 CMC, vol.57, no.1, pp.49-68, 2018

Snow Cover Mapping for Mountainous Areas by Fusion of MODIS L1B and Geographic Data Based on Stacked Denoising Auto-Encoders

Xi Kan[1], Yonghong Zhang[2,*], Linglong Zhu[2], Liming Xiao[2], Jiangeng Wang[3], Wei Tian[4] and Haowen Tan[5]

Abstract: Snow cover plays an important role in meteorological and hydrological researches. However, the accuracies of currently available snow cover products are significantly lower in mountainous areas than in plains, due to the serious snow/cloud confusion problem caused by high altitude and complex topography. Aiming at this problem, an improved snow cover mapping approach for mountainous areas was proposed and applied in Qinghai-Tibetan Plateau. In this work, a deep learning framework named Stacked Denoising Auto-Encoders (SDAE) was employed to fuse the MODIS multispectral images and various geographic datasets, which are then classified into three categories: Snow, cloud and snow-free land. Moreover, two independent SDAE models were trained for snow mapping in snow and snow-free seasons respectively in response to the seasonal variations of meteorological conditions. The proposed approach was verified using in-situ snow depth records, and compared to the most widely used snow products MOD10A1 and MYD10A1. The comparison results show that our method got the best performance: Overall accuracy of 98.95% and F-measure of 73.84%. The results indicated that our method can effectively improve the snow recognition accuracy, and it can be further extended to other multi-source remote sensing image classification issues.

Keywords: Snow cover, remote sensing, deep learning, Qinghai-Tibetan Plateau, MODIS L1B.

1 Introduction

Snow cover is a key influencing factor in many research areas of earth science. In meteorology, snow cover is a sensitive indicator of global climate system [Bi, Xie, Huang

[1] School of Atmospheric Science, Nanjing University of Information Science and Technology, Nanjing, 210044, China.

[2] School of Information and Control, Nanjing University of Information Science and Technology, Nanjing, 210044, China.

[3] School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing, 210044, China.

[4] School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, 210044, China.

[5] Department of Computer Engineering, Chosun University, Gwangju, 501-759, South Korea.

[*] Corresponding Author: Yonghong Zhang. Email: zyh@nuist.edu.cn.

et al. (2015)]. Meanwhile, snow cover has a significant impact on radiation budget of the earth surface, due to its high albedo [Hori, Sugiura, Kobayashi et al. (2017)]. In hydrology, snow cover contains a huge quantity of fresh water that can be released suddenly in spring, which makes snow cover an important component of regional water cycle, especially in mountainous regions [Yang, Wu and Qin (2014)]. The snow cover in Qinghai-Tibetan Plateau (QT Plateau) has particular importance in the climate and hydrology system, due to its unique geographical location and topography [Zhang, Cao, Kan et al. (2016)]. Thus, the snow cover in QT Plateau exerts a significant thermal forcing on the local atmospheric general circulation [Pu and Xu (2009)]. Therefore, the accurately and timely snow cover monitoring over the QT Plateau is important for many metrologic and hydrologic issues [Tang, Shrestha, Li et al. (2013)].

So far, various optical remote sensors have been used to extract snow cover extent, mainly including the Advanced Very High Resolution Radiometer (AVHRR) [Allen, Durkee, Wash et al. (1990)], the Moderate Resolution Imaging Spectroradiometer (MODIS) on board the Terra and Aqua satellites [Hall, Riggs and Salomonson (1995, 2002); Hall, Riggs, Foster et al. (2010)], the Landsat Thematic Mapper (TM) [Rosenthal and Dozier (1996)], the Systeme Probatoire d'Observation de la Terre (SPOT) [Dankers and De Jong (2004)] and other optical sensors [Yang, Jiang, Shi et al. (2014); Yang, Wu, Qin et al. (2014); Zhang, Kan, Ren et al. (2017)].

Currently, some available satellite-based snow cover products have been developed [[Dietz, Wohner and Kuenzer (2012); Dozier and Painter (2004)], most of which are based on Normalized Difference Snow Index (NDSI), utilizing the specific spectral characteristic of snow cover: Strong reflection in the visible bands and the strong absorption in the shortwave infrared bands. Among these snow products, the MODIS snow cover products are most widely used in meteorology [Dietz, Wohner and Kuenzer (2012)] and hydrology [Thirel, Salamon, Burek et al. (2013)]. The data quality of MODIS snow products has been verified by many studies with positive results [Hall and Riggs (2007)]. However, previous studies showed that the cloud obstruction and snow misclassification are the two major limitations of MODIS snow products [Dong and Menzel (2016a)].

Although MODIS snow products have good classification accuracies in general, the data quality in some specific regions and periods is seriously decreased by the snow misclassification errors [Dariane and Khoramian (2017)]. The main source of misclassification errors are the snow/cloud confusion errors. The error may be caused by either misidentifying cloud cover as snow or identifying thin, sparse snow cover as cloud in the cloud mask [Tang, Shrestha, Li et al. (2013)]. Moreover, the snow/cloud confusion caused misclassification errors could increase even more dramatically under complicated geographical and meteorological conditions, typically, in mountainous areas. This problem is particularly acute in QT Plateau with the highest average elevation and complicated topography, because the snow cover in QT Plateau are often shallow, patchy and frequently of short duration [Qin, Liu and Li (2006)]. Liu et al. [Liu, Jin and Ke (2014)] shows that the accuracies of currently available snow cover products are significantly lower in QT Plateau than in other areas around the world. A comprehensive evaluation [Yang, Jiang, Ménard et al. (2015)] of MODIS and IMS (Interactive Multisensor Snow and Ice Mapping System) snow cover

products shows that the IMS product presented a serious overestimation error in grassland area of the QT Plateau; the MODIS snow cover product exhibited a high underestimation error in forested regions of the QT Plateau. Pu et al. [Pu, Xu, Vincent et al. (2007)] evaluated the accuracies of the MODIS 8-day snow cover product over the QT Plateau from 2001 to 2003, results show that the overall accuracy is about 90%, and the major omission errors may be caused by the shallow snow cover.

