean ice thickness and volume for Chhota Shigri Glacier are 109 ± 17 m and 1.69 ± 0.26 km³, respectively.

The ABT and AAT estimates of ice thickness showed good agreement with the GPR measurements available for five cross sections. Mean bias errors for ice thickness along the measured profiles ranging from 6 to 21 m for ABT and from 1 to 13 m

for AAT are lower than the average ice-thickness errors ranging from 9 to 22 m estimated by Ramsankaran and others (2018) for all transverse cross sections. Based on the calculated RMSE values, ANN models trained with and without ice thickness observations provide reasonable estimates of ice thickness. The ANN architecture used in this study requires minimal input data that can be extracted for other locations. However, model transferability remains an open question, and future research should be directed towards CNNs and additional model inputs such as surface velocity (Dehecq and others, 2019) and mass balance (e.g. Wu and others, 2019). The more the model is constrained with input predictors, the less dependency on data it will have.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/jog.2021.19>

Acknowledgements. Mohd Anul Haq would like to thank the Deanship of Scientific Research at Majmaah University for supporting this work under Project No. R-2021-10. This work has also been supported by IFCPAR/CEFIPRA within project no. 3900-W1 and by the French GLACIOCLIM observation service. MFA acknowledges the research grant provided by the INSPIRE Faculty award (IFA-14-EAS-22) from the Department of Science and Technology (DST, India). The authors thank three anonymous reviewers for their constructive comments and suggestions. The authors would like to warmly thank Scientific Editor Dr Joseph Shea for numerous editorial revisions and suggestions. The authors are highly thankful to Holger Frey for sharing 90 m resolution GlabTop model data and Raaj Ramsankaran and Ankur Pandit for sharing GlabTop2 model data. The authors are also thankful to Dr Munir Nayak and Dr Pottakkal George Jose for language improvement.

References

- Abraham A (2005) Artificial neural networks. In Sydenham PH and Thorn R (eds), *Handbook of Measuring System Design*. London: John Wiley and Sons, pp. 901–908.
- Agatonovic-Kustrin S and Beresford R (2000) Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis* **22**(5), 717–727.
- Agrawal A and Tayal S (2013) Assessment of volume change in East Rathong glacier, Eastern Himalaya. *International Journal of Geoinformatics* **9**(1), 73–82.
- Azam MF and 10 others (2012) From balance to imbalance: a shift in the dynamic behaviour of Chhota Shigri Glacier, western Himalaya, India. *Journal of Glaciology* **58**(208), 315–324.
- Azam MF and 5 others (2014a) Reconstruction of the annual mass balance of Chhota Shigri Glacier, Western Himalaya, India, since 1969. *Annals of Glaciology* **55**(66), 69–80. doi: [10.3189/2014AoG66A104](https://doi.org/10.3189/2014AoG66A104).
- Azam MF and 6 others (2014b) Processes governing the mass balance of Chhota Shigri Glacier (Western Himalaya, India) assessed by point-scale surface energy balance measurements. *The Cryosphere* **8**(6), 2195–2217.
- Azam MF and 10 others (2016) Meteorological conditions, seasonal and annual mass balances of Chhota Shigri Glacier, Western Himalaya, India. *Annals of Glaciology* **57**(71), 328–338. doi: [10.3189/2016AoG71A570](https://doi.org/10.3189/2016AoG71A570).
- Azam MF and 5 others (2018) Review of the status and mass changes of Himalayan-Karakoram glaciers. *Journal of Glaciology* **64**(243), 61–74. doi: [10.1017/jog.2017.86](https://doi.org/10.1017/jog.2017.86).
- Azam MF and 7 others (2019) Snow and ice melt contributions in a highly glacierized catchment of Chhota Shigri Glacier (India) over the last five decades. *Journal of Hydrology* **574**, 760–773. <https://doi.org/10.1016/j.jhydrol.2019.04.075>.
- Bahr DB, Pfeffer WT and Kaser G (2015) A review of volume-area scaling of glaciers. *Reviews of Geophysics* **53**, 95–140. doi: [10.1002/2014RG000470](https://doi.org/10.1002/2014RG000470).
- Baumhoer CA, Dietz AJ, Kneisel C and Kuenzer C (2019) Automated extraction of antarctic glacier and ice shelf fronts from sentinel-1 imagery using deep learning. *Remote Sensing* **11**, 2529.
