Snow grain size mapping in Upper Himalayas using Hyperion data

Anisha Narendran and RAAJ Ramsankaran[#]

Department of Civil Engineering
Indian Institute of Technology Bombay, Mumbai -400076, India
[#]Email: ramsankaran@civil.iitb.ac.in

KEY WORDS: snow grain size, hyperspectral, SAM classifier, GSI classifier, SVM classifier

ABSTRACT

Estimation of physical property of snow such as grain size through classification of hyperspectral images is very challenging task due to minute differences between the classes which is complicated by limited availability of field measurements (both spectra and physical property). To overcome this limitation, in this study a Support Vector Machine (SVM) based classification using Radial basis kernel has been tried to classify the snow cover data of Hyperion for Upper Himalayan regions. SVM defines the classification model by the concept of margin maximization using few training pixels. The generalization capability of SVM is due to the selection of hyperplane that maximizes the geometrical margin between the classes and with the hyperspectral data, this allows the full exploitation of the discrimination capability of the relatively few training samples available. Due to unavailability of field data, grain size map generated using Spectral Angle Mapper (SAM) from Negi et al., 2013 has been considered as reference map. Initially, a visual comparison made between the snow grain size maps obtained from the SVM method and Grain Size Index (GSI) method indicates that the former is able to capture the spatial distribution of different grain sizes such as coarse, medium and fine. Further, a detailed comparison of the estimated snow grain sizes map obtained from SVM method and Grain Size (GSI) method and SAM method was also carried out using matrix method, which gave an overall matching of 79.08% and 79.63% against GSI and SAM methods respectively.

1. INTRODUCTION

Apart from the poles and Alps, the Himalayan mountain ranges have a large concentration of glaciers and huge seasonal snow cover. As glaciers and snow are natural resources that affect the human life and ecosystem by contributing to stream flow, melt run-off, regional climate, hydropower generation, winter tourism and other development strategies, thus, studying the Himalayan region is of high importance. Due to the rugged terrain and harsh climatic conditions in the Himalayas, it is difficult to use conventional methods to monitor snow and ice cover. In such scenario, remote sensing tools have emerged as an alternative solution. Snow grain size is one of the physical properties which contributes mainly to the metamorphic changes in snow cover, spectral albedo, snow melting, avalanches, etc (Negi *et al*., 2013) and thus its estimation is necessity. For purpose of snow physical parameter estimations, multispectral satellite imagery is being used widely.

With the recent advances in remote sensing technologies, the use of hyperspectral data would facilitate an improved and accurate snow-cover-characteristics as even minute information can be captured by hyperspectral sensor and this improves the discrimination among similar ground cover classes than the traditional multispectral images (Moughal, 2013). Hyperspectral sensors are able to record reflected light from land surface objects in numerous narrow continuous spectral bands from the visible to the shortwave infrared parts of the electromagnetic spectrum, acquiring a vast amount of spectral information (Petropoulos *et* al., 2011). The only imaging data available over the Himalayan regions is from spaceborne EO-1, Hyperion sensor. It collects data at 220 unique spectral channels ranging from 357 to 2576nm with a 10nm bandwidth (EO-1 User Guide 2003). These data collected in turn can be effectively used to provide multiple snow cover characteristics like grain size, albedo, fractional snow cover, contaminated snow, etc (Negi *et al*., 2013).