However, there is infrequent research effort focus on reducing these misclassification errors [Dong and Menzel (2016a)]. Tang et al. [Tang, Shrestha, Li et al. (2013)] presented an improved snow mapping algorithm for the QT Plateau using MODIS data. In this work, MODIS cloud mask product (MOD35) and MODIS land surface temperature product (MOD11) are used to improve the snow cover recognition accuracy. Thompson et al. [Thompson, Paull and Lees (2015)] employed an improved liberal cloud mask to address snow/cloud confusion errors of MODIS snow products, and the Normalized Difference Vegetation Index (NDVI) is calculated to reduce confusion errors between snow-free and snow-covered forests. Dong et al. [Dong and Menzel (2016b)] used two meteorological filters and topographical correction to reduce the overestimation errors of MODIS 8-day snow products. Sirguey et al. [Sirguey, Mathieu and Arnaud (2009)] presented a comprehensive method for subpixel snow cover mapping, in this work, an iterative approach was used to correct imagery for both atmospheric and topographic effects using daily observations of atmospheric parameters and terrain factors (including slope, aspect and curvature). From these works, we can see that the potential of utilizing meteorological parameters and geographical information (mainly consist of topography and vegetation coverage information in this issue) can improve MODIS snow products. But the effects are still limited in mountainous areas. Researches show that existing cloud and temperature masks, which were widely used in above improved methods, have limited effect in mountainous areas, because of the very high altitude and rough terrain [Yang, Jiang, Ménard et al. (2015); Riggs and Hall (2016)]. Thus, it is obvious that exactly distinguishing snow from clouds in mountainous areas is very difficult, which is closely related to many impact factors: Altitude, slope, aspect, viewing angles, land cover type, etc. It is hopeless to develop a good enough mathematical model for snow recognition in mountainous areas, which have to take all the above impact factors into consideration.

Machine learning methods have been used in many pattern recognition problems. Except the above NDSI based mathematical models, previous studies have tried using statistical and machine learning methods for snow cover recognition, such as Support Vector Machine (SVM) [Zhu, Xiao, Feng et al. (2014)] and Artificial Neural Network (ANN) [Dobreva and Klein (2011)]. However, there are several critical problems in snow cover recognition using MODIS multispectral data and other auxiliary data for QT plateau: (a) Curse of dimensionality; (b) Limited number of labeled training samples; (c) Large spatial variability of spectral signature [Chen, Lin, Zhao et al. (2017)]. Most of these traditional statistical classifiers only have less than three hidden layers, it is hard to extract more abstract, invariant features from high dimensional data. Previous studies showed that human brains have strong ability to deal with delicate task like object recognition, because of its multiple stages of processing from retina to cortex [Kruger, Janssen, Kalkan et al. (2013)]. Depend on this theory, the deep learning (DL) method has

been proposed by Hinton et al. [Hinton, Osindero and Teh (2006)], these DL models with multiple layers architecture have the ability to learn a set of rich nonlinear representations directly from the input data with no assumptions or prior knowledge [Marmanis, Datcu, Esch et al. (2016)]. Therefore, DL based classifiers are believed to have the ability of classifying high dimensional data and getting higher classification accuracy than traditional shallower classifiers [Zhang, Zhang and Du (2016)]. In recent years, dozens of DL architectures have been proposed by Krizhevsky et al. [Krizhevsky, Sutskever and Hinton (2012); Vincent, Larochelle, Lajoie et al. (2010); Salakhutdinov and Hinton (2009)] and applied in many fields such as image classification, face recognition, pedestrian detection, natural language processing, and of course, the remote sensing recognition [Marmanis, Datcu, Esch et al. (2016); Zhang, Zhang and Du (2016); Li, Fu, Yu et al. (2016)]. These studies showed that these DL architectures can greatly improve the accuracies of classification or regression tasks in many fields, where they usually outperform the traditional methods [Li, Fu, Yu et al. (2016)].

In this paper, we proposed a DL based snow recognition method for the QT Plateau. Firstly, pre-processing the MODIS L1B multi-spectral remote sensing data and other geo-datasets, geo-coding all these data with a uniform geographic coordinates and spatial resolution. Secondly, building and initializing the SDAE model [Vincent, Larochelle, Lajoie et al. (2010)], which is one of the mainstream deep learning architectures, then initializing and training it through a large amount of unsupervised learning iterations to make the SDAE model be able to discover the complicated non-linear relations between multispectral data, topographic factors and the land cover types. Finally, initializing and supervised training a classifier network by the labeled samples, making it be able to classify the snow-free, snow and cloud. Moreover, in response to the seasonal variations of climate and environmental conditions, two independent SDAE models were trained for snow recognition in snow and snow-free seasons respectively.

2 Study area and data source

2.1 Study area

The QT Plateau is located in southwest China, the highest plateau with an average elevation of 4,000 m, with the Pamirs plateau to the west, the Hengduan Mountains to the east, the southern margin of the Himalaya Mountains to the south, the Kunlun Mountains to the north of Qilian Mountains. The scope of latitude and longitude is between 26°00’12”N to 39°46’50”N and 73°18’52”E to 104°46’59”E. From east to west, the Plateau covers a distance of about 2945 km, from north to south about 1532 km. The QT Plateau is the birthplace of Yangtze, the Yellow and other rivers, and of numerous lakes, accounting for about half of the total area of the lake which is supplied by the surrounding mountain snow melt water. The vegetation types of the QT Plateau are mainly dominated by meadow vegetation, and supplemented by grassland vegetation and desert vegetation. Fig. 1 shows the QT Plateau digital elevation and Meteorological stations. The warm season of QT Plateau is from April to September, and the cold season is from October to March. There is a small and uneven distribution of precipitation, decreasing from southeast to northwest, the annual precipitation in QT Plateau is about 480 mm. The wet season of QT Plateau is from May to September, because nearly 90% of annual

precipitation is concentrated in these months. The dry season each year is from October to April of the following year [Qin, Sun and Chen (2015)].

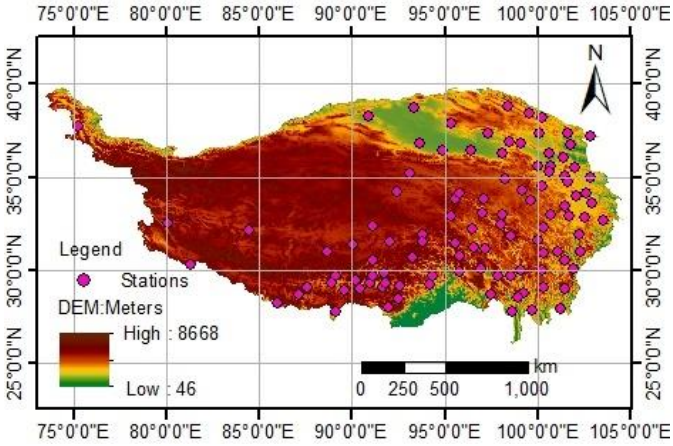

Figure 1: QT Plateau digital elevation and Meteorological stations

2.2 Data source

2.2.1 MODIS Level_1B

MODIS is a key sensor aboard the Terra and Aqua satellites, which has a viewing swath width of 2330 km and views the entire surface of the Earth every one to two days. Its detectors measure 36 spectral bands at three spatial resolutions: 250 m, 500 m and 1000 m. In this work, the following Aqua MODIS Level_1B (L1B) datasets were used: MYD02HKM, MYD021KM and MYD03. The MOD02HKM and MYD021KM data are calibrated MODIS L1B 500 m and 1 km granule products. All bands of MOD02HKM and MYD021KM were used in this work, except the Band 6, Band 13 (lo and hi), Band 14 (lo and hi) and Band 36, because on Aqua MODIS, most detectors of the above bands no longer function. The MOD03 data are the geo-location fields calculated for each 1 km MODIS Instantaneous field of view for all daytime orbits. MOD03 data include geodetic latitude, longitude, surface height above the geoid, solar zenith and azimuth angles, satellite zenith and azimuth angles.

