

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Prediction of Relative Humidity in a High Elevated Basin of Western Karakoram by Using Different Machine Learning Models

*Muhammad Adnan, Rana Muhammad Adnan, Shiyin Liu, Muhammad Saifullah, Yasir Latif and Mudassar Iqbal*

## Abstract

Accurate and reliable prediction of relative humidity is of great importance in all fields concerning global climate change. The current study has employed Multivariate Adaptive Regression Spline (MARS) and M5 Tree (M5T) models to predict the relative humidity in the Hunza River basin, Pakistan. Both the models provided the best prediction for the input scenario S6 (RHt-1, RHt-2, RHt-3, Tt-1, Tt-2, Tt-3). The statistical analysis displayed that the MARS model provided a better prediction of relative humidity as compared to M5T at all meteorological stations, especially, at Ziarat followed by Khunjerab and Naltar. The values of root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination ( $R^2$ ) were (5.98%, 5.43%, and 0.808) for Khunjerab; (6.58%, 5.08%, and 0.806) for Naltar; and (5.86%, 4.97%, 0.815) for Ziarat during the testing of MARS model whereas, the values were (6.14%, 5.56%, and 0.772) for Khunjerab; (6.19%, 5.58% and 0.762) for Naltar and (6.08%, 5.46%, 0.783) for Ziarat during the testing of M5T model. Both the models performed slightly better in training as compared to the testing stage. The current study encourages future research to be conducted at high altitude basins for the prediction of other meteorological variables using machine learning tools.

**Keywords:** relative humidity, MARS, M5T, Hunza, machine learning

## 1. Introduction

The relative humidity is defined as the amount of water vapor in the air in comparison with the full saturation [1, 2]. Being the important indicator of precipitation forecasting, its prediction plays a significant part in improving the accuracy of weather forecasting [3]. The relative humidity changes with respect to change in saturated vapor pressure which further depends on wind speed, solar radiation, pressure, temperature, and moisture content in the air [1]. The relative humidity is a function of temperature and is regarded as a sensitive parameter in the field of science [4]. Relative humidity plays a vital role in plant growth, agricultural and industrial production and in the prevention and control of air pollution [5];

economic stability of a region, water systems and also in managing renewable and solar energy systems [1, 6], weather and climate [7, 8]. Moreover, it has also an impact on ozone concentration and adaptive thermal comfort [9]. Keeping in view the importance of relative humidity, the research on its prediction is increasingly important [7].

The relative humidity is an important aspect of the hydrological phase [8] and has a role in alpine hydrology, especially, in a cold and dry climate; any change in temperature and humidity causes larger variations in the ablation of glaciers [10]. The warm environment glaciers are subjected to be influenced more by the change in relative humidity. Few other studies e.g. [11–13] also observed that tropical glaciers are sensitive to subtle changes in relative humidity, precipitation, and cloudiness. Relative humidity and clouds play an important role in the energy balance of glaciers by controlling the number of outgoing longwave radiation. Moreover, relative humidity and wind speed influence the turbulent latent heat flux which supplies all energy for sublimation and thus they indirectly control the equilibrium line altitude (ELA) [14]. Another study conducted by [15] observed that relative humidity has an effect on evaporation and there is an inverse relation between them. Evaporation further controls the water balance of closed lakes in hilly areas and evapotranspiration, especially, in irrigated agricultural areas.

Regardless of relative humidity is an important component of hydrology, meteorology, and climate, only a few studies are available for its prediction. A study conducted by [1] used artificial neural networks (ANNs) and genetic expression programming (GEP) models for the prediction of relative humidity as a function of three meteorological variables: wind speed, temperature, and pressure in two Californian gauging stations. They observed that both the models can successfully predict one-year relative humidity data into the future. Another study done by [5] predicted relative humidity by establishing time series models such as Extreme Gradient Boosting (XGBoost), Seasonal Auto-Regressive Integrated Moving Average (SARIMA), and Holt-Winters (HW). The XGBoost was found more accurate because of its robust capability to resist a fitting. The study conducted by [3] found that the performance of an autoregressive integrated moving average (ARIMA) model is better than the Long Short-Term Memory (LSTM) Network for the prediction of relative humidity. On contrary, [8] observed that the LSTM network is capable of predicting complex univariate relative humidity time series with robust no-stationarity. However, Least Square Support Vector Machine (LSSVM) and Adaptive Network-Based Fuzzy Inference System (ANFIS) models were used by [2] for prediction of relative humidity in terms of dry bulb temperature and wet bulb depression and found satisfactory.

Another study conducted by [16] proposed four ANNs models to predict the relative humidity and temperature in a swine livestock warehouse located in Puerto Gaitan–Meta. They observed that the models used in the study are suitable for the prediction of humidity in barns not equipped with humidity sensors. However, [17] used an improved backpropagation (BP) neural network for the prediction of indoor relative humidity and temperature every 10 min and 6–72 hours in advance based on a cloud database in Chongqing, China. Both temperature and humidity predictions have a strong correlation with the observed data. Similarly, another study conducted by [18] used BP neural network for the prediction of one day ahead mean air temperature and relative humidity of greenhouse located in the sub-humid sub-tropical regions of India. The results displayed that the BP neural network model provided the best prediction for inside temperature and relative humidity. However, a study done by [19] used daily minimum air temperature ( $T_n$ ) downscaled from INMCM4 general circulation model (GCM) to predict the relative humidity for climate change studies but relative humidity predictions were poor in

few months especially in March, July, August, and October. Moreover, a study conducted by [20] proposed a Functional Link Neural Network (FLNN) which comprises of a single layer of tunable weight trained with the Modified Cuckoo Search algorithm (MCS) for prediction of daily temperature and relative humidity. It was observed that FNN when trained with MCS produced less prediction error. Further, an attempt has been made for the prediction of relative humidity and temperature at different locations inside tobacco dryer by [21] by using a fitting ANN model. Another study performed by [22] also used different ANN models to successfully forecast indoor relative humidity and temperature in the education building of Izmir, Turkey.

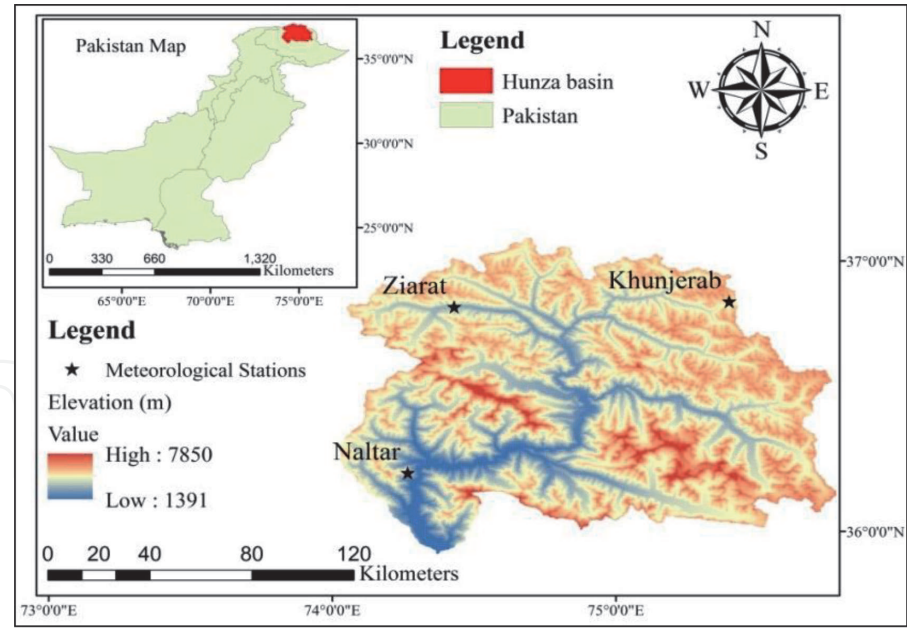
Formerly, no attempt has been made for the prediction of relative humidity in the alpine catchment where there is an issue of data scarcity. The current study is unique because it uses two machine learning models such as MARS and M5T to predict the relative humidity in the Hunza basin (glaciated basin), Pakistan. MARS model was selected because it requires a short training process and has the ability to model complex nonlinear processes deprived of strong model assumptions as compared to ANNs models [23, 24] whereas the M5T model was selected because of its small computation cost and ease in large data treatment as compared to support vector machine (SVM) and ANN [25, 26]. In previous studies, mostly these models were used for the prediction of runoff in poorly gauged basins. A study conducted by [27] suggested that the MARS method is capable of predicting short-term runoff forecast in mountainous watersheds whereas MARS was successfully used for the prediction of streamflows with inadequate data input in the mountainous catchment by [28]. Similarly, the M5T model was found useful in the prediction of streamflows of several tributaries by [29] and it was observed that predictions are good in rainless periods. Another study conducted by [30] found the M5T algorithm reliable in the prediction of streamflows. Several other studies also encouraged the researchers to use MARS and M5T models for the prediction of runoff e.g. [31–37]. Apart from runoff prediction, MARS and M5T models were also used for the prediction of evapotranspiration (ET) and Pan Evaporation (Ep). A study conducted by [38] compared the performance of M5T, MARS along with calibrated Hargreaves-Samani (CHS), MLP, and Stephens-Stewart (SS) models and observed that MARS performed better in the prediction of Ep. Another study conducted by [39] found that the M5T model outperformed compared to Ritchie Equation for the prediction of ET. Similarly, [40] successfully predicted reference evapotranspiration by using M5T and ANN models.