Grain size classification can be done by Spectral Angle Mapping (SAM), Grain Size Index (GSI) method and Asymptotic Radiative transfer (ART) (Negi *et al*.,2013). SAM and GSI are qualitative methods whereas ART is a quantitative method. Analysis of hyperspectral data is a quite tedious task due to large spatial variability of the spectral signature of each land cover class, atmospheric effects and high dimensionality (Moughal, 2013). To overcome this tedious task, machine learning techniques can be adopted. Support Vector Machines (SVM) have been extensively used for regression and classification of multispectral (Mitra *et al*., 2004) and hyperspectral images (Mercier and Lennon, 2003). This method further improves the classification accuracy of hyperspectral data (Moughal, 2013). SVM does not require an estimation of the statistical distribution of classes to classify an image but only requires a model which can be trained using few training pixels to exploit the concept of margin maximization. SVM is also an effective method for small samples learning, has better generalization ability and high efficiency for learning (Moughal, 2013). SVM will also provide qualitative estimation of grain size. These techniques are essential where quantitative information is not essential, such as in the case when output is needed within a short duration. Therefore, in this study the potential of SVM classifier on hyperspectral imagery for snow grain size estimation has been assessed using the available limited data, wherein the obtained result was compared with the classification maps generated by SAM and GSI respectively.

2. STUDY AREA AND DATA USED

Due to non-availability of field data at the time of satellite overpass, the same study area of upper western Himalayas used in Negi *et* al., 2013 has been adopted to visually compare the classification results. The study area is the Hyperion image of 25th November, 2003 covering parts of the upper western Himalayas (Figure 1)**Error! Reference source not found.** with the following coordinates (Latitudinal extent 34°35′00″-34°55′00″ and longitudinal extent 77°40′00″-77°55′00″).

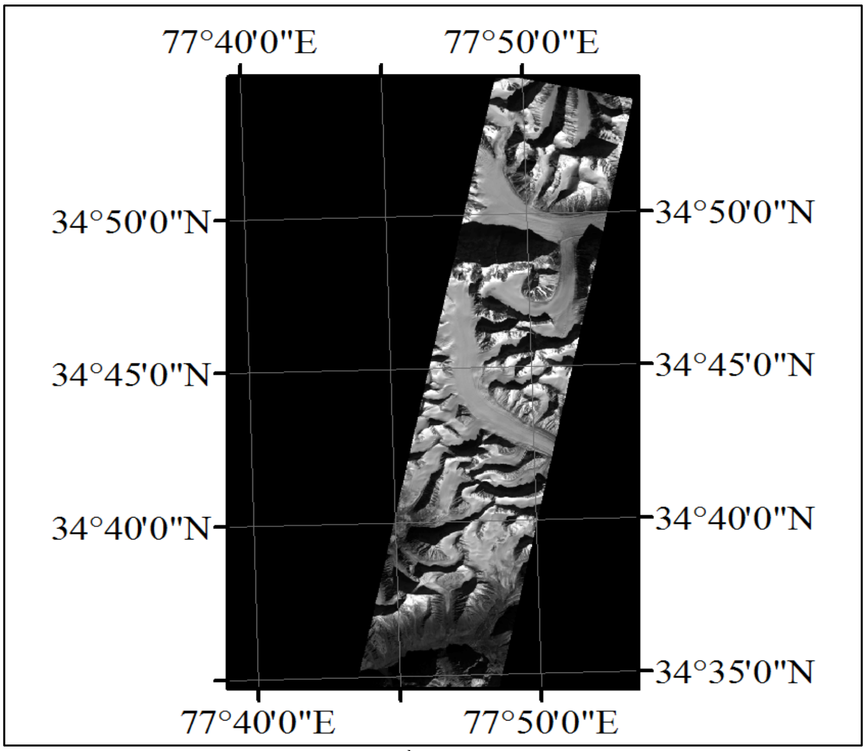

Figure 1. Hyperion satellite image of 25th November, 2003 of Upper Himalaya

The upper western Himalayan zone is extremely cold, receives dry snow and is highly glaciated with little or no vegetation has permanent snow which resembles continental snow conditions and an altitude greater than 5000m (Negi and Kokhanovsky, 2011). This area is characterized by low temperature, light snowfall and severe wind activities and has a gradual slope less than 5° and it avoids fractional snow cover (Negi *et al.*, 2013).

3. METHODOLOGY

The overall methodology adopted in this study is presented in the flowchart given in Figure 2.