2.2.2 Snow cover products

Three snow products have been used in this work: The MODIS daily snow products MOD10A1 and MYD10A1, and the Chinese FengYun-3 daily snow product FY3A_MULSS. These snow products were used for training samples preparation and result validation.

The MODIS daily snow cover products MOD10A1 and MYD10A1 have been available through the National Snow and Ice Data Center (NSIDC) Distributed Active Archive Center (DAAC) since September 13, 2000. The MODIS snow mapping algorithm was developed by Hall et al. [Hall, Riggs and Salomonson (1995)], based on the NDSI. MOD10A1 and MYD10A1consist of 1200 km by 1200 km tiles of 500 m resolution data

gridded in a sinusoidal projection. The version we used in this study is V5 (MODIS/Terra & Aqua Snow Cover Daily L3 Global 500 m Grid, Version 5). Although V6 is the most recent version available from NSIDC, but the fractional snow cover (FSC) and the binary snow-covered area (SCA) map are no longer calculated in V6. Users who wish to construct a binary SCA map should choose their own NDSI thresholds and apply them to the V6 NDSI data. Therefore, in this paper, we still use the binary SCA maps from V5 MODIS snow products.

Chinese FengYun-3 daily snow product FY3A_MULSS has been available through the Chinese National Satellite Meteorological Centre since August 2, 2009, with 1 km resolution in a geographic Latitude/Longitude projection.

2.2.3 Geographic data

The geographic information data include the altitude, slope, aspect and land cover data, these data are used for expressing the terrain features of the QT Plateau. The altitude data can be directly extracted from a DEM data, the slope and aspect data can be calculated from the DEM data. In this study, the DEM that we used comes from the Shuttle Radar Topography Mission (SRTM), at 90 m resolution, in a geographic Latitude/Longitude projection. The DEM data are currently distributed free of charge by the United States Geological Survey (USGS) and are available for downloading from the website of National Map Seamless Data Distribution System. The land cover information comes from the MODIS yearly land cover product: MCD12Q1 (V051), with a high spatial resolution of 500 m, in a sinusoidal projection. The primary land cover scheme identifies 17 land cover classes defined by the International Geosphere Biosphere Programme (IGBP), which includes 11 natural vegetation classes, 3 developed and mosaicked land classes, and 3 non-vegetated land classes. MCD12Q1 data are available for downloading from the website of Land Processes Distributed Active Archive Center (LP DAAC).

2.2.4 Meteorological data

To train and test the snow recognition model, the in-situ daily snow depth measurements, and daily minimum air temperature measurements were referred as the objective judgment standard. These in-situ measurements were observed from 96 ground meteorological stations over the QT Plateau from 2008 to 2013. The daily snow depth data were used for making labelled samples. The daily minimum air temperature data were used as an indicator for the seasonal patterns switching before the snow recognition process. These in-situ measurements were obtained from the National Meteorological Information Center, China.

3 Methods

3.1 Data pre-processing

The data pre-processing consists of three procedures: a) Data calibration; b) Geo-referencing and re-projection; c) Image mosaic and clip.

The data calibration processing is only used for MODIS L1B data. It transforms the original MODIS L1B data from digital numbers (DN) to physical values. It converts the DN values of visible and near-infrared bands into reflectance values, and converts the DN

values of infrared bands into the radiant brightness values. The geo-referencing and re-projection procedure were used to unify the resolution and projection of these multi-source remote sensing data, which will be registered by a uniform geographical coordinate. After this processing, all these data and snow products were uniformly converted into the GeoTIFF format, and were resampled to 0.005° resolution in a geographic Lat/Lon projection and WGS84 geographic coordinates. Because of the huge spatial scale of QT Plateau, more than two tiles of MOIDS L1B data can just cover the whole study area every day. Thus, through the image mosaic processing, these multiple tiles acquired from different times in one day can be spliced together and merged into one large overall image. Then the complete area of QT Plateau can be clipped out from this overall image. Eventually, the daily multi-source remote sensing data over the QT Plateau can be produced.

3.2 Snow recognition model based on SDAE

In this paper, a deep learning network based on Stacked Denoising Auto-Encoders (SDAE) was built and trained for snow cover recognition. First, creating and initializing a SDAE network; then, pre-training the SDAE network to enable the network to extract and express the features form the remote sensing data and geographic information; finally, fine tuning the SDAE network to classify the feature codes into three types: Snow cover, snow-free and cloud. The flow chart of the SDAE training process is shown in Fig. 2

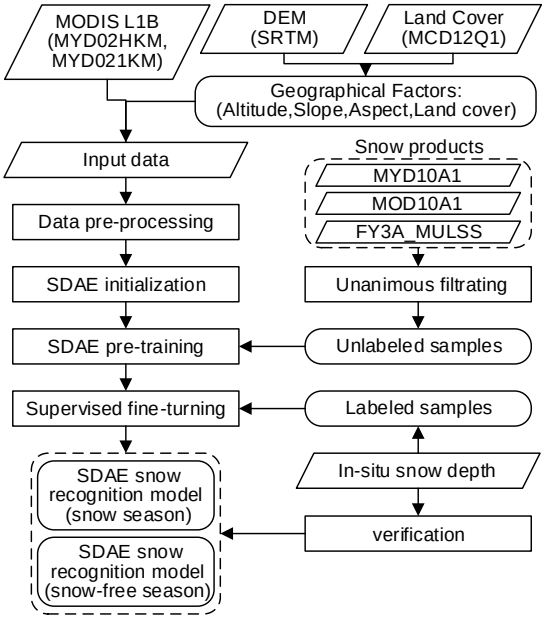

Figure 2: The flow chart of the SDAE training process for snow recognition

3.2.1 Stacked denoising Auto-Encoders

SDAE is one of the most mainstream deep learning structures, which is stacked by multiple Denoising Auto-Encoders (DAE) [Vincent, Larochelle, Lajoie et al. (2010)].