- Beale R and Jackson T (1990) *Neural Computing: An Introduction*. Bristol, England: Adam Hilger. doi: [10.1887/0852742622](https://doi.org/10.1887/0852742622).
- Bishop MP, Shroder JF Jr. and Hickman BL (1999) SPOT Panchromatic imagery and neural networks for information extraction in a complex mountain environment. *Geocarto International* **14**(2), 19–28. doi: [10.1080/10106049908542100](https://doi.org/10.1080/10106049908542100).
- Bliss A, Hock R and Cogley JG (2013) A new inventory of mountain glaciers and ice caps for the Antarctic periphery. *Annals of Glaciology* **54**(63), 191–199. doi: [10.3189/2013AoG63A377](https://doi.org/10.3189/2013AoG63A377).
- Bolch T and 10 others (2019) Status and change of the cryosphere in the extended Hindu Kush Himalaya region. In Koji Fujita ed. *The Hindu Kush Himalaya Assessment*. Cham: Springer, pp. 209–255. https://doi.org/10.1007/978-3-319-92288-1_7.
- Bolibar J and 5 others (2020) Deep learning applied to glacier evolution modelling. *The Cryosphere* **14**, 565–584. doi: [10.5194/tc-2019-163](https://doi.org/10.5194/tc-2019-163).
- Brun F, Berthier E, Wagnon P, Kääb A and Treichler D (2017) A spatially resolved estimate of high mountain Asia glacier mass balances from 2000 to 2016. *Nature Geosciences* **10**(9), 668–673. doi: [10.1038/ngeo2999](https://doi.org/10.1038/ngeo2999).
- Caiping C and Yongjian D (2009) The application of artificial neural networks to simulate meltwater runoff of Keqikaer Glacier, south slope of Mt. Tuomuer, Western China. *Environmental Geology* **57**(8), 1839–1845. doi: [10.1007/s00254-008-1471-1](https://doi.org/10.1007/s00254-008-1471-1).
- Clarke GKC, Berthier E, Schoof CG and Jarosch AH (2009) Neural networks applied to estimating subglacial topography and glacier volume. *Journal of Climate* **22**(8), 2146–2160. doi: [10.1175/2008JCLI2572.1](https://doi.org/10.1175/2008JCLI2572.1).
- Czyzowska-Wisniewski EH, van Leeuwen WJ, Hirschboeck KK, Marsh SE and Wisniewski WT (2015) Fractional snow cover estimation in complex alpine-forested environments using an artificial neural network. *Remote Sensing of Environment* **156**, 403–417. doi: [10.1016/j.rse.2014.09.026](https://doi.org/10.1016/j.rse.2014.09.026).
- Dehecq, A and 6 others (2019) Twenty-first century glacier slowdown driven by mass loss in high mountain Asia. *Nature Geoscience* **12**(1), 22–27. <https://doi.org/10.1038/s41561-018-0271-9>.
- Demuth H and Beale M (2000) *Neural Network Toolbox, User's Guide*. US: The MathWorks, Inc..
- De Villiers J and Barnard E (1993) Backpropagation neural nets with one and two hidden layers. *IEEE Transactions on Neural Networks* **4**, 136–141. doi: [10.1109/72.182704](https://doi.org/10.1109/72.182704).
- Doberva ID and Klein AG (2009) Artificial neural networks approach to fractional snow cover mapping. *Proceedings of 66th Eastern Snow Conference Niagara-on-the-Lake, 9–11 June, Ontario, Canada*.
- Dobhal DP, Kumar S and Mundepi, AK (1995) Morphology and glacier dynamics studies in monsoon-arid transition zone: an example from Chhota Shigri Glacier, Himachal-Himalaya, India. *Current Science* **68**(9), 936–944.
- Eluyode OS and Akomolafe DT (2013) Comparative study of biological and artificial neural networks. *European Journal of Applied Engineering and Scientific Research* **2**(1), 36–46.
- Farinotti D and 10 others (2017) How accurate are estimates of glacier ice thickness? Results from ITMIX, the ice thickness models intercomparison experiment. *The Cryosphere* **11**(2), 949–970. doi: [10.5194/tc-11-949-2017](https://doi.org/10.5194/tc-11-949-2017).
- Farinotti D and others (2020) Results from the ice thickness models intercomparison experiment Phase 2. *Frontiers of Earth Science* **8**, 571923. doi: [10.3389/feart.2020.571923](https://doi.org/10.3389/feart.2020.571923).