## 2. Study area

Hunza is a glaciated sub-catchment of the Upper Indus Basin (UIB) and is located in the western Karakoram Himalayan region of Pakistan (**Figure 1**). The basin lies within the extent of 74°02'–75°48'E and 35°54'–37°05'N and encompasses 13,671 km<sup>2</sup> of the catchment area.

The elevation of the basin ranged from 1391 to 7850 m. About 20% catchment area of the basin is covered by glaciers [41] and there are 110 glacial lakes in the basin [42]. It is the main tributary of the Indus Basin Irrigation System (IBIS) and it contributes about 12% of UIB streamflows upstream of Tarbela dam [43]. The climate of the Hunza basin is arid to semi-arid and is normally categorized by two seasons, October to March as winter and April to September as summer. The weather conditions vary within the basin. At low altitudes, weather is hot whereas at high altitudes winters are cold and there are extensive variations in temperature extremes [44]. The mean total annual precipitation varies with respect to altitude;





**Figure 1.**  
*Location map of the study area.*

low altitude station such as Naltar (2858 m) receives more precipitation i.e. 660 mm as compared to high altitude station Khunjerab (4730 m) which receives 165 mm of precipitation. The meteorological station installed in between Naltar and Khunjerab (i.e. Ziarat, 3669 m) receives 292 mm of precipitation [45, 46].

The temporal variations in meteorological variables of Khunjerab station (using data of 1995–2009) are displayed in **Table 1**. **Table 1** shows that the maximum temperature varies between  $-11.1^{\circ}\text{C}$  (January) to  $11.6^{\circ}\text{C}$  (July) whereas minimum temperature varies from  $-21.3^{\circ}\text{C}$  (January) to  $1.3^{\circ}\text{C}$  (July). The maximum relative humidity in the basin varies from 59% (March) to 91% (August) while minimum relative humidity varies from 23% (March) to 52% (December). The daily solar radiation in the Hunza basin varies from 2563 (December) to 5148 (May) watt/m<sup>2</sup>.

| Month     | Maximum Temperature (°C) | Minimum Temperature (°C) | Maximum Relative Humidity (%) | Minimum Relative Humidity (%) | Solar Radiation (watt/m <sup>2</sup> ) |
|-----------|--------------------------|--------------------------|-------------------------------|-------------------------------|--|
| January   | -11.1                    | -21.3                    | 62                            | 30                            | 2933                                   |
| February  | -11.0                    | -19.7                    | 77                            | 34                            | 3500                                   |
| March     | -4.5                     | -16.9                    | 59                            | 23                            | 4394                                   |
| April     | 0.2                      | -9.7                     | 78                            | 30                            | 4750                                   |
| May       | 5.5                      | -4.6                     | 81                            | 26                            | 5148                                   |
| June      | 7.7                      | -1.3                     | 87                            | 44                            | 5102                                   |
| July      | 11.6                     | 1.3                      | 86                            | 38                            | 4858                                   |
| August    | 10.5                     | -0.3                     | 91                            | 34                            | 4711                                   |
| September | 4.6                      | -4.5                     | 86                            | 27                            | 4227                                   |
| October   | 0.8                      | -10.4                    | 78                            | 35                            | 4003                                   |
| November  | -5.8                     | -16.1                    | 68                            | 39                            | 3452                                   |
| December  | -11.0                    | -18.9                    | 80                            | 52                            | 2563                                   |

**Table 1.**  
*Monthly average variations in meteorological variables of Khunjerab (1995–2012).*

### 3. Material and methods

#### 3.1 Topography

The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), Global Digital Elevation Model (GDEM) was used to delineate the catchment boundary of the Hunza basin. The Hunza basin was delineated using ASTER GDEM v3 data in Arc GIS. The data was acquired from the website: <https://lpdaac.usgs.gov/tools/data-pool/>. The format of the downloaded tiles was Geo-TIFF and has the gridding resolution i.e. (30 m) and tile structure (1°x 1°).

#### 3.2 Meteorological data

There are four meteorological stations in the Hunza River basin such as Hunza, Naltar, Khunjerab, and Ziarat (**Table 2**). The Hunza meteorological station was installed by the Pakistan Meteorological Department (PMD) and the record is available from 2007 to onward whereas the other three stations were installed and managed by Water and Power Development Authority (WAPDA) and the record is available from 1995 to onward. The current study has employed daily data of temperature, precipitation, solar radiation, and relative humidity of Ziarat, Naltar, and Khunjerab meteorological stations. The required data of the aforementioned stations were acquired from the Surface Water Hydrology Project of the Water and Power Development Authority (SWHP-WAPDA), Pakistan from 1995 to 2009 (**Table 2**).

#### 3.3 Machine learning models

The current study has employed two machine learning models such as M5 Tree and MARS for the prediction of relative humidity at three meteorological stations of the Hunza basin. Their detailed description is given below:

##### 3.3.1 M5 tree model

The M5T model was first introduced by [47]. Model trees simplify the theories of regression trees and there are constant values at their leaves [48]. M5T model is established in relation to a binary decision tree where linear regression functions are placed in the terminal node (leaf) and a relationship is developed between dependent and independent variables through it [49]. Model development involves two stages; the first stage involves in creation of a decision tree by using a split criterion

| Meteorological Station | Latitude (DD) | Longitude (DD) | Elevation (m) | Data                          | Agency     |
|------------------------|---------------|----------------|---------------|-------------------------------|------------|
| Hunza                  | 36.320        | 74.640         | 2374          | $P, T_{max}, T_{min}, RH, SR$ | PMD        |
| Naltar                 | 36.216        | 74.266         | 2858          | —                             | SWHP-WAPDA |
| Khunjerab              | 36.850        | 75.400         | 4730          | —                             | SWHP-WAPDA |
| Ziarat                 | 36.830        | 74.430         | 3669          | —                             | SWHP-WAPDA |

Note: DD = Degree decimal; P= Precipitation;  $T_{max}$ = Maximum temperature;  $T_{min}$ = Minimum temperature; RH = Relative humidity; SR = Solar radiation.

**Table 2.**  
List of meteorological stations in the Hunza basin.

whereas in the second stage overgrown tree is pruned for designing the model tree [25]. The splitting stage in the M5T model is composed of regression function at the leaves instead of class labels and continuous numerical attributes can be estimated through it [36]. The splitting criterion for the M5T model procedure is based on the standard deviation reduction (SDR) function achieved in every node. This criterion points out the error in that node and the minimum expected error is calculated by the model because of testing each attribute in that node [50, 51]. The SDR in the M5T model can be calculated by the following Equation [47]:

$$SDR = sd(M) - \sum \frac{|M_i|}{|M|} sd(M_i) \quad (1)$$

Where  $SDR$  specifies the standard deviation reduction and  $sd$  indicates standard deviation;  $M$  specifies a set of examples that reaches the node; whereas  $M_i$  signifies the subset of examples that have the  $i^{th}$  outcome of the potential set.

Because of the splitting or branching process, data in child nodes (smaller nodes) have less SD than parent nodes (greater nodes). The division process often results in producing a large tree-like structure which causes overfitting and this issue can be resolved by pruning back the tree [52], for instance by substituting a subtheme with a leaf. Pruning the overgrown tree and substitution of subthemes with linear regression functions are performed in the second stage of model designing. This method of producing the model tree separates the parameter space into subspaces and builds in each of them a linear regression model.

### 3.3.2 MARS algorithm

MARS model was first developed by [53]. Its working procedure involved establishing a relationship among a set of input variables and the target-dependent that involve connections with less number of variables [54]. MARS produces flexible models to facilitate the solution space to be divided into several intervals of independent parameters whereas individual splines are fit to each interval [53]. This method is non-parametric and non-linear and it involves a forward-backward procedure to predict a continuous dependent parameter in high-dimensional data [55]. No assumptions have been made about the fundamental functional relationships between independent and dependent variables by the MARS model. In MARS, the splines are connected smoothly together to form piecewise curves which are also known as basis functions (BFs), and these form a flexible model which is capable of handling both linear and non-linear behavior [54]. Two stages are involved in setting up the MARS model which includes forward (constructing the model) and backward (a pruning procedure) stages. In the first stage (forward), to define a pair of BF candidates, knots are placed within the range of each predictor variable. To produce a maximum reduction in sum-of-squares residual error, the model adjusts the knot and its corresponding pair of BF in each step. This process of adding BF lasts and generally a very complex and overfitted model is produced. However, the overfitted model is pruned by deleting the less important redundant BF in the backward stage [54, 55].