Essentially, DAE is a single-hidden layer artificial neural network, but it is not used for outputting a classification or fitting result, the purpose of DAE is to make the output possibly as similar as the input. In DAE's hidden layer, the random noise was added into input data, by randomly setting a part of input data as zero. According to the training rules, reconstructed data should be very close to the noise-free input data, thus DAE can acquire the ability to eliminate the noise disturbance of input data.

Let $x^{(K)}$ be the *K*-th training sample, which is an n-dimensional vector. Every input data x will be stochastically corrupted to $\tilde{x}$. The encoder then maps it into an m-dimensional feature vector h, which is called "encoding", as is shown in Eq. (1). Then the feature code h will be propagated from the hidden layer to output layer and transformed into an n-dimensional reconstructed vector $\hat{x}$. This process is called "decoding", as is shown in Eq. (2). The nonlinear activation function of DAE's neuron is sigmoid function, which is shown in Eq. (3):

$$h = f(x) = S\left(Wx + B_1\right) \tag{1}$$

$$\hat{x} = g(y) = S\left(W^T y + B_2\right) \tag{2}$$

$$S(t) = \frac{1}{1+e^{-t}} \tag{3}$$

Where W is the weight matrix of all neuron connections between input layer and hidden layer, W^T is the weight matrix between hidden layer and input layer. In DAE's structure, the weight matrix W^T is the transpose of the matrix W. B_1 and B_2 are the bias vectors. The iterative training process of DAE is searching for a group of optimal parameters matrixes θ^*, which are used for minimizing the cost function $L\left(W, B_1, B_2\right)$, which is shown in Eq. (4) to Eq. (6).

$$\theta^* = \arg\min_{W,b_1,b_2} L(W, b_1, b_2) \tag{4}$$

$$L\left(W, b_1, b_2\right) = \left[\frac{1}{K}\sum_{i=1}^{K} J\left(x^{(i)}, \hat{x}^{(i)}\right)\right] + \frac{\lambda}{2}\sum_{l=1}^{n-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_{l+1}}\left(w_{ij}^{(l)}\right)^2 + \beta\sum_{i=1}^{s_2} KL\left(\rho \middle\| \overline{\rho_i}\right) \tag{5}$$

$$J\left(x^{(i)}, \hat{x}^{(i)}\right) = \frac{1}{2}\left\|x^{(i)} - \hat{x}^{(i)}\right\|^2 = \frac{1}{2}\left\|x^{(i)} - g_{W^T,B_2}\left(f_{W,B_1}(x^{(i)})\right)\right\|^2 \tag{6}$$

The first item in the right-hand side of Eq. (5) is used for measuring the overall training error. The second item in the right-hand side of Eq. 5 is the regularization term, which is used to avoid overfitting, where λ is the weight coefficient, n is the sequence number of layers (in DAE's network structure n =3), s_l is the neuron number of the *l*-th layer. $w_{ij}^{(l)}$ is the connect weight between the *i*-th neuron of the *l*-th layer and the *j*-th neuron of *(l+1)*th layer. The third item is the sparse penalty term, where $KL(.)$ is the KL divergence between the mean activation rate of the *i*-th neuron in hidden layer $\overline{\rho_i}$ and the expected

activation rate ρ, β is the weight coefficient of the sparse penalty term, s_2 is the neuron number of DAE's hidden layer. The goal of this sparse penalty term is to make the number of activated neuron as less as possible, thus the hidden layer can learn the sparse representation of the input data. The error back propagation and gradient descent algorithm can be used for the iterative training process of the cost function minimizing.

SDAE is stacked by multiple DAEs. The training process of SDAE has two stages: unsupervised pre-training and supervised fine-tuning. In the unsupervised pre-training, the outputs of each hidden layer are used as the inputs of the next layer. The per-training of SDAE was layer by layer, only when the training of previous hidden layer meets the request can the next layer start training. This layer by layer training strategy is named as the "Greedy layer-wise training". This pre-training is unsupervised, which is not used for training classifier but used for training the hidden layers to learn how to extract linearly separable feature from the input. Once all the hidden layers have been well pre-trained, the supervised fine-tuning should be performed. In this process, an output layer would be added to the end of the hidden layers. This output layer is normally a supervised multi-class classifier, such as Softmax, which is used for classifying the final feature code into the specified categories. The training process is the same as the Back Propagation (BP) neural network. After a large number of supervised training, this SDAE network can be used as a high-efficiency classifier. With this network, the input data could be correctly classified.

3.2.2 Producing of training and testing sample

The unsupervised pre-training for SDAE needs amounts of unlabeled training samples. In this work, only unanimous results of various snow cover products can be used for making unlabeled samples: Checking the consistency of different snow products with the same location and the same date. Only if all the results are unanimous, this recognition result can be selected as a candidate unlabeled sample. By this filtrating process, these selected unlabeled samples should be very separable. It is very important to improve the pre-training performance.

The supervised fine-tuning needs sufficient labelled samples. In general, the in-situ observation records of snow depth are used for making labels. But in QT plateau, only using the in-situ observation records were not enough for training and testing, because the distribution density of meteorological stations is extremely low and the distribution is very unevenly. Thus, in this work, as a supplement to in-situ observation records, we manually selected a fair number of labelled samples from the above-mentioned unlabeled samples. Because of the high separability of the unlabeled samples, we could save a lot of labor and time.

3.2.3 Creation and initialization of SDAE for snow recognition

In this work, after amount of experiments, we designed the SDAE architecture with a suitable network depth, which has three hidden layers: First hidden layer with 128 units, the following two hidden layers with 32 units and 8 units. The input layer has 41 units, corresponding to the 33 MODIS bands and 8 geographical factors. The output layer has 3 units, corresponding to 3 outcomes: Snow cover, snow-free and cloud. The architecture

we designed for snow recognition is shown in Fig. 3.

The number of neurons in the input layer depends on the dimension of input data, which is the dimension of the training sample vector. The number of the first hidden layer's units is significantly higher than input layer, and the numbers of three hidden layers' units are progressive decreasing layer by layer. Thus, the input data can be first projected into a higher-dimensional feature code, and then this feature code would be fused and dimension reduction through the following hidden layers. Finally, the 8-dimensional feature code (outputted from the forth hidden layer) would be inputted to the output layer, the output layer is a supervised multi-class classifier, in this paper we choose a logistic regression classifier: Softmax. With this classifier, the feature code could be classified into three labels: Snow cover, cloud and snow-free land.