- Farinotti D, Huss M, Bauder A and Funk M (2009b) An estimate of the glacier ice volume in the Swiss Alps. *Global and Planetary Change* **68**(3), 225–231. doi: [10.1016/j.gloplacha.2009.05.004](https://doi.org/10.1016/j.gloplacha.2009.05.004).
- Farinotti D, Huss M, Bauder A, Funk M and Truffer M (2009a) A method to estimate ice volume and ice-thickness distribution of alpine glaciers. *Journal of Glaciology* **55**(191), 422–430. doi: [10.3189/002214309788816759](https://doi.org/10.3189/002214309788816759).
- Fortuna L, Ferla G and Imbruglia A (2002) *Microelectronics and Microsystems: Emergent Design Techniques*. London: Springer.
- Frey H and 9 others (2014) Estimating the volume of glaciers in the Himalayan-Karakoram region using different methods. *The Cryosphere* **8**(6), 2313–2333. doi: [10.5194/tc-8-2313-2014](https://doi.org/10.5194/tc-8-2313-2014).
- Gallant SI (1986) Three constructive algorithms for network learning. *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, 652–660.
- Gantayat P, Kulkarni AV and Srinivasan J (2014) Estimation of ice thickness using surface velocities and slope: case study at Gangotri Glacier, India. *Journal of Glaciology* **60**(220), 277–282. doi: [10.3189/2014JoG13J078](https://doi.org/10.3189/2014JoG13J078).
- Gavin HP (2019) *The Levenberg-Marquardt Method for Nonlinear Least Squares Curve-Fitting Problems*. Duke University: Computer Science.
- Gergan JT, Dobhal DP and Kaushik R (1999) Ground penetrating radar ice thickness measurements of Dokriani bamak (glacier), Garhwal Himalaya. *Current Science* **77**, 169–173.
- Haerberli W and Hoelzle M (1995) Application of inventory data for estimating characteristics of and regional climate-change effects on mountain glaciers: a pilot study with the European Alps. *Annals of Glaciology* **21**, 206–212.

- Hagan MT and Menhaj MB (1994) Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks* 5(6), 989–993. doi: [10.1109/72.329697](https://doi.org/10.1109/72.329697).
- Haq MA, Jain K and Menon KPR (2014) Modelling of Gangotri glacier thickness and volume using an artificial neural network. *International Journal of Remote Sensing* 35(16), 6035–6042. doi: [10.1080/01431161.2014.943322](https://doi.org/10.1080/01431161.2014.943322).
- Hubbard B and Glasser NF (2005) *Field Techniques in Glaciology and Glacial Geomorphology*. USA: John Wiley and Sons. ISBN: 978-0-470-84426-7.
- Huss M and Farinotti D (2012) Distributed ice thickness and volume of all glaciers around the globe. *Journal of Geophysical Research: Earth Surface* 117(F4), 1–10. doi: [10.1029/2012JF005253](https://doi.org/10.1029/2012JF005253).
- Iwata S, Watanabe O and Fushimi H (1980) Surface morphology in the ablation area of the Khumbu glacier. *Journal of the Japanese Society of Snow and Ice* 41, 9–17.
- Kisi Ö and Uncuoğlu E (2005) Comparison of three back-propagation training algorithms for two case studies. *Indian Journal of Engineering & Materials Sciences* 12, 434–442.
- Kwok T-Y and Yeung D-Y (1995) Constructive feedforward neural networks for regression problems: A survey (Tech. Rep. No. HKUST-CS95-43). Clear Water Bay, Kowloon, Hong Kong: Department of Computer Science, Hong Kong University of Science and Technology.
- Levenberg K (1944) A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics* 2(2), 164–168. doi: [10.1090/qam/10666](https://doi.org/10.1090/qam/10666).
- Little MA, Varoquaux G and Saeb S (2017) Using and understanding cross-validation strategies. Perspectives on Saeb and others. *Gigascience* 6(5), 1–6. doi: [10.1093/gigascience/gix020](https://doi.org/10.1093/gigascience/gix020).
- Lliboutry L (1987) Realistic, yet simple bottom boundary conditions for glaciers and ice sheets. *Journal of Geophysical Research: Solid Earth* 92(B9), 9101–9109. doi: [10.1029/JB092iB09p09101](https://doi.org/10.1029/JB092iB09p09101).