The MARS model  $f(X)$  is generally expressed by the following equation;

$$f(x) = \delta_o + \sum_{m=1}^M \delta_m h_m(X) \quad (2)$$

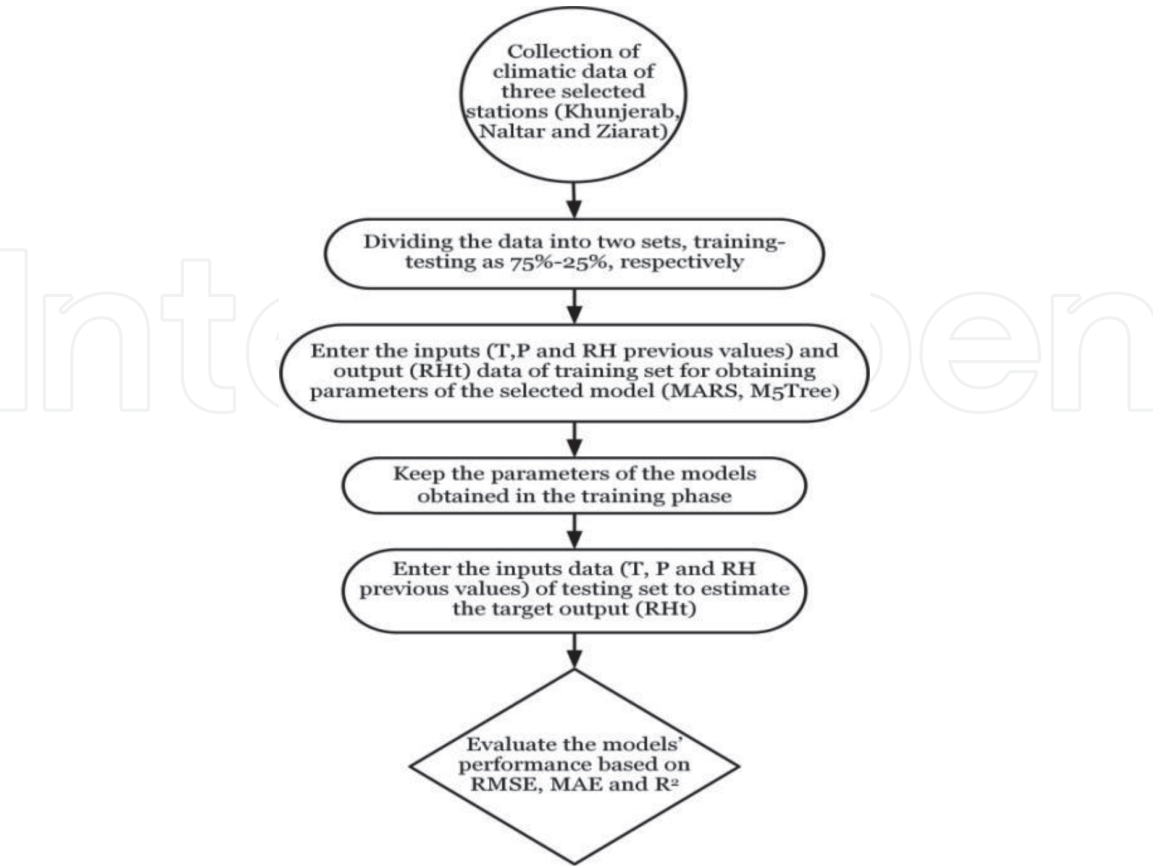
Where  $\delta_o$  and  $\delta_m$  denote the coefficients which are calculated by the least sum of squared errors from splines functions, whereas  $h_m(X)$  represents the spline

functions, and  $M$  denotes the number of functions. The pruning stage improves the forecasting accuracy of the model and  $M$  is determined during this phase [55].

### 3.4 Model setup

The current study compares the accuracy of two machine learning methods such as MARS and M5Tree, for the prediction of daily relative humidity using different input data combinations of precipitation, temperature, and relative humidity. These machine learning models were applied on three meteorological stations such as Khunjerab, Naltar, and Ziarat one by one. The flowchart of the current study is displayed in **Figure 2**. Each model was applied on these stations separately with different input data combinations for the prediction of relative humidity (RH). Ten input data combinations were developed for each meteorological station by each model to decide the best input data combination for the prediction of relative humidity. Initially, three preceding relative humidity (RH) input combinations such as (i) RHt-1, (ii) RHt-1 and RHt-2, and (iii) RHt-1, RHt-2, and RHt-3 were tried to both the models to predict current RH (RHt). After that, three precipitation (i.e. (i) Pt-1, (ii) Pt-1, Pt-2, (iii) Pt-1,Pt-2,Pt-3) and temperature inputs (i.e. (i) Tt-1, (ii) Tt-1,Tt-2, (iii) Tt-1,Tt-2,Tt-3) combinations were separately added to the best RH combination whereas in the last input combination (10th); best temperature and precipitation inputs were added together with the best RH input combination to see the combine effect of both parameters on model's accuracy in predicting relative humidity.

The current analysis involves daily data of precipitation, temperature, and relative humidity from 1995 to 2009. About 75% of input data i.e. from 1995 to 2006 was used for training whereas 25% of input data i.e. from 2007 to 2009 was used for testing in both machine learning models for prediction of relative humidity.



**Figure 2.**  
Flowchart of the study.



However, [8] used only two-year data i.e. 2008 to 2009 for training the LSTM model which might not be enough for reliable predictions.

### 3.5 Models evaluation criteria

The models' accuracy in relative humidity prediction against observed data was evaluated using the following statistics which are normally used in the related literature. The statistics include  $R^2$ , RMSE, and MAE as shown in Eqs. (3)-(5).

$$R^2 = 1 - \frac{\frac{1}{n} \sum_{i=1}^n (RH_i - \overline{RH})^2}{\frac{1}{n} \sum_{i=1}^n (rh_i - \overline{rh})^2} \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (RH_{iO} - RH_{iM})^2}{N}} \quad (4)$$

$$MAE = \frac{\sum_{i=1}^N |RH_{iO} - RH_{iM}|}{N} \quad (5)$$

Where  $\overline{rh}$  indicates the observed mean relative humidity;  $\overline{RH}$  is the mean of the predicted relative humidity  $RH_i$ ;  $N$  signifies the number of data points. Moreover,  $RH_{iO}$  is observed relative humidity and  $RH_{iM}$  is modeled relative humidity. Previous studies such as [56–61] suggested that a single statistical indicator cannot examine well the prediction accuracy of soft computing models. Therefore, the current study used three statistical indicators to judge the model prediction accuracy with confidence. When the error distributions of the models are normal and uniform in that case the use of error statistics such as RMSE and MAE is more suitable. For an ideal model, the values of RMSE and MAE should equal to 0, whereas,  $R^2$  should equal to 1. The model having relatively small values of MAE and RMSE as compared to other models is considered the best model.

## 4. Results and discussions

### 4.1 Performance evaluation of MARS model in predicting relative humidity

The performance evaluation statistics of the MARS model for the prediction of relative humidity at Khunjerab, Naltar, and Ziarat are presented in **Tables 3–5**, respectively. The MARS model performed excellent for the prediction of relative humidity at all meteorological stations both during training and testing processes especially, it provided the best predictions for the 6th scenario (S6) of input data combination which is highlighted in bold. The RMSE, MAE, and  $R^2$  values during the training (5.58%, 4.51%, 0.852) and testing (5.98%, 5.43%, 0.808) stages for Khunjerab meteorological station are displayed in **Table 3**. The MARS model performed better during training as compared to testing at Khunjerab. However, the MARS model did not perform well for the S1, S2, and S3 scenarios. Our study results were found better than the study conducted by [1]. They described that GEP and ANNs models can predict relative humidity reliably at two Californian stations (RMSE= 10.7%, MAE= 7.6% and  $R^2 = 0.73$ ) during training; and (RMSE= 10.1%, MAE= 7.5% and  $R^2 = 0.714$ ) during testing stage in the case of GEP model. However, ANN model produced better results as compared to GEP such as (RMSE= 7.8%, MAE= 3.6% and  $R^2 = 0.826$ ) during training, and (RMSE= 8.2%, MAE= 4.1% and  $R^2 = 0.751$ ) during testing stage.