The satellite image was first pre-processed for removing atmospheric errors and the resultant image was then classified using SAM and GSI approach to reproduce the snow grain size maps following the criteria mentioned in Negi *et al*., 2013. The snow grains were classified into coarse, medium and fine classes as given in Table 1. Due to the absence of ground data, based on the conclusions by Negi *et al*., 2013, the map generated by SAM was taken as the true map and used to collect the training samples for the SVM model. Thus, obtained snow grain size map by SVM was then visually compared with the maps generated by GSI and SAM classifiers separately. It was also evaluated by matrix matching method against the GSI and SAM. Details of the SAM, GSI and SVM implementation are explained in the following sections.

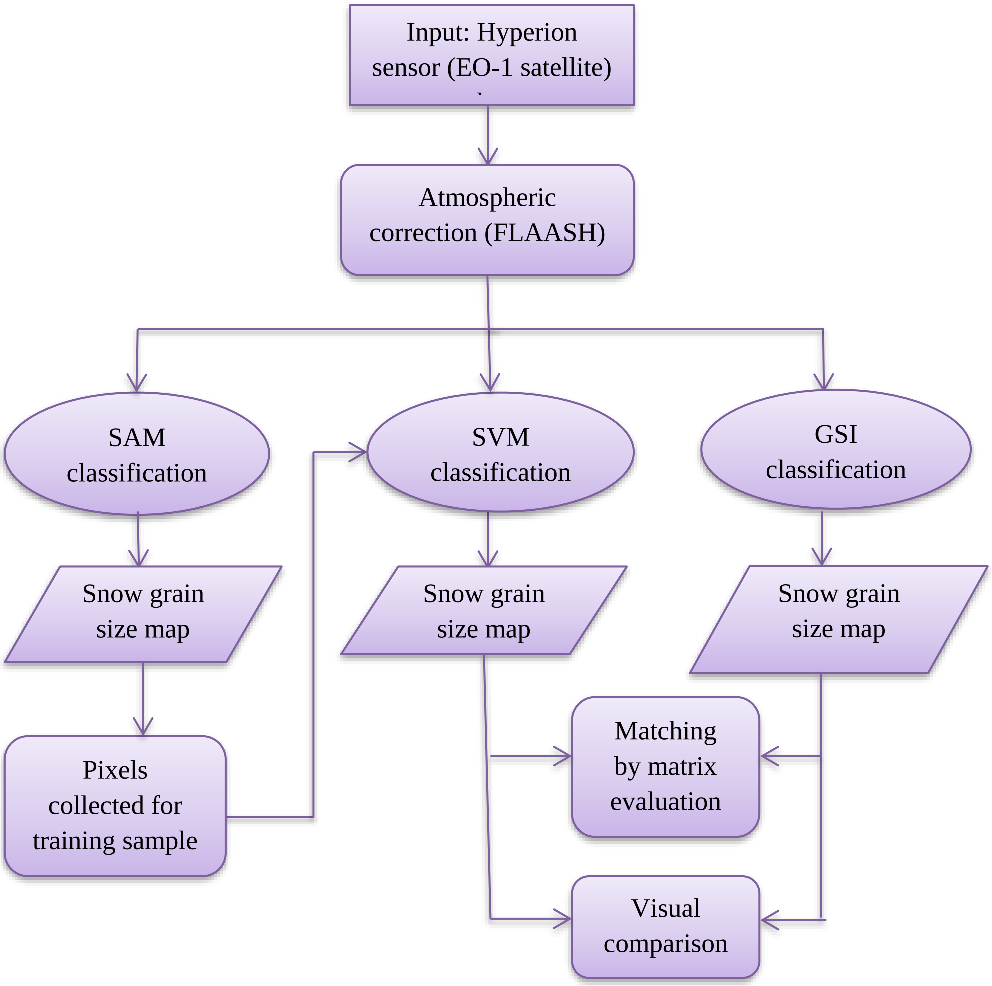

Figure 2. Flowchart of the overall methodology adopted

Table 1 Snow grain size classes and details adapted from Negi *et al.*, 2013

SL No	Class	Description: Range of snow grain sizes in mm	Threshold for GSI
1	Coarse	1 to 2	0.26 to 0.35
2	Medium	0.5 to 1	0.17 to 0.26
3	Fine	< 0.5	0 to 0.17
4	Unclassified	Non-snow-covered area	NA