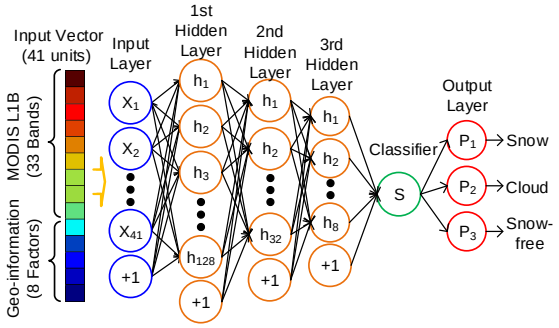

Figure 3: The framework diagram of the snow recognition model based on SDAE

3.2.4 SDAE training for different seasonal models

Because of the huge climatic and environmental difference among different seasons in QT plateau, in this work, two independent SDAE models were trained for snow recognition in different seasonal patterns: correspond respectively with the snow and snow-free season. We divided one year of QT Plateau into two seasons: Snow season and snow-free season, based on the statistical analysis of snow depth measurements from 96 ground meteorological stations over the QT Plateau from 2008 to 2013. The statistical result is shown in Fig. 4. The snow season of QT Plateau is from October to May of the following year, which contains almost all snow cover records of a year. The snow-free season is from June to September, which has very rare snow cover records.

In this work, a total of 10 million unlabeled samples and 50 thousand labelled samples were used to train these two seasonal SDAE models. These samples were extracted from the MODIS L1B data and associated geographic information data in the period from 2010 to 2011.The samples from October to next May were used to train snow season model, the samples from June to September were used to train snow-free season model.

Each seasonal SDAE model has gotten 2000 epochs pre-training and 2000 epochs fine-tuning. The whole training process took a long time of totally 16 h 41 min. Although the training process of SDAE was very time consuming, once the SDAE models were well trained, the snow products could be mapped very speedy: Taking less than 3 s to process one day's preprocessed multi-source data over the QT plateau.

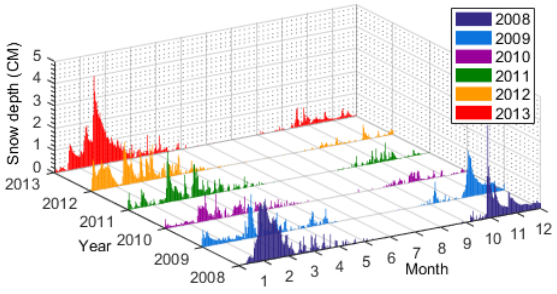

Figure 4: In-situ snow depth measurements over the QT Plateau from 2008 to 2013

3.3 Producing of snow cover product

The well-trained seasonal SDAE snow recognition model can be used for fast snow mapping and batch processing daily snow cover products. Fig. 5 shows the flow chart of the snow cover product producing process in the proposed method.

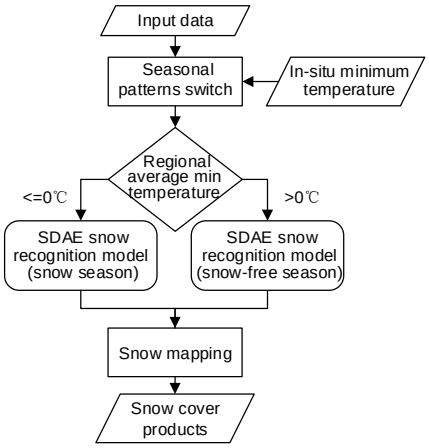

Figure 5: The flow chart of the snow cover product producing process

In the production process, the seasonal patterns switching cannot only rely on the divide of snow and snow-free season, because this monthly divide is too crude to be used in daily snow mapping. Therefore, the daily minimum air temperature from 96 ground meteorological stations over the QT Plateau form 2008 to 2013 is used as the judgement standard for seasonal patterns switching. Fig. 6 shows the monthly statistics analysis of daily minimum air temperature over the QT Plateau from 2008 to 2013. If the regional average daily minimum air temperature of QT Plateau is higher (or lower) than zero degree centigrade for five days in succession, the seasonal patterns should switch to snow-free season (or snow season), and the corresponding seasonal SDAE snow recognition model will be applied. The daily snow cover product based on this method were named as MYD_DL.

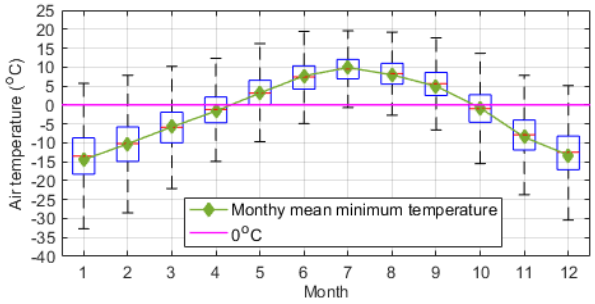

Figure 6: Monthly statistics analysis of daily minimum air temperature over the QT Plateau from 2008 to 2013

3.4 Validation metrics

To show the effectiveness of the proposed method, we compared a set of performance indicators of MYD_DL against the MOD10A1 and MYD10A1. In this work, four common statistical metrics were used: Overall Accuracy (OA), Precision, Recall and F-measure. These metrics rely on the common classification of four possible outcomes: true positive (TP), true negative (TN), false positive (FP) and false negative (FN). The in-situ snow depth measurements were used to verify the recognition results of snow cover products. Thus, in the coincidence matrix for this two-class classification problem (snow or snow-free), the classification outcomes from in-situ measurements were regarded as the true class; the classification outcomes of snow cover products were regarded as the predicted class. The detailed definitions of TP, TN, FP and FN are shown in Tab. 1.

Table 1: Coincidence matrix for snow cover products against in-situ observations

		In-situ (True Class)	
		Snow	Snow-free
Snow product (Predicted Class)	Snow	TP	FP
	Snow-free	FN	TN

The Recall (also be called hit rate or true positive rate) is estimated by dividing the correctly classified snow cover samples (the true positive count, TP) by all the samples indeed were snow cover (total positive count, TP+FN). The Precision is estimated by dividing the correctly classified snow cover samples (TP) by all the samples were classified as snow cover (TP+FP). The OA is estimated by dividing the total correctly classified snow and snow-free samples by the total number of samples. In this work, OA can mislead where the snow coverage rate is significantly less than snow-free coverage rate. For example, if only 10% of total samples are true snow cover, even classifying all these snow samples as snow-free, we could still have a high OA score: 90% [Rittger, Painter, Dozier et al. (2013)].