| Scenario  | Input Combinations                                      | Training    |             |                | Testing     |             |                |
|-----------|---|-------------|-------------|----------------|-------------|-------------|----------------|
|           |   | RMSE (%)    | MAE (%)     | R <sup>2</sup> | RMSE (%)    | MAE (%)     | R <sup>2</sup> |
| S1        | RHt-1   | 11.00       | 8.49        | 0.480          | 13.59       | 10.51       | 0.381          |
| S2        | RHt-1, RHt-2  | 10.88       | 8.39        | 0.491          | 13.58       | 10.51       | 0.385          |
| S3        | RHt-1, RHt-2, RHt-3                                     | 10.86       | 8.36        | 0.493          | 13.53       | 10.45       | 0.388          |
| S4        | RHt-1, RHt-2, RHt-3, Tt-1                               | 5.78        | 4.64        | 0.823          | 6.43        | 5.62        | 0.782          |
| S5        | RHt-1, RHt-2, RHt-3, Tt-1, Tt-2                         | 5.64        | 4.57        | 0.831          | 6.12        | 5.57        | 0.801          |
| <b>S6</b> | <b>RHt-1, RHt-2, RHt-3, Tt-1, Tt-2, Tt-3</b>            | <b>5.58</b> | <b>4.51</b> | <b>0.852</b>   | <b>5.98</b> | <b>5.43</b> | <b>0.808</b>   |
| S7        | RHt-1, RHt-2, RHt-3, Pt-1                               | 9.76        | 7.62        | 0.602          | 11.67       | 8.75        | 0.532          |
| S8        | RHt-1, RHt-2, RHt-3, Pt-1, Pt-2                         | 9.16        | 7.28        | 0.643          | 10.87       | 8.38        | 0.544          |
| S9        | RHt-1, RHt-2, RHt-3, Pt-1, Pt-2, Pt-3                   | 9.08        | 7.19        | 0.649          | 10.73       | 8.29        | 0.552          |
| S10       | RHt-1, RHt-2, RHt-3, Pt-1, Pt-2, Pt-3, Tt-1, Tt-2, Tt-3 | 6.21        | 5.13        | 0.802          | 6.03        | 5.53        | 0.803          |

*Bold values represent the best input data combination.*

**Table 3.**  
*The statistical evaluation of the MARS model at Khunjerab.*

| Scenario  | Input Combinations                                      | Training    |             |                | Testing     |             |                |
|-----------|---|-------------|-------------|----------------|-------------|-------------|----------------|
|           |   | RMSE (%)    | MAE (%)     | R <sup>2</sup> | RMSE (%)    | MAE (%)     | R <sup>2</sup> |
| S1        | RHt-1   | 11.42       | 8.95        | 0.588          | 14.22       | 10.69       | 0.499          |
| S2        | RHt-1, RHt-2  | 11.09       | 8.65        | 0.612          | 13.99       | 10.31       | 0.518          |
| S3        | RHt-1, RHt-2, RHt-3                                     | 11.02       | 8.61        | 0.616          | 13.98       | 10.32       | 0.517          |
| S4        | RHt-1, RHt-2, RHt-3, Tt-1                               | 5.84        | 4.73        | 0.812          | 6.84        | 5.34        | 0.783          |
| S5        | RHt-1, RHt-2, RHt-3, Tt-1, Tt-2                         | 5.76        | 4.62        | 0.818          | 6.73        | 5.25        | 0.792          |
| <b>S6</b> | <b>RHt-1, RHt-2, RHt-3, Tt-1, Tt-2, Tt-3</b>            | <b>5.63</b> | <b>4.53</b> | <b>0.826</b>   | <b>6.58</b> | <b>5.08</b> | <b>0.806</b>   |
| S7        | RHt-1, RHt-2, RHt-3, Pt-1                               | 10.24       | 7.63        | 0.673          | 11.73       | 8.36        | 0.624          |
| S8        | RHt-1, RHt-2, RHt-3, Pt-1, Pt-2                         | 10.08       | 7.46        | 0.692          | 11.29       | 8.07        | 0.645          |
| S9        | RHt-1, RHt-2, RHt-3, Pt-1, Pt-2, Pt-3                   | 9.36        | 7.13        | 0.724          | 10.76       | 7.93        | 0.663          |
| S10       | RHt-1, RHt-2, RHt-3, Pt-1, Pt-2, Pt-3, Tt-1, Tt-2, Tt-3 | 5.71        | 4.71        | 0.815          | 6.74        | 5.18        | 0.796          |

*Bold values represent the best input data combination.*

**Table 4.**  
*The statistical evaluation of the MARS model at Naltar.*

Similarly, the MARS model provided the best prediction of relative humidity for the S6 input data scenario at Naltar both during training and testing stages as shown in **Table 4**. The RMSE, MAE and R<sup>2</sup> values for the best input parameter combination were 5.63%, 4.53%, and 0.826 respectively, during training whereas 6.58%, 5.08%, and 0.806, were during testing (**Table 4**). The MARS model did not perform well for S1, S2, and S3 input combinations. However, a study conducted by [5] observed that the XGBoost model provided the best prediction of relative humidity (MAE= 2.29%) as compared to SARIMA (MAE= 2.97%) and HW additive (MAE= 2.74%).

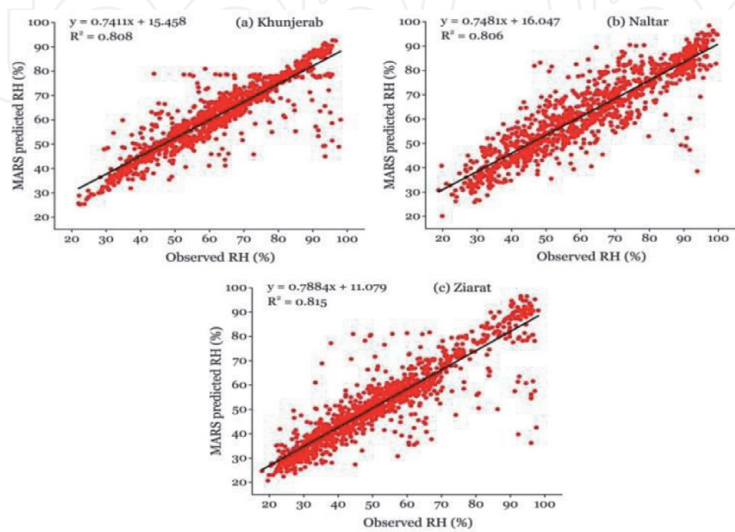
| Scenario  | Input Combinations                                      | Training    |             |                | Testing     |             |                |
|-----------|---|-------------|-------------|----------------|-------------|-------------|----------------|
|           |   | RMSE (%)    | MAE (%)     | R <sup>2</sup> | RMSE (%)    | MAE (%)     | R <sup>2</sup> |
| S1        | RHt-1   | 11.29       | 8.58        | 0.530          | 14.75       | 10.90       | 0.420          |
| S2        | RHt-1, RHt-2  | 11.17       | 8.49        | 0.540          | 14.73       | 10.86       | 0.424          |
| S3        | RHt-1, RHt-2, RHt-3                                     | 11.12       | 8.46        | 0.544          | 14.77       | 10.89       | 0.421          |
| S4        | RHt-1, RHt-2, RHt-3, Tt-1                               | 5.73        | 4.92        | 0.801          | 6.13        | 5.23        | 0.792          |
| S5        | RHt-1, RHt-2, RHt-3, Tt-1, Tt-2                         | 5.58        | 4.76        | 0.813          | 6.02        | 5.06        | 0.807          |
| <b>S6</b> | <b>RHt-1, RHt-2, RHt-3, Tt-1, Tt-2, Tt-3</b>            | <b>5.26</b> | <b>4.59</b> | <b>0.833</b>   | <b>5.86</b> | <b>4.97</b> | <b>0.815</b>   |
| S7        | RHt-1, RHt-2, RHt-3, Pt-1                               | 10.03       | 7.13        | 0.624          | 12.75       | 8.34        | 0.542          |
| S8        | RHt-1, RHt-2, RHt-3, Pt-1, Pt-2                         | 9.75        | 6.48        | 0.687          | 11.78       | 8.07        | 0.568          |
| S9        | RHt-1, RHt-2, RHt-3, Pt-1, Pt-2, Pt-3                   | 9.48        | 6.29        | 0.698          | 11.38       | 7.84        | 0.597          |
| S10       | RHt-1, RHt-2, RHt-3, Pt-1, Pt-2, Pt-3, Tt-1, Tt-2, Tt-3 | 5.38        | 4.68        | 0.820          | 5.94        | 5.02        | 0.812          |

*Bold values represent the best input data combination.*

**Table 5.**  
The statistical evaluation of the MARS model at Ziarat.

However, the MARS model performed the best (RMSE= 5.86, MAE= 4.97%, R<sup>2</sup> = 0.815) for prediction of relative humidity at Ziarat for the S6 input combination during the testing stage as shown in **Table 5**. The MARS model also performed fairly well during training stage (RMSE= 5.26%, MAE= 4.59%, R<sup>2</sup> = 0.833) for S6 input combination. The MARS model provided a poor prediction of relative humidity for S1, S2, and S3 input scenarios (**Table 5**). Overall, the MARS model performed fairly well at Khunjerab (R<sup>2</sup>= 0.852) and showed slightly low performance at Naltar (R<sup>2</sup> =0.826) for the S6 input combination during the training stage (**Tables 3–5**).

The MARS model performance was also evaluated by drawing scatter plots. The scatter plots had been drawn between observed and predicted relative humidity from 2007 to 2009 on daily data as displayed in **Figure 3**. Scatter plots also displayed that the MARS model outperformed for prediction of relative humidity at



**Figure 3.**  
Scatter plots between observed and predicted relative humidity by using MARS model at (a) Khunjerab; (b) Naltar and (c) Ziarat.

all meteorological stations, especially, at Ziarat with  $R^2 = 0.815$  for the S6 input combination during the testing stage (**Figure 3**).