3.1 Atmospheric correction

Initially, the Fast Line-of-sight Atmospheric Analysis of Spectral Hypercube (FLAASH) atmospheric correction was done and TOAR (Top of Atmosphere Reflectance) image is obtained. In FLAASH module, there are six atmospheric models as given in Table 2. This accounts for 120 models of correction to be applied. Out of these, the data was atmospherically corrected for three models namely, sub-arctic winter, mid-latitude winter and tropical atmospheric models. The dataset is of November 2003, which implies winter season and hence, sub-arctic winter and mid latitude winter atmospheric models were chosen. Also, the image is of Upper Western Himalayas, which lies in the tropical belt and hence, tropical atmospheric model was taken. This narrows down the models to 18 (Table 1). These selected models are in confirmation with the ones used by Negi et al., 2013 but the information regarding the combinations used were not given in their study. Hence, the FLAASH was run for all possible combinations and snow grain size maps were generated for all FLAASH corrected reflectance images.

Table 2. Total models available in FLAASH module

Atmospheric model	6	Sub-arctic winter, Mid-latitude winter, U. S. Standard, Sub-arctic summer, Mid-latitude summer, Tropical
Water retrieval	4	None, at 820nm, at 940nm, at 1135nm
Aerosol model	5	None, Rural, Urban, Maritime, Tropospheric
Total combinations	6x4x5=120	

Due to anomalies observed in the spectra of the processed data, the bad bands were removed at the end of FLAASH correction. After band removal, the number of bands was reduced from 242 to 131. The hyperspectral image after correction having smooth spectra without anomalies was used as input for the classifiers.

Table 3. FLAASH models applied in this study

Atmospheric model	Water retrieval	Aerosol model	Total combinations
Sub-arctic winter	None, 820nm, 1135nm	Rural, urban	3x2 = 6
Mid- latitude winter	None, 820nm, 1135nm	Rural, urban	3x2 = 6
Tropical	None, 820nm, 1135nm	Rural, urban	3x2 = 6
Total	6x3= 18		

3.2 Spectral Angle Mapper (SAM)

Spectral Angle Mapper (SAM) is a supervised classification algorithm which works on the principle of similarity. It utilizes the similarity between the spectral signature of feature class and the spectra of pixel derived from the hyperspectral image data. It classifies the image by calculating the spectral similarity between the image spectrums to reference reflectance spectra (Moughal, 2013). The reference spectra of the feature class and max thresholding SAM angle is fed as inputs to the classifier. The reference spectra can be either collected from field using a spectroradiometer, extracted directly from image or from spectral library which is of John Hopkins University (JHU) collected under lab conditions. The field spectra *(obtained from Dr. Negi through email communication)* was used to reproduce snow grain size map by SAM as done in Negi *et al*., 2013. Following Negi *et al*., 2013's work, the maximum SAM angle for thresholding was taken as 8° for snow grain size classification.

3.3 Grain Size Index (GSI)

The Grain Size index (GSI) method proposed by Negi *et al*., 2010 was implemented to reproduce the snow grain size map as done by Negi *et al*.,2013 and used for validation against the SVM classified image. In this method, first the snow pixels were identified using Normalized Difference Snow Index (NDSI). NDSI uses the green and Shortwave Infrared (SWIR) difference band ratio and its value range between -1 to +1 (Negi *et al*., 2010). The NDSI threshold is raised from 0.4 to 0.6 to avoid fractional snow cover and heavily contaminated snow. NDSI is also used to separate snow from cumulus clouds (Negi *et al*., 2013). Then the GSI was calculated on the pure snow pixels generated from NDSI, which uses Visible and Infrared (VNIR) bands i.e., 590nm and 1050nm. The thresholds used to identify the classes and their details are given in Table 1. The thresholds given by Negi *et al*., 2010 were used to generate snow grain size maps.