- Mandal A and others (2020) Understanding the interrelationships among mass balance, meteorology, discharge and surface velocity on Chhota Shigri Glacier over 2002–2019 using in situ measurements. *Journal of Glaciology* 66(259), 1–15. <https://doi.org/10.1017/jog.2020.42>.
- Marquardt DW (1963) An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* 11(2), 431–441.
- McCarthy M, Pritchard H, Willis I and King E (2017) Ground-penetrating radar measurements of debris thickness on Lirung Glacier, Nepal. *Journal of Glaciology* 63(239), 543–555. doi: [10.1017/jog.2017.18](https://doi.org/10.1017/jog.2017.18).
- Meier MF and Bahr DB (1996) Counting glaciers: use of scaling methods to estimate the number and size distribution of the glaciers of the world. In Colbeck SC ed. *Glaciers, Ice Sheets and Volcanoes: A Tribute to Mark F. Meier*, vol. 96. USA: US Army Corps of Engineer, 89–94.
- Mishra A, Negi BDS, Banerjee A, Nainwal HC and Shankar R (2018) Estimation of ice thickness of the Satopanth Glacier, Central Himalaya using ground penetrating radar. *Current Science* 114(4), 785–791. <https://doi.org/10.18520/cs/v114/i04/785-791>.
- Mishra B, Tripathi NK and Babel MS (2014) An artificial neural network-based snow cover predictive modeling in the higher Himalayas. *Journal of Mountain Science* 11(4), 825–837. doi: [10.1007/s11629-014-2985-5](https://doi.org/10.1007/s11629-014-2985-5).
- Mohajerani Y (2019) Understanding Regional Ice Sheet Mass Balance: Remote Sensing, Regional Climate Models, and Deep Learning. UC Irvine. ProQuest ID: Mohajerani_uci_0030D_16077. Merritt ID: arkl/13030/m54f6z56. Retrieved from <https://escholarship.org/uc/item/7jj888jg>.
- Mosbeux C, Gillet-Chaulet F and Gagliardini O (2016) Comparison of adjoint and nudging methods to initialise ice sheet model basal conditions. *Geoscientific Model Development* 9, 2549–2562. doi: [10.5194/gmd-9-2549-2016](https://doi.org/10.5194/gmd-9-2549-2016).
- Nye J (1952) The mechanics of glacier flow. *Journal of Glaciology* 2(12): 82–93. doi: [10.3189/S0022143000033967](https://doi.org/10.3189/S0022143000033967).
- Pattyn F, Delcourt C, Samyn D, De Smedt B and Nolan M (2009) Bed properties and hydrological conditions underneath McCall Glacier, Alaska, USA. *Annals of Glaciology* 50(51), 80–84. doi: [10.3189/172756409789097559](https://doi.org/10.3189/172756409789097559).
- Paul F and Linsbauer A (2012) Modeling of glacier bed topography from glacier outlines, central branch lines, and a DEM. *International Journal of Geographical Information Science* 26(7), 1173–1190. doi: [10.1080/13658816.2011.627859](https://doi.org/10.1080/13658816.2011.627859).
- Pettersson, R, Jansson P and Holmlund P (2003) Cold surface layer thinning on Storglaciären, Sweden, observed by repeated ground penetrating radar surveys. *Journal of Geophysical Research: Earth Surface* 108(F1), 1–9. doi: [10.1029/2003JF000024](https://doi.org/10.1029/2003JF000024).
- Rabatel A, Sanchez O, Vincent C and Six D (2018) Estimation of glacier thickness from surface mass balance and ice flow velocities: a case study on Argentière Glacier, France. *Frontiers in Earth Science* 6(12), 1–16. doi: [10.3389/feart.2018.00112](https://doi.org/10.3389/feart.2018.00112).
- Radić V and Hock R (2011) Regionally differentiated contribution of mountain glaciers and ice caps to future sea-level rise. *Nature Geoscience* 4(2), 91. doi: [10.1038/NNGEO1052](https://doi.org/10.1038/NNGEO1052).
- Ramsankaran R, Pandit A and Azam MF (2018) Spatially distributed ice-thickness modelling for Chhota Shigri Glacier in western Himalayas, India. *International Journal of Remote Sensing* 39(10), 3320–3343.
- Raper SCB and Braithwaite RJ (2005) The potential for sea level rise: new estimates from glacier and ice cap area and volume distributions. *Geophysical Research Letters* 32(5), 1–4. doi: [10.1029/2004GL021981](https://doi.org/10.1029/2004GL021981).