4.2 Performance evaluation of M5T model in predicting relative humidity

The performance evaluation of the M5T model for the prediction of relative humidity at Khunjerab, Naltar, and Ziarat is displayed in **Tables 6–8**, respectively. The M5T model also performed well for the prediction of relative humidity at all

| Scenario  | Input Combinations                                 | Training    |             |              | Testing     |             |              |
|-----------|--|-------------|-------------|--------------|-------------|-------------|--------------|
|           |  | RMSE (%)    | MAE (%)     | $R^2$        | RMSE (%)    | MAE (%)     | $R^2$        |
| S1        | RHt-1  | 11.08       | 8.56        | 0.476        | 13.64       | 10.54       | 0.378        |
| S2        | RHt-1, RHt-2                                       | 10.94       | 8.45        | 0.486        | 13.61       | 10.53       | 0.382        |
| S3        | RHt-1, RHt-2, RHt-3                                | 10.90       | 8.41        | 0.491        | 13.56       | 10.49       | 0.391        |
| S4        | RHt-1, RHt-2, RHt-3,Tt-1                           | 6.71        | 5.76        | 0.758        | 6.94        | 6.32        | 0.726        |
| S5        | RHt-1, RHt-2, RHt-3,Tt-1,Tt-2                      | 6.62        | 5.64        | 0.765        | 6.89        | 6.13        | 0.748        |
| <b>S6</b> | <b>RHt-1, RHt-2, RHt-3,Tt-1,Tt-2,Tt-3</b>          | <b>5.94</b> | <b>5.08</b> | <b>0.796</b> | <b>6.14</b> | <b>5.56</b> | <b>0.772</b> |
| S7        | RHt-1, RHt-2, RHt-3,Pt-1                           | 9.79        | 7.68        | 0.598        | 11.77       | 8.82        | 0.529        |
| S8        | RHt-1, RHt-2, RHt-3,Pt-1,Pt-2                      | 9.20        | 7.32        | 0.639        | 10.92       | 8.44        | 0.541        |
| S9        | RHt-1, RHt-2, RHt-3,Pt-1,Pt-2,Pt-3                 | 9.13        | 7.24        | 0.642        | 10.83       | 8.36        | 0.548        |
| S10       | RHt-1, RHt-2, RHt-3,Pt-1,Pt-2,Pt-3, Tt-1,Tt-2,Tt-3 | 6.43        | 5.32        | 0.772        | 6.23        | 5.81        | 0.752        |

*Bold values represent the best input data combination.*

**Table 6.**  
The statistical evaluation of the M5T model at Khunjerab.

| Scenario  | Input Combinations                                 | Training    |             |              | Testing     |             |              |
|-----------|--|-------------|-------------|--------------|-------------|-------------|--------------|
|           |  | RMSE (%)    | MAE (%)     | $R^2$        | RMSE (%)    | MAE (%)     | $R^2$        |
| S1        | RHt-1  | 11.12       | 8.57        | 0.476        | 13.68       | 10.58       | 0.378        |
| S2        | RHt-1, RHt-2                                       | 10.92       | 8.45        | 0.486        | 13.62       | 10.54       | 0.381        |
| S3        | RHt-1, RHt-2, RHt-3                                | 10.89       | 8.38        | 0.491        | 13.57       | 10.48       | 0.384        |
| S4        | RHt-1, RHt-2, RHt-3,Tt-1                           | 6.76        | 5.79        | 0.752        | 6.90        | 6.36        | 0.721        |
| S5        | RHt-1, RHt-2, RHt-3,Tt-1,Tt-2                      | 6.58        | 5.61        | 0.760        | 6.81        | 6.18        | 0.742        |
| <b>S6</b> | <b>RHt-1, RHt-2, RHt-3,Tt-1,Tt-2,Tt-3</b>          | <b>5.82</b> | <b>5.12</b> | <b>0.791</b> | <b>6.19</b> | <b>5.58</b> | <b>0.762</b> |
| S7        | RHt-1, RHt-2, RHt-3,Pt-1                           | 9.84        | 7.67        | 0.598        | 11.76       | 8.79        | 0.529        |
| S8        | RHt-1, RHt-2, RHt-3,Pt-1,Pt-2                      | 9.28        | 7.38        | 0.638        | 10.96       | 8.48        | 0.541        |
| S9        | RHt-1, RHt-2, RHt-3,Pt-1,Pt-2,Pt-3                 | 9.11        | 7.23        | 0.646        | 10.77       | 8.32        | 0.550        |
| S10       | RHt-1, RHt-2, RHt-3,Pt-1,Pt-2,Pt-3, Tt-1,Tt-2,Tt-3 | 6.52        | 5.46        | 0.767        | 6.37        | 5.94        | 0.758        |

*Bold values represent the best input data combination.*

**Table 7.**  
The statistical evaluation of the M5T model at Naltar.



| Scenario  | Input Combinations                                      | Training    |             |                | Testing     |             |                |
|-----------|---|-------------|-------------|----------------|-------------|-------------|----------------|
|           |   | RMSE (%)    | MAE (%)     | R <sup>2</sup> | RMSE (%)    | MAE (%)     | R <sup>2</sup> |
| S1        | RHt-1   | 11.12       | 8.54        | 0.476          | 13.64       | 10.56       | 0.378          |
| S2        | RHt-1, RHt-2  | 10.95       | 8.43        | 0.487          | 13.61       | 10.54       | 0.381          |
| S3        | RHt-1, RHt-2, RHt-3                                     | 10.92       | 8.39        | 0.491          | 13.58       | 10.49       | 0.384          |
| S4        | RHt-1, RHt-2, RHt-3, Tt-1                               | 6.67        | 5.70        | 0.758          | 6.82        | 6.27        | 0.728          |
| S5        | RHt-1, RHt-2, RHt-3, Tt-1, Tt-2                         | 6.47        | 5.52        | 0.764          | 6.72        | 6.10        | 0.752          |
| <b>S6</b> | <b>RHt-1, RHt-2, RHt-3, Tt-1, Tt-2, Tt-3</b>            | <b>5.74</b> | <b>5.04</b> | <b>0.796</b>   | <b>6.08</b> | <b>5.46</b> | <b>0.783</b>   |
| S7        | RHt-1, RHt-2, RHt-3, Pt-1                               | 9.79        | 7.66        | 0.599          | 11.72       | 8.78        | 0.530          |
| S8        | RHt-1, RHt-2, RHt-3, Pt-1, Pt-2                         | 9.20        | 7.32        | 0.640          | 10.90       | 8.42        | 0.541          |
| S9        | RHt-1, RHt-2, RHt-3, Pt-1, Pt-2, Pt-3                   | 9.12        | 7.21        | 0.645          | 10.78       | 8.32        | 0.550          |
| S10       | RHt-1, RHt-2, RHt-3, Pt-1, Pt-2, Pt-3, Tt-1, Tt-2, Tt-3 | 6.26        | 5.16        | 0.800          | 6.08        | 5.58        | 0.778          |

*Bold values represent the best input data combination.*

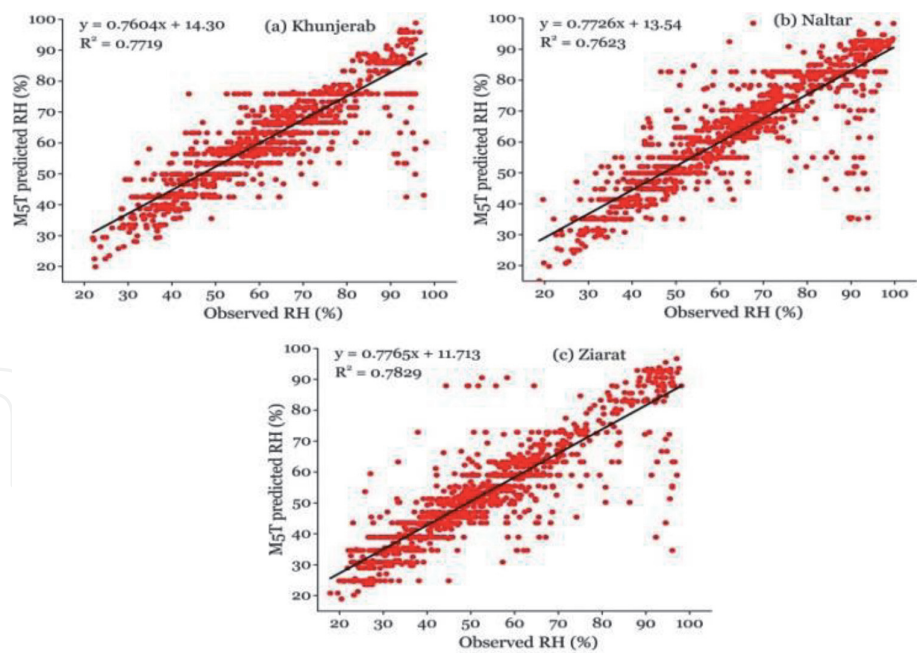
**Table 8.**  
*The statistical evaluation of the M5T model at Ziarat.*

meteorological stations both during training and testing stages; however, it provided the best predictions of relative humidity for the 6th input data combination (S6) at all stations which are highlighted in bold. Overall, the M5T model performance was slightly lower as compared to MARS. The M5T model also performed better during training as compared to testing at all meteorological stations. However, the M5T model provided the best prediction of relative humidity at Ziarat as compared to Naltar and Khunjerab (**Table 8**). However, the M5T model did not perform well for the prediction of relative humidity for the S1, S2, and S3 scenarios with  $R^2 < 0.50$  at all meteorological stations (**Tables 6–8**). A previous study conducted by [8] observed that the LSTM model is capable of forecasting complex univariate relative humidity time series. On contrary, [3] suggested that ARIMA can provide a better prediction of relative humidity as compared to LSTM.