3.4 Support Vector Machine (SVM)

The implementation of Support Vector Machines (SVM) on hyperspectral imagery has gained popularity in the recent times. SVM outperforms classical supervised classifiers like maximum likelihood and is a robust alternative for pattern recognition algorithms with hyperspectral data, as it is based on a geometric point of view and there is no statistical estimation. It performs better especially in scenarios where number of training samples are limited and when their number of bands is high (Mercier and Lennon, 2003).

SVM is a non-parametric binary classifier that locates the optimal hyper plane between the two classes to separate them in a mulita dimensional feature space by taking into account only the training samples that lie on the edge of the class distributions known as support vectors (Moughal, 2013). This provides best separation between two classes in a multidimensional feature space. SVM works well with hyperspectral imagery as it is insensitive to dimensionality. A variety of kernels like polynomial, radian basis function (RBF) and sigmoid are used to represent more complex shapes than linear hyperplanes (Petropoulos *et al*., 2012). SVM method is found to be effective in cases of heterogeneous classes where only few pixels are available for training. SVM classification can also be extended to multiclass classification by splitting the optimization problem into a series of binary class separations (Petropoulos *et al*., 2012).

In this study, due to lack of field data, training samples for the training of Support Vector Machine learning algorithm was collected from the SAM classified grain size map. Therefore, the snow grain size map generated by SAM was considered as a true map. Pure pixels representative of all grain size classes, along with unclassified class were collected from the Hyperion imagery. The training sites are selected in a way that the sites are restricted to pure pixels only. Approximately 1000 pixels per class were selected and used to train the classifier. RBF kernel was chosen as it requires few parameters and gives good results (Petropoulos *et al*., 2012). The parametrization of the RBF kernel was done in reference to ENVI User's Guide (2008). The penalty number which deals with misclassification was set as 100, which implies no misclassification of pixels. The classification probability threshold was set to zero. This means that all image pixels have to be classified into one class and hence, for good results an "unclassified" class was defined and unclassified pixels were collected for training the same. Hyperion image was classified at full resolution and for this the pyramid parameter was set to zero.

The map generated by SVM was then visually validated against the maps produced by SAM and GSI approaches mentioned in Secs. 3.2 and 3.3. Furthermore, a matrix matching of SVM against SAM and GSI was also done to find the overall matching area of the different snow grain classes.

4. RESULTS & DISCUSSIONS

By taking Negi *et* al., 2013 as reference, snow grain size maps by SAM and GSI methods have been reproduced as shown in Figure 3(a) and Figure 3(b) respectively to compare the snow grain size map generated using SVM classifier Figure 3(c). The three classifiers were run for all the FLAASH combinations mentioned in Table 3. Only the best results are shown here. The best snow grain size maps were obtained for sub-arctic winter atmospherically model with water absorption at 820nm with no aerosol. It is to be noted that all of these methods give a qualitative estimate of the snow grain size. On visual comparison with the reference maps created by reproducing the work done by Negi *et al*., 2013, it can be observed that the map produced by SVM classifier gives similar results. Thus, this method can also be adopted for grain size classification. From Figure 3, it is seen that SVM classified map has almost same spatial distribution of the different snow classes with respect to the reference maps (Figure 3(a) and Figure 3(b)). Due to the absence of ground truth data, the three classifiers can only be compared qualitatively against each other. Hence, based on this visual comparison, it could be primarily concluded that SVM classifier outperforms GSI. To get a detailed understanding of the areas covered by each class, the SVM map was compared against GSI map and SAM map using matrix method. For this, two confusion matrices of SVM classified image was generated with GSI classified image and SAM classified image respectively. Here, both the matrices are represented together in Table 4. Instead of pixel count of classes, percentage area of classes was used to fill the confusion matrix as the former gave huge numbers, hence, for easy interpretation and representation, latter was used. The percentage area of grain size classes was then matched against the methods. By taking the sum of diagonal elements, it gave an overall matching of 79.08% and 79.63% against GSI and SAM methods respectively.