Therefore, we used the F-measure score. F-measure can balance Precision and Recall, penalize both underestimation (falsely classifying snow cover as snow-free) and overestimation (falsely classifying snow-free as snow cover), avoiding the influence of

the large proportion of true negative (snow-free) samples. All these performance indicators were calculated under cloud-free conditions. They are calculated as:

$$\begin{aligned}
&\text{Recall} = \frac{TP}{TP+FN} \\
&\text{Precision} = \frac{TP}{TP+FP} \\
&\text{Overall Accuracy}(OA) = \frac{TP+TN}{TP+TN+FP+FN} \\
&\text{F-measure} = 2\Big/\left(\frac{1}{\text{Precision}}+\frac{1}{\text{Recall}}\right) = \frac{2TP}{2TP+FP+FN}
\end{aligned} \tag{7}$$

4 Results and discussions

To demonstrate the superiority of the proposed method in this paper, we produced the daily snow products MYD_DL for QT plateau of the whole year of 2012. Then we quantitatively and qualitatively compared the precision of our snow cover product MYD_DL against the MOD10A1 and MYD10A1. The daily in-situ snow depth measurements in 2012 at 96 stations over the QT plateau were used to evaluate the recognition accuracy of these snow cover products.

4.1 Comparison of snow map effects

Fig. 7 shows the contrast of the overall effects between MYD_DL and MYD10A1. Fig. 7(a) is the true RGB colour composite image from MYD021KM (red channel: Band 1; green channel: Band 4; blue channel: Band 3) over the QT plateau on February 3rd, 2012. Fig. 7(b) and Fig. 7(c) are the mapping effects of MYD10A1 and MYD_DL. From Fig. 7, it can be seen that these snow cover products share similar characteristics, but the snow coverage in MYD_DL is relatively higher and has lower cloud coverage than MYD10A1.

Fig. 8 shows the local detail contrast of snow mapping results of the MYD_LD and MYD10A1. Fig. 8(a) is the RGB color composite image from MYD02HKM in local area of the QT Plateau on January 31st, 2012. It shows very clear outlines of the snow cover on top of the mountains, and there is almost no cloud cover this region. Fig. 8(b) and Fig. 8(c) are the snow mapping results of MYD_LD and MYD10A1 corresponding to Fig. 8(a). The comparison of Fig. 8(b) and Fig. 8(c) shows that the snow cover in MYD10A1 is relatively less than MYD_DL. Moreover, the shape of snow cover in MYD10A1 is fragmented, because a mass of small patches of snow cover were misjudged as cloud on mountains.

Fig. 8(d) is the RGB color composite image on February 3rd, 2012, which shows quite clear distinction between snow cover and cloud. Fig. 8(e) shows a good classification result of MYD_DL. But in Fig. 8(f), MYD10A1 tends to classify almost all snow cover as cloud. Fig. 8(g) is the color image on February 3rd, 2012, where gauzy cloud or mist can be seen in some areas. Although the atmospheric transmittance is relatively low in some gauzy cloud covered region, the land cover information can still be identified. In Fig. 8(h), some cloud covered areas can still be recognized in MYD_DL, which have high enough atmospheric transmittance. But in Fig. 8(i), MYD10A1 tends to ignore all the gauzy cloud covered regions. The reason for this is most probably that the MODIS cloud mask product (MYD35) was applied to the snow recognition methods of MYD10A1. In this cloud mask product, different kinds of clouds, fog and haze were not separated.

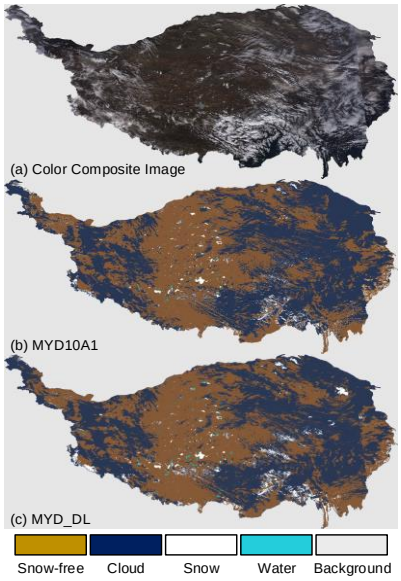

Figure 7: Overall effects of MYD_LD and MYD10A1 over the QT plateau on February 3rd, 2012: (a) Color composite image from MYD02HKM; (b) MYD10A1; (c) MYD_DL

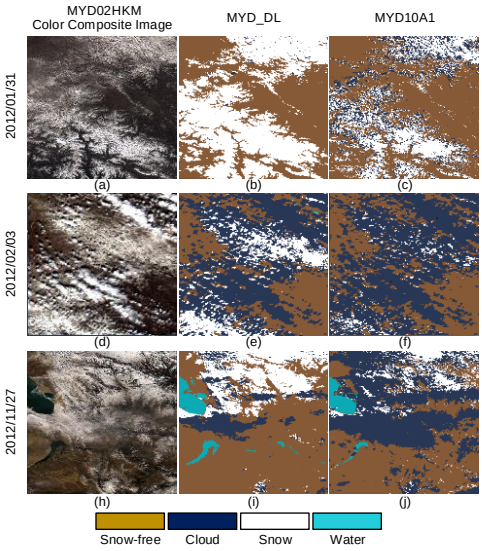

Figure 8: Local detail comparison between MYD_LD and MYD10A1 respectively on January 31, February 3 and November 27 of 2012. Left column (a, d, g): RGB composite images from MYD02HKM. Middle column (b, e, h): Snow recognition results of MYD_DL. Right column (c, f, i): Snow recognition results of MYD10A1

4.2 Comparison of weekly mean overall accuracies

Fig. 9 shows a time series of weekly mean overall accuracy (OA) curves of MOD10A1, MYD10A1 and MYD_DL over the QT Plateau in 2012. Fig. 9 shows that the MYD_DL curve is obvious higher than MYD10A1 and MOD10A1. The annual mean OA supports this comparison result: the annual mean OA of MYD_DL is 98.95%, higher than MOD10A1 (98.66%) and MYD10A1 (97.81%). In Fig. 9, all the OA curves have a similar variation tendency: OA is higher in snow-free season than in snow seasons. There are two main reasons. On the one hand, new snow cover stop generating in snow-free season, meanwhile, the permanent snow at high altitudes of mountains are hard to thaw in summer, so the distribution of snow cover over the QT plateau is very stable in summer, which can be readily detected. On the other hand, OA can mislead where the snow cover records from in-situ measurements (true positives) are significantly less than snow-free samples (true negative samples), however in snow-free season, there are hardly any in-situ snow cover records. Therefore, the OA curves make little sense in snow-free season. Despite all this, it is noteworthy that MYD_DL performs well in snow season. Fig. 9 shows that the OA curve of MYD_DL is higher than MYD10A1 and MOD10A1 in the period from late January to February. It further confirms the superiority of deep learning based snow recognition method in disposing of multi-factor coupling classification problems.