At Khunjerab station, the M5T model performed well (RMSE= 5.94%, MAE = 5.08%,  $R^2$ = 0.796) in case of S6 input combination during model training stage whereas it displayed low prediction performance (RMSE= 6.14%, MAE= 5.56%,  $R^2$ = 0.772) during testing stage as shown in **Table 6**. Similarly, the M5T model did not perform well for the S1, S2, and S3 scenarios ( $R^2 < 0.50$ ). Similarly, at Naltar station, the M5T model performed reasonably well (RMSE= 5.82%, MAE= 5.12%,  $R^2$ = 0.791) for S6 input combination during training stage whereas it exhibited a slightly low performance (RMSE= 6.19%, MAE= 5.58%,  $R^2$ = 0.762) during testing stage as presented in **Table 7**.

However, the M5T model provided the best prediction of relative humidity at the Ziarat station for the S6 input combination (**Table 8**). The M5T model performed better during training (RMSE= 5.74%, MAE= 5.04%,  $R^2$ = 0.796) as compared to testing (RMSE= 6.08%, MAE= 5.46%,  $R^2$ = 0.783) stage as displayed in **Table 8**.

The M5T model performance was also evaluated by drawing scatter plots. The scatter plots were drawn between observed and predicted relative humidity from 2007 to 2009 on daily data as displayed in **Figure 4**. Scatter plots showed that, the M5T model can also predict relative humidity fairly well at all meteorological stations, especially, at Ziarat ( $R^2$ = 0.782) for the S6 input combination during the testing stage (**Figure 4**).

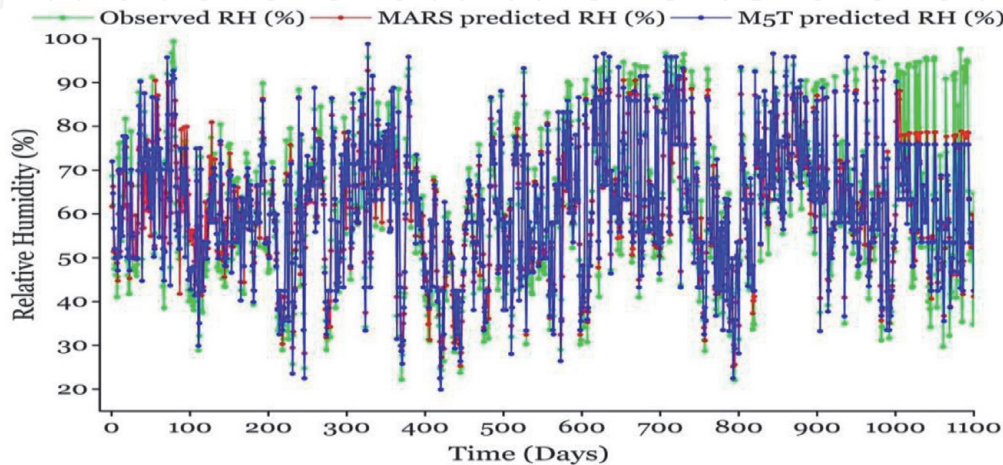


**Figure 4.**  
Scatter plots between observed and predicted relative humidity by using the M5T model at (a) Khunjerab; (b) Naltar and (c) Ziarat.

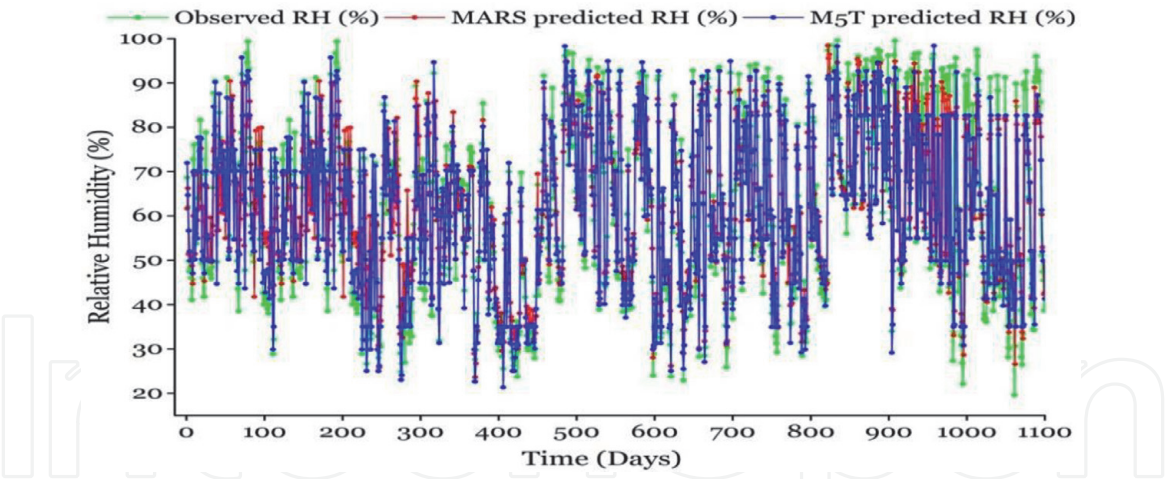
**4.3 Time variations of the observed and predicted relative humidity by MARS and M5T models**

Time variations of the observed and predicted relative humidity by MARS and M5T model at Khunjerab, Naltar, and Ziarat meteorological stations are displayed in **Figures 5–7**. Time variations plots have been drawn by using the best-predicted data of relative humidity (i.e. S6 scenario). The daily data has been drawn from 2007 to 2009. **Figure 5** showed that both the models captured time-series variations of predicted relative humidity very well with reference to observed data at Khunjerab station but slightly underestimated the values from 900 to 1100 days. Moreover, these models slightly underestimated the low and high values of predicted relative humidity with reference to observed data at few points throughout the time series. Overall, the MARS model performed better as compared to M5T for the prediction of daily relative humidity data at Khunjerab.

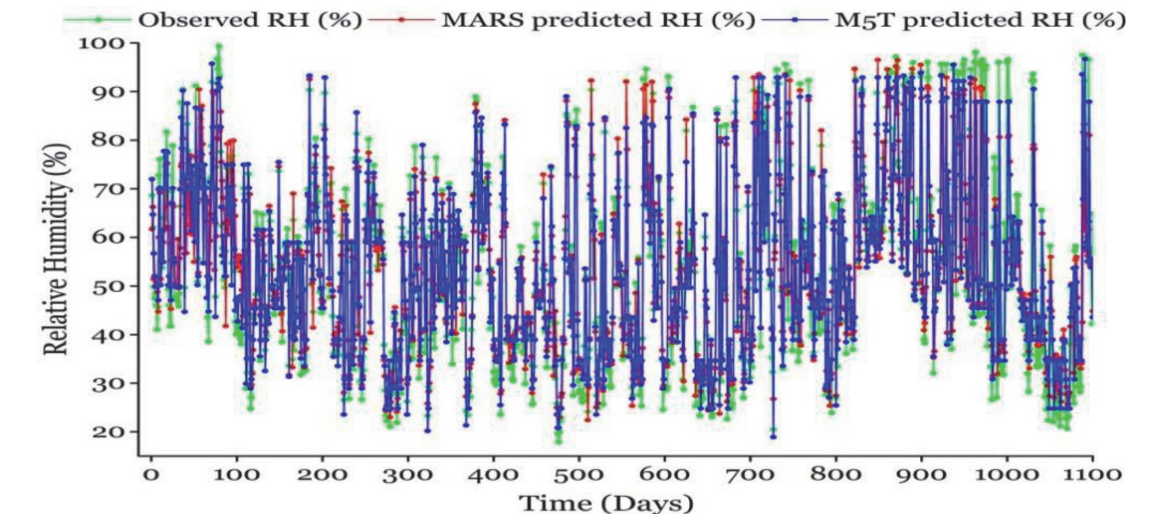
The MARS and M5T models also captured time-series variation of relative humidity superbly with respect to observed data at Naltar station for the S6 input



**Figure 5.**  
Time variation of the observed and predicted relative humidity by MARS and M5T model at Khunjerab station.