- Raymond-Pralong, M and Gudmundsson, G. (2011) Bayesian Estimation of basal conditions on Rutford Ice Stream, West Antarctica, from surface data. *Journal of Glaciology* 57, 315–324. doi: [10.3189/002214311796406004](https://doi.org/10.3189/002214311796406004).
- Sauter T, Schneider C, Kilian R and Moritz M (2009) Simulation and analysis of runoff from a partly glaciated meso-scale catchment area in Patagonia using an artificial neural network. *Hydrological Processes* 23(7), 1019–1030. doi: [10.1002/hyp.7210](https://doi.org/10.1002/hyp.7210).
- Sauter T, Weitzkamp B and Schneider C (2010) Spatio-temporal prediction of snow cover in the Black Forest mountain range using remote sensing and a recurrent neural network. *International Journal of Climatology* 30(15), 2330–2341. doi: [10.1002/joc.2043](https://doi.org/10.1002/joc.2043).
- Sheela KG and Deepa SN (2013) Review on methods to fix number of hidden neurons in neural networks. *Mathematical Problems in Engineering* 2013, 1–11. doi: [10.1155/2013/425740](https://doi.org/10.1155/2013/425740).
- Singh V (2011) *Encyclopedia of Snow, Ice and Glaciers*, 1st edition. New York: Springer.
- Singh SK, Rathore BP, Bahuguna IM, Ramnathan and AL and Ajai (2012) Estimation of glacier ice thickness using ground penetrating radar in the Himalayan region. *Current Science* 68–73.
- Steiner D, Walter A and Zumbühl HJ (2005) The application of a non-linear back-propagation neural network to study the mass balance of Grosse Aletschgletscher, Switzerland. *Journal of Glaciology* 51(173), 313–323.
- Vanlooy JA, Miede C, Vandeberg GS and Forster RR (2014) Ice volume estimation inferred from ice thickness and surface measurements for Continental Glacier, Wind River Range, Wyoming, USA. *Journal of Glaciology* 60(221), 478–488. doi: [10.3189/2014JG13J162](https://doi.org/10.3189/2014JG13J162).
- Vashishth P, Pandey M, Ramanathan AL, Tayal S and Jackson M (2017) Comparative assessment of volume change in Kolahoi and Chhota Shigri Glaciers, Western Himalayas, using empirical techniques. *Journal of Climate Change* 3(1), 37–48.
- Vincent C and 10 others (2013) Balanced conditions or slight mass gain of glaciers in the Lahaul and Spiti region (northern India, Himalaya) during the nineties preceded recent mass loss. *The Cryosphere* 7(2), 569–582. doi: [10.5194/tc-7-569-2013](https://doi.org/10.5194/tc-7-569-2013).
- Vincent C and 10 others (2016) Reduced melt on debris-covered glaciers: investigations from Changri Nup Glacier, Nepal. *The Cryosphere* 10, 1845–1858. doi: [10.5194/tc-10-1845-2016](https://doi.org/10.5194/tc-10-1845-2016).
- Wagnon P and 10 others (2007) Four years of mass balance on Chhota Shigri Glacier, Himachal Pradesh, India, a new benchmark glacier in the Western, Himalaya. *Journal of Glaciology* 53(183), 603–610.
- Wagnon P and 11 others (2013) Seasonal and annual mass balances of Mera and Pokalde glaciers (Nepal Himalaya) since 2007. *The Cryosphere* 7, 1769–1786. doi: [10.5194/tc-7-1769-2013](https://doi.org/10.5194/tc-7-1769-2013).
- Wilson P and Mantooth HA (2013) Model-based optimization techniques. In Wilson P and Mantooth HA (eds), *Model-Based Engineering for Complex Electronic Systems*, Newnes 2013, 347–367. ISBN 9780123850850.
- Wu K, Liu S, Jiang Z, Xu J and Wei J (2019) Glacier mass balance over the central Nyainqentanglha Range during recent decades derived from remote-sensing data. *Journal of Glaciology* 65(251), 422–439. doi: [10.1017/jog.2019.20](https://doi.org/10.1017/jog.2019.20).
- Zumbühl HJ, Steiner D and Nussbaumer SU (2008) 19th Century glacier representations and fluctuations in the central and Western European Alps: an interdisciplinary approach. *Global and Planetary Change* 60(1–2), 42–57. doi: [10.1016/j.gloplacha.2006.08.005](https://doi.org/10.1016/j.gloplacha.2006.08.005).