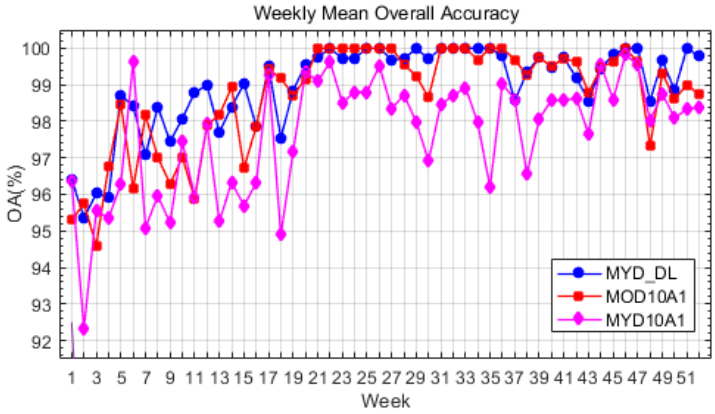

Figure 9: Weekly mean overall accuracy (OA) curves of MOD10A1, MYD10A1 and MYD_DL over the QT Plateau in 2012

4.3 Reduction of cloud contamination

Fig. 10(a) shows the time series of weekly mean cloud coverage rates of MYD_DL, MYD10A1 and MOD10A1, over the QT Plateau in 2012. It shows that the MYD_DL can remove more cloud obscuration than MYD10A1, and similar with MOD10A1, without reducing the overall accuracy (shown in Fig. 9). Fig. 10(b) shows the time series of weekly mean snow coverage rate of MYD_DL, MYD10A1, and MOD10A1. The curves show the MYD_DL detected more snow cover areas than MYD10A1 and MOD10A1 in the QT Plateau. Associate Fig. 10(a) with Fig. 10(b), it also proves that the effect of removing cloud is significant of we proposed method. Fig. 10(c) shows the time series of weekly mean snow-free land coverage rates of MYD_DL, MYD10A1, and MOD10A1.

The curves seem to be very close, because most identification confusions in snow cover products are between cloud and snow cover.

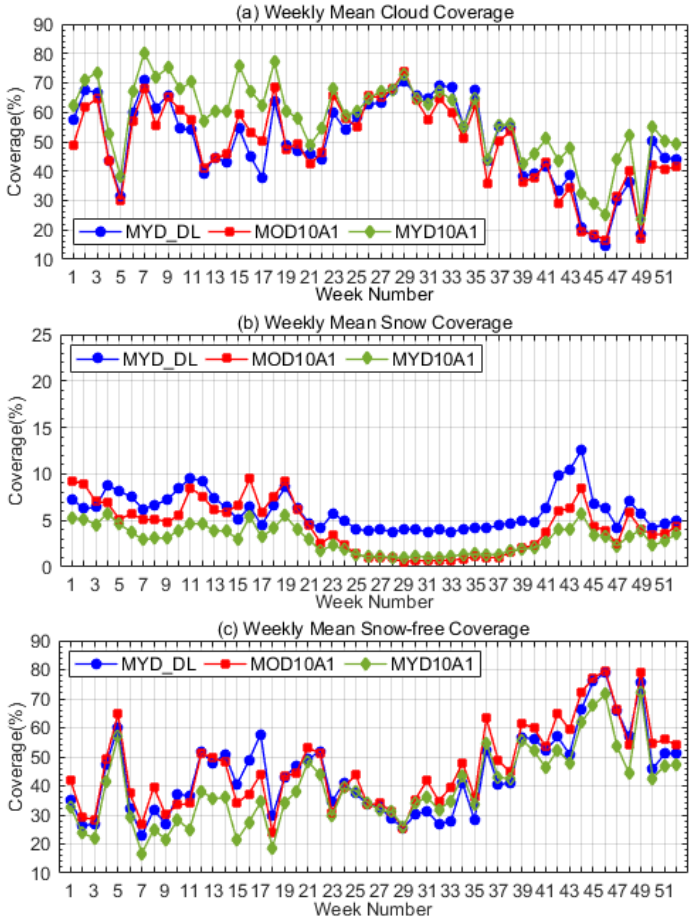

Figure 10: Weekly mean coverage rate comparison between various snow products over the QT Plateau in 2012: (a) Cloud, (b) Snow cover, (c) Snow-free land

4.4 Comparison of performance indices

We calculated the annual averages of all the above performance indices of MYD_DL, MYD10A1 and MOD10A1, which are shown in Tab. 2. It can be seen that MYD_DL has the best performance of OA (98.95%), Recall (28.83%) and Precision (69.75). Meanwhile, MYD_DL has the highest annual mean snow coverage rate (6.02%), and a lower annual mean cloud coverage rate (49.80%) than MYD10A1 (57.60%). It shows that MYD_DL can effectively reduce the cloud cover disturbance. Although the annual mean cloud coverage rate of MOD10A1 (49.17%) is the lower than MYD_DL, but in consideration of the observation time are different between MYD_DL and MOD10A1, this disadvantage can be accepted. The high classification accuracy and low cloud coverage rate of MYD_DL show that our snow recognition method has the obvious advantage of solving snow/cloud confusion problem in mountainous region.

Table 2: Performance indices of various snow cover products

Snow products	Annual mean accuracy (%)				Annual mean coverage (%)		
	OA	Recall	Precision	F-measure	Cloud	Snow	Snow-free
MYD10A1	97.81	54.95	31.28	39.87	57.60	3.02	39.38
MOD10A1	98.66	74.19	56.23	63.98	**49.17**	4.32	**46.51**
MYD_DL	**98.95**	**78.43**	**69.75**	**73.84**	49.80	**6.02**	44.18

5 Conclusion

The Snow cover in QT plateau is very important to the water circulation, weather evaluation and global climate change. However, due to its complex topography and high altitude, the recognition accuracies of existing snow products in QT plateau are significantly lower than other regions. In addition, due to the Band 6 in Aqua MODIS no longer functions, the overall accuracy of MYD10A1 is significantly lower than MOD10A1.