**Figure 6.**  
*Time variation of the observed and predicted relative humidity by MARS and M5T model at Naltar station.*



**Figure 7.**  
*Time variation of the observed and predicted relative humidity by MARS and M5T model at Ziarat station.*

combination as displayed in **Figure 6**. Both the models slightly underestimated the predicted relative humidity from 850 days to 1100. Moreover, these models slightly underestimated the predictions of low and high values of relative humidity at some points throughout the study period. Overall, the MARS model provided better predictions of relative humidity as compared to M5T at Naltar (**Figure 6**).

However, both the machine learning models provided the best prediction of relative humidity at Ziarat which is a mid-altitude meteorological station as shown in **Figure 7**. Both the models captured the temporal variations of relative humidity very well throughout the period with reference to observed data for the S6 input combination. Furthermore, the models underestimated the low and high values of predicted relative humidity with reference to observed data. The MARS model predicted low and high values of relative humidity fairly well but it slightly underestimated the values at few points throughout the study period. Overall, the MARS model provided better predictions of relative humidity as compared to M5T at Ziarat (**Figure 7**).

5. Conclusions

Relative humidity has an important impact on plant growth, human health, industry, weather, and climate. Any change in temperature and relative humidity



may result in droughts, heatwaves, floods, and hurricanes. Thus the relative humidity is one of the important factors to measure environmental changes. Keeping in view the importance of relative humidity, the current study has attempted to predict the relative humidity in a high elevated alpine basin (Hunza) of western Karakoram by using the MARS and M5T machine learning models. The current study is novel in that respect that previously nobody tried to predict the relative humidity in a high elevation alpine basin.

Statistical analysis of the model outputs suggested that both the models produced reliable predictions of relative humidity at Khunjerab, Naltar, and Ziarat meteorological stations of the Hunza basin during both training and testing stages. Out of 10 input data combinations of temperature, precipitation, and relative humidity, the 6th combination (i.e. RHt-1, RHt-2, RHt-3, Tt-1, Tt-2, Tt-3) produced the best results for each station by each model. The statistical indicators confirmed the excellent performance of both the models at all stations. For the MARS model, RMSE, MAE, and  $R^2$  values ranged from 5.26–5.63%, 4.51–4.59%, and 0.826–0.856, respectively, during the training stage while they ranged from 5.86–6.58%, 4.97–5.43%, and 0.806–0.815, respectively, during the testing stage. However, in the case of the M5T model, the RMSE, MAE, and  $R^2$  values ranged from 5.74–5.94%, 5.04–2.12%, and 0.791–0.796, respectively, during the training stage whereas the values ranged from 6.08–6.19%, 5.46–5.58%, and 0.762–0.783, respectively, during the testing stage of M5T model. Both the models showed poor performance such as ( $R^2 < 0.50$ ) in the case of S1, S2, and S3 input combinations at all stations. Moreover, it was observed that both the models performed better in training as compared to the testing stage. Both the models outperformed at Ziarat as compared to other stations. Overall, the MARS model performed better than M5T at all stations. The current study is important and it will provide a baseline for future studies to predict the other meteorological variables such as temperature, wind speed, solar radiation, and evapotranspiration by using machine learning tools in high altitude and remote basins which face the issue of data scarcity.

## Acknowledgements

This study is supported by the Second Tibetan Plateau Scientific Expedition and Research Program (STEP, Grant No. 2019QZKK0208), and the Research Fund for Introducing Talents of Yunnan University (No. YJRC3201702). We are thankful to the Surface Water Hydrology Project of the Water and Power Development Authority (SWHP-WAPDA) of Pakistan for providing the required meteorological data to conduct this study.

## Conflict of interest

The authors declare no conflict of interest.



## Author details

Muhammad Adnan<sup>1</sup>, Rana Muhammad Adnan<sup>2</sup>, Shiyin Liu<sup>1\*</sup>,  
Muhammad Saifullah<sup>3</sup>, Yasir Latif<sup>4</sup> and Mudassar Iqbal<sup>5</sup>

1 Institute of International Rivers and Eco-Security, Yunnan University,  
Kunming, China

2 State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering,  
College of Hydrology and Water Resources, Hohai University, Nanjing, China


3 Department of Agriculture Engineering, Muhammad Nawaz Sharif University of  
Agriculture, Multan, Pakistan

4 Key Laboratory of Tibetan Environment Changes and Land Surface Processes,  
Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China

5 Centre of Excellence in Water Resources Engineering, University of Engineering  
and Technology, Lahore, Pakistan

\*Address all correspondence to: shiyin.liu@ynu.edu.cn

## IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Khatibi R, Naghipour L, Ghorbani MA, Aalami MT. Predictability of relative humidity by two artificial intelligence techniques using noisy data from two Californian gauging stations. *Neural Computing and Applications*. 2013; **23**(7):2241-52.
- [2] Ghadiri M, Marjani A, Mohammadinia S, Shokri M. Machine Learning Approaches for Accurate Prediction of Relative Humidity based on Temperature and Wet-Bulb Depression. 2020.
- [3] Li Z, Zou H, Qi B. Application of ARIMA and LSTM in Relative Humidity Prediction. In 2019 IEEE 19th International Conference on Communication Technology (ICCT). 2019; **16**:1544-1549.
- [4] Azzouni A, Pujolle G. A long short-term memory recurrent neural network framework for network traffic matrix prediction. *arXiv preprint*. 2017; arXiv: 1705.05690.
- [5] Li H, Yang Y, Cheng Y, Jin Y, Luo H, Zhang L. Application of Time Series Model in Relative Humidity Prediction. In *Journal of Physics: Conference Series*. 2020; **1584** (1):012017.
- [6] Gangishetty MK, Scott RW, Kelly TL. Effect of relative humidity on crystal growth, device performance and hysteresis in planar heterojunction perovskite solar cells. *Nanoscale*. 2016; **8** (12):6300-7.
- [7] Yu X. Indication of relative humidity of ECMWF in precipitation forecast in Hainan Prefecture. *Qinghai Meteorology*. 2009; **3**:17-20.
- [8] Hutapea MI, Pratiwi YY, Sarkis IM, Jaya IK, Sinambela M. Prediction of relative humidity based on long short-term memory network. In *AIP Conference Proceedings*. 2020. **2221** (1): 060003.
- [9] Quansah E, Amekudzi LK, Preko K. The influence of temperature and relative humidity on indoor ozone concentrations during the Harmattan. *Journal of Emerging Trends in Engineering and Applied Sciences*. 2012; **3**(5):863-7.
- [10] Ohno H, Ohata T, Higuchi K. The influence of humidity on the ablation of continental-type glaciers. *Annals of Glaciology*. 1992; **16**:107-14.
- [11] Hastenrath S. Recession of tropical glaciers. *Science*. 1994; **265**(5180): 1790-1.
- [12] Kaser G, Hardy DR, Mölg T, Bradley RS, Hyera TM. Modern glacier retreat on Kilimanjaro as evidence of climate change: observations and facts. *International Journal of Climatology: A Journal of the Royal Meteorological Society*. 2004; **24**(3):329-39.
- [13] Mölg T, Hardy DR. Ablation and associated energy balance of a horizontal glacier surface on Kilimanjaro. *Journal of Geophysical Research: Atmospheres*. 2004; **109**(D16).
- [14] Rupper S, Roe G. Glacier changes and regional climate: A mass and energy balance approach. *Journal of Climate*. 2008; **21**(20):5384-401.
- [15] Farhat N. Effect of relative humidity on evaporation rates in Nabatieh region. *Lebanese Science Journal*. 2018; **19**(1):59.
- [16] Molano-Jimenez A, Orjuela-Cañón AD, Acosta-Burbano W. Temperature and Relative Humidity Prediction in Swine Livestock Buildings. In 2018 IEEE Latin American Conference on Computational Intelligence (LA-CCI). 2018; 1-4.
- [17] Shi X, Lu W, Zhao Y, Qin P. Prediction of indoor temperature and relative humidity based on cloud

database by using an improved BP neural network in Chongqing. *IEEE Access*. 2018; **6**:30559-66.

[18] Singh VK, Tiwari KN. Prediction of greenhouse micro-climate using artificial neural network. *Applied Ecology and Environmental Research*. 2017; **15**(1):767-78.

[19] Gunawardhana LN, Al-Rawas GA, Kazama S. An alternative method for predicting relative humidity for climate change studies. *Meteorological Applications*. 2017; **24**(4):551-9.

[20] Bakar SZ, Ghazali RB, Ismail LH. Implementation of modified cuckoo search algorithm on functional link neural network for temperature and relative humidity prediction. In *Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013)*. 2014; 151-158.

[21] Martínez-Martínez V, Baladrón C, Gomez-Gil J, Ruiz-Ruiz G, Navas-Gracia LM, Aguiar JM, Carro B. Temperature and relative humidity estimation and prediction in the tobacco drying process using artificial neural networks. *Sensors*; **12**(10):14004-21.

[22] Özbalta TG, Sezer A, Yıldız Y. Models for prediction of daily mean indoor temperature and relative humidity: education building in Izmir, Turkey. *Indoor and Built Environment*. 2012; **21**(6):772-81.