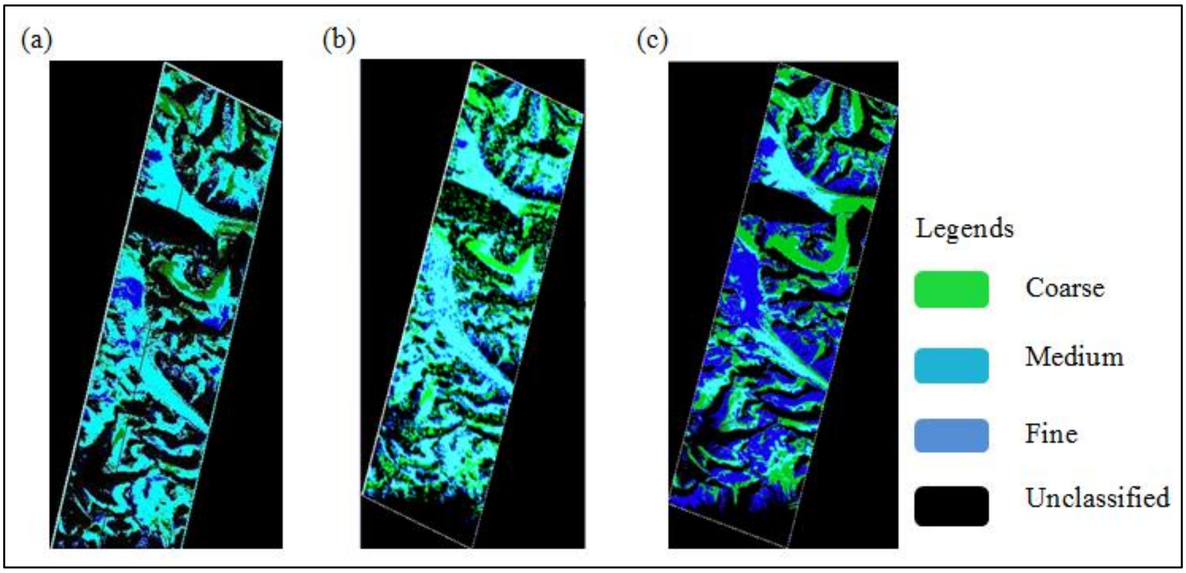

Figure 3. Snow grain size map generated by (a) SAM (b) GSI (c) SVM, for sub-arctic winter model with water absorption at 820nm with no aerosol (FLAASH) model

Table 4. Matrix comparing different snow grain size classes (% area of classes) using results of SVM method to GSI method and SAM method

	Support Vector Machine Classification					
GSI Classification; (SAM Classification)	**Grain Size Classes (% area of classes)**	**Coarse**	**Medium**	**Fine**	**Unclassified**	**Total**
	Coarse	2.72 (5.29)	0.37 (0.06)	0.06 (0.74)	2.48 (0.01)	5.63 (6.1)
	Medium	3.46 (5.12)	5.53 (1.55)	4.99 (0.35)	1.59 (0.06)	15.57 (7.08)
	Fine	0.12 (0.75)	0.14 (0)	4.22 (8.86)	1.23 (5.32)	5.71 (14.93)
	Unclassified	0.77 (0.95)	0.06 (5.37)	5.65 (1.64)	66.61 (63.93)	73.09 (71.89)
	Total	7.07 (12.11)	6.1 (6.98)	14.92 (11.59)	72 (69.32)	100 (100)

Note: Confusion matrix is filled with values representing % area of different snow grain size classes. First value in the confusion matrix is the values of grain size classes of SVM against GSI method; second value in the confusion matrix given within the parenthesis are the values of grain size classes of SVM against SAM method

Based on the obtained results from 18 atmospheric models, an interesting observation is made on the effect of FLAASH atmospheric models on the classifiers. Obtained results indicate that the SAM and GSI classifiers performed differently with each of the FLAASH atmospheric models. SAM gave best results with sub-arctic winter model whereas GSI performed well with tropical model. SAM performed poorly with tropical model and even GSI did not give any result in case of some tropical models without aerosol. In contrast to this non-uniformity with SAM and GSI, the maps generated by SVM gave almost uniform and similar results for all 18 models. However, detailed investigations are needed to understand this behavior of SVM with different atmospheric correction models.

5. CONCLUSIONS

The snow grain size map generated using SVM had an overall matching area of about 79.08% and 79.63% against both GSI and SAM classified maps respectively. Thus, it can be concluded that SVM can be effectively used for the classification snow grain size in snow covered areas even with less training dataset. Due to data unavailability of ground truth data, quantitative estimation of property cannot be carried out. Based on the obtained results, it is suggested to investigate further on the effect of FLAASH atmospheric models on SVM.

REFERENCES

Beck, R. (2003). EO-1 User Guide. *Ohio: Department of Geography*.

Guide, E. U. S. (2008). ENVI on-line software user’s manual. *ITT Visual Information Solutions*.

Mercier, G., & Lennon, M. (2003, July). Support vector machines for hyperspectral image classification with spectral-based kernels. In *Geoscience and Remote Sensing Symposium, 2003. IGARSS'03. Proceedings. 2003 IEEE International* (Vol. 1, pp. 288-290). IEEE.

Mitra, P., Shankar, B. U., & Pal, S. K. (2004). Segmentation of multispectral remote sensing images using active support vector machines. *Pattern recognition letters*, *25*(9), 1067-1074.

Moughal, T. A. (2013). Hyperspectral image classification using Support Vector Machine. In *Journal of Physics: Conference Series* (Vol. 439, No. 1, p. 012042). IOP Publishing.

Negi, H. S., Jassar, H. S., Saravana, G., Thakur, N. K., Snehmani, & Ganju, A. (2013). Snow-cover characteristics using Hyperion data for the Himalayan region. *International journal of remote sensing*, *34*(6), 2140-2161.

Negi, H. S., & Kokhanovsky, A. (2011). Retrieval of snow grain size and albedo of western Himalayan snow cover using satellite data. *The Cryosphere*, *5*(4), 831.

Negi, H. S., Shekhar, C., & Singh, S. K. (2015). Snow and glacier investigations using hyperspectral data in the Himalaya. *Current Science*, *108*(5), 892.

Negi, H. S., Singh, S. K., Kulkarni, A. V., & Semwal, B. S. (2010). Field-based spectral reflectance measurements of seasonal snow cover in the Indian Himalaya. *International Journal of Remote Sensing*, *31*(9), 2393-2417.

Petropoulos, G. P., Kalaitzidis, C., & Vadrevu, K. P. (2012). Support vector machines and object-based classification for obtaining land-use/cover cartography from Hyperion hyperspectral imagery. *Computers & Geosciences*, *41*, 99-107.

Sharma, S. S., & Ganju, A. (2000). Complexities of avalanche forecasting in Western Himalaya—an overview. *Cold regions science and technology*, *31*(2), 95-102.

Yu, H., Xie, T., Paszczynski, S., & Wilamowski, B. M. (2011). Advantages of radial basis function networks for dynamic system design. *IEEE Transactions on Industrial Electronics*, *58*(12), 5438-5450.