[23] Chou SM, Lee TS, Shao YE, Chen IF. Mining the breast cancer pattern using artificial neural networks and multivariate adaptive regression splines. *Expert Systems with Applications*. 200; **27**(1):133-42.

[24] Adnan RM, Petroselli A, Heddami S, Santos CA, Kisi O. Short term rainfall-runoff modelling using several machine learning methods and a conceptual event-based model. *Stochastic*

*Environmental Research and Risk Assessment*. 2021; **35**(3):597-616.

[25] Singh KK, Pal M, Singh VP. Estimation of mean annual flood in Indian catchments using back propagation neural network and M5 model tree. *Water Resources Management*. 2010; **24**(10):2007-19.

[26] Adnan RM, Liang Z, Trajkovic S, Zounemat-Kermani M, Li B, Kisi O. Daily streamflow prediction using optimally pruned extreme learning machine. *Journal of Hydrology*. 2019; **577**:123981.

[27] Sharda VN, Prasher SO, Patel RM, Ojasvi PR, Prakash C. Performance of Multivariate Adaptive Regression Splines (MARS) in predicting runoff in mid-Himalayan micro-watersheds with limited data/Performances de régressions par splines multiples et adaptives (MARS) pour la prévision d'écoulement au sein de micro-bassins versants Himalayens d'altitudes intermédiaires avec peu de données. *Hydrological Sciences Journal*. 2008; **53** (6):1165-75.

[28] Adamowski J, Chan HF, Prasher SO, Sharda VN. Comparison of multivariate adaptive regression splines with coupled wavelet transform artificial neural networks for runoff forecasting in Himalayan micro-watersheds with limited data. *Journal of Hydroinformatics*. 2012; **14**(3):731-44.

[29] Štravs L, Brilly M. Development of a low-flow forecasting model using the M5 machine learning method. *Hydrological Sciences Journal*. 2007; **52** (3):466-77.

[30] Sattari MT, Pal M, Apaydin H, Ozturk F. M5 model tree application in daily river flow forecasting in Sohu Stream, Turkey. *Water Resources*. 2013; **40**(3):233-42.

[31] Yaseen ZM, Kisi O, Demir V. Enhancing long-term streamflow

forecasting and predicting using periodicity data component: application of artificial intelligence. *Water Resources Management*. 2016(b); **30** (12):4125-51.

[32] Yin Z, Feng Q, Wen X, Deo RC, Yang L, Si J, He Z. Design and evaluation of SVR, MARS and M5Tree models for 1, 2 and 3-day lead time forecasting of river flow data in a semiarid mountainous catchment. *Stochastic Environmental Research and Risk Assessment*. 2018; **32**(9):2457-76.

[33] Nourani V, Davanlou Tajbakhsh A, Molajou A, Gokcekus H. Hybrid wavelet-M5 model tree for rainfall-runoff modeling. *Journal of Hydrologic Engineering*. 2019; **24**(5):04019012.

[34] Al-Sudani ZA, Salih SQ, Yaseen ZM. Development of multivariate adaptive regression spline integrated with differential evolution model for streamflow simulation. *Journal of Hydrology*. 2019; **573**:1-2.

[35] Mehdizadeh S, Fathian F, Safari MJ, Adamowski JF. Comparative assessment of time series and artificial intelligence models to estimate monthly streamflow: a local and external data analysis approach. *Journal of Hydrology*. 2019; **579**:124225.

[36] Adnan RM, Liang Z, Heddam S, Zounemat-Kermani M, Kisi O, Li B. Least square support vector machine and multivariate adaptive regression splines for streamflow prediction in mountainous basin using hydro-meteorological data as inputs. *Journal of Hydrology*. 2020; **586**:124371.

[37] Fathian F, Mehdizadeh S, Sales AK, Safari MJ. Hybrid models to improve the monthly river flow prediction: Integrating artificial intelligence and non-linear time series models. *Journal of Hydrology*. 2019; **575**:1200-13.

[38] Kisi O, Heddam S. Evaporation modelling by heuristic regression

approaches using only temperature data. *Hydrological Sciences Journal*. 2019; **64**(6):653-72.

[39] Kaya YZ, Mamak M, Üneş F, Demirci M. Evapotranspiration prediction using M5T method and Ritchie equation for St. Johns, FL, USA. 2017.

[40] Alipour A, Yarahmadi J, Mahdavi M. Comparative study of M5 model tree and artificial neural network in estimating reference evapotranspiration using MODIS products. *Journal of Climatology*. 2014; **2014**.

[41] Shrestha S, Nepal S. Water balance assessment under different glacier coverage scenarios in the Hunza Basin. *Water*. 2019; **11**(6):1124.

[42] Saifullah M, Liu S, Adnan M, Ashraf M, Zaman M, Hashim S, Muhammad S. Risks of Glaciers Lakes Outburst Flood along China Pakistan Economic Corridor. In *Glaciers and Polar Environment*. 2020. IntechOpen.

[43] Hewitt K. Glacier change, concentration, and elevation effects in the Karakoram Himalaya, Upper Indus Basin. *Mountain Research and Development*. 2011; **31**(3):188-200.

[44] Ali AF, Zhang XP, Adnan M, Iqbal M, Khan G. Projection of future streamflow of the Hunza River Basin, Karakoram Range (Pakistan) using HBV hydrological model. *Journal of Mountain Science*. 2018; **15**(10):2218-35.

[45] Tahir AA, Chevallier P, Arnaud Y, Ahmad B. Snow cover dynamics and hydrological regime of the Hunza River basin, Karakoram Range, Northern Pakistan. *Hydrology and Earth System Sciences*. 2011; **15**(7):2275-90.

[46] Shrestha M, Koike T, Hirabayashi Y, Xue Y, Wang L, Rasul G, Ahmad B. Integrated simulation of snow and



- glacier melt in water and energy balance based, distributed hydrological modeling framework at Hunza River Basin of Pakistan Karakoram region. *Journal of Geophysical Research: Atmospheres*. 2015; 120(10):4889-919.
- [47] Quinlan JR. Learning with continuous classes. In 5th Australian joint conference on artificial intelligence. 1992; 92:343-348.
- [48] Witten IH, Frank E, Hall MA, Pal CJ. Practical machine learning tools and techniques. Morgan Kaufmann. 2005:578.
- [49] Rahimikhoob A, Asadi M, Mashal M. A comparison between conventional and M5 model tree methods for converting pan evaporation to reference evapotranspiration for semi-arid region. *Water Resources Management*. 2013; 27(14):4815-26.
- [50] Adnan RM, Petroselli A, Heddad S, Santos CA, Kisi O. Comparison of different methodologies for rainfall-runoff modeling: machine learning vs. conceptual approach. *Natural Hazards*. 2021; 105(3):2987-3011.
- [51] Adnan RM, Liang Z, Yuan X, Kisi O, Akhlaq M, Li B. Comparison of LSSVR, M5RT, NF-GP, and NF-SC models for predictions of hourly wind speed and wind power based on cross-validation. *Energies*. 2019; 12(2):329.
- [52] Pal M, Deswal S. M5 model tree based modelling of reference evapotranspiration. *Hydrological Processes: An International Journal*. 2009; 23(10):1437-43.
- [53] Friedman JH. Multivariate adaptive regression splines. *The Annals of Statistics*. 1991:1-67.
- [54] Zhang W, Goh AT. Multivariate adaptive regression splines and neural network models for prediction of pile drivability. *Geoscience Frontiers*. 2016; 7(1):45-52.
- [55] Kisi O, Parmar KS. Application of least square support vector machine and multivariate adaptive regression spline models in long term prediction of river water pollution. *Journal of Hydrology*. 2016; 534:104-12.
- [56] Liu Z, Zhou P, Chen G, Guo L. Evaluating a coupled discrete wavelet transform and support vector regression for daily and monthly streamflow forecasting. *Journal of Hydrology*. 2014; 519:2822-31.
- [57] Yuan X, Wu X, Tian H, Yuan Y, Adnan RM. Parameter identification of nonlinear Muskingum model with backtracking search algorithm. *Water Resources Management*. 2016; 30(8): 2767-83.
- [58] Meng E, Huang S, Huang Q, Fang W, Wu L, Wang L. A robust method for non-stationary streamflow prediction based on improved EMD-SVM model. *Journal of Hydrology*. 2019; 568:462-78.
- [59] Zhou Y, Guo S, Chang FJ. Explore an evolutionary recurrent ANFIS for modelling multi-step-ahead flood forecasts. *Journal of Hydrology*. 2019; 570:343-55.
- [60] Alizamir M, Kisi O, Muhammad Adnan R, Kuriqi A. Modelling reference evapotranspiration by combining neuro-fuzzy and evolutionary strategies. *Acta Geophysica*. 2020; 68:1113-26.
- [61] Yuan X, Chen C, Lei X, Yuan Y, Adnan RM. Monthly runoff forecasting based on LSTM-ALO model. *Stochastic environmental research and risk assessment*. 2018; 32:2199-212.