In this paper, we proposed a deep learning based snow recognition method to reduce the snow/cloud confusions and improve the accuracies of snow mapping in mountainous region. The SDAE framework was employed to train a pixel-level classifier for snow recognition, this SDAE snow recognition model can effectively fuse the high-dimensional input data which include 33 spectral bands and 8 geographical influence factors and can extract the linearly separable feature vectors from the input data. Finally, these feature vectors can be easily classified into three categories: snow, cloud and snow-free land. Compared to the traditional snow cover recognition method based on NDSI, this method can better adapt to mountainous regions with complex topographical and geographic characteristics, without relying so heavily on some specific bands, like Band 6 in MODIS. Moreover, two independent SDAE models were trained for snow recognition in wet and dry seasons respectively, in response to the huge climatic and environmental difference between wet season and dry season in QT plateau. These seasonal snow recognition models can better express the spectral and geographical differences between different seasons in QT plateau. Experiments show that the overall accuracy and snow recognition accuracy of the proposed method are significant higher than MODIS snow cover products MOD10A1 and MYD10A1. Furthermore, the snow coverage rate of our snow cover products is the highest. The above results prove the efficiency and superiority of this method.

Acknowledgement: This research was supported by National Natural Science Foundation of China (Grant Nos. 41661144039, 91337102, 41401481), and Natural Science Foundation of Jiangsu Province of China (Grant No. BK20140997).

References

Allen, R. C. J.; Durkee, P. A.; Wash, C. H. (1990): Snow/cloud discrimination with multispectral satellite measurements. *Journal of Applied Meteorology*, vol. 29, no. 10, pp.

994-1004.

Bi, Y.; Xie, H.; Huang, C.; Ke, C. (2015): Snow cover variations and controlling factors at upper Heihe River Basin, Northwestern China. *Remote Sensing*, vol. 7, no. 6, pp. 6741-6762.

Chen, Y.; Lin, Z.; Zhao, X.; Wang, G.; Gu, Y. (2017): Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, vol. 7, no. 6, pp. 2094-2107.

Dankers, R.; De Jong, S. M. (2004): Monitoring snow-cover dynamics in Northern Fennoscandia with SPOT vegetation images. *International Journal of Remote Sensing*, vol. 25, pp. 2933-2949.

Dariane, A. B.; Khoramian, A.; Santi, E. (2017): Investigating spatiotemporal snow cover variability via cloud-free MODIS snow cover product in Central Alborz Region. *Remote Sensing of Environment*, vol. 202, pp. 152-165.

Dietz, A. J.; Kuenzer, C.; Gessner, U.; Dech, S. (2014): Remote sensing of snow-A review of available methods. *International Journal of Remote Sensing*, vol. 33, no. 13, pp. 4094-4134.

Dietz, A. J.; Wohner, C.; Kuenzer, C.; (2014): European snow cover characteristics between 2000 and 2011 derived from improved MODIS daily snow cover products. *Remote Sensing*, vol. 4, no. 8, pp. 2432-2454.

Dobreva, I. D.; Klein, A. G. (2011): Fractional snow cover mapping through artificial neural network analysis of MODIS surface reflectance. *Remote Sensing of Environment*, vol. 115, no. 12, pp. 3355-3366.

Dong, C.; Menzel, L. (2016a): Producing cloud-free MODIS snow cover products with conditional probability interpolation and meteorological data. *Remote Sensing of Environment*, vol. 186, pp. 439-451.

Dong, C.; Menzel, L. (2016b): Improving the accuracy of MODIS 8-day snow products with in situ temperature and precipitation data. *Journal of Hydrology*, vol. 534, pp. 466-477.

Dozier, J.; Painter, P. H. (2004): Multispectral and hyperspectral remote sensing of alpine snow properties. *Annual Review of Earth & Planetary Sciences*, vol. 32, pp. 465-494.

Hall, D. K.; Riggs, G. A. (2007): Accuracy assessment of the MODIS snow products. *Hydrological Processes*, vol. 21, pp. 1537-1547.

Hall, D. K.; Riggs, G. A.; Foster, J. L.; Kuma, S. V. (2010): Development and evaluation of a cloud-gap-filled MODIS daily snow-cover product. *Remote Sensing of Environment*, vol. 114, no. 3, pp. 496-503.

Hall, D. K.; Riggs, G. A.; Salomonson, V. V. (1995): Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data. *Remote Sensing of Environment*, vol. 54, no. 2, pp.127-140.

Hall, D. K.; Riggs, G. A.; Salomonson, V. V.; Digirolamo, N. E.; Bayr, K. J. (2002): MODIS snow-cover products. *Remote Sensing of Environment*, vol. 83, no. 1, pp. 181-194.

Hinton, G. E.; Osindero, S.; Teh, Y. (2016): A fast learning algorithm for deep belief nets. *Neural Computation*, vol. 18, no. 7, pp. 1527-1554.

Hori, M.; Sugiura, K.; Kobayashi, K. (2017): A 38-year (1978-2015) Northern

Hemisphere daily snow cover extent product derived using consistent objective criteria from satellite-borne optical sensors. *Remote Sensing of Environment*, vol. 191, pp. 402-418.

Krizhevsky, A.; Sutskever, I.; Hinton, G. E. (2012): ImageNet classification with deep convolutional neural networks. *International Conference on Neural Information Processing Systems*, vol. 60, no. 2, pp. 1097-1105.

Kruger, N.; Janssen, P.; Kalkan, S.; Lappe, M.; Leonardis, A. et al. (2013): Deep hierarchies in the primate visual cortex: what can we learn for computer vision? *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 35, no. 8, pp. 1847-1871.

Li, W.; Fu, H; Yu, L.; Gong, P.; Feng, D. et al. (2016): Stacked autoencoder-based deep learning for remote-sensing image classification: A case study of African land-cover mapping. *International Journal of Remote Sensing*, vol. 37, no. 23, pp. 5632-5646.

Liu, X.; Jin, X.; Ke, C. (2014): Accuracy evaluation of the IMS snow and ice products in stable snow covers regions in China. *Journal of Glaciology and Geocryology*, vol. 36, no. 3, pp. 500-507

Marmanis, D.; Datcu, M.; Esch, T.; Stilla, U. (2016): Deep learning earth observation classification using ImageNet pretrained networks. *IEEE Geoscience & Remote Sensing Letters*, vol. 13, no. 1, pp. 105-109.

Pu, Z. X.; Xu, L. (2009): MODIS/Terra observed snow cover over the Tibet Plateau: Distribution, variation and possible connection with the East Asian Summer Monsoon (EASM). *Theoretical & Applied Climatology*, vol. 97, no. 3-4, pp. 265-278.

Pu, Z. X.; Xu, L.; Salomonson, V. V. (2007): MODIS/Terra observed seasonal variations of snow cover over the Tibetan Plateau. *Geophysical Research Letters*, vol. 34, no. 6, pp. 137-161.

Qin, D.; Liu, S.; Li, P. (2006): Snow cover distribution, variability, and response to climate change in western China. *Journal of Climate*, vol. 19, pp. 1820-1833.

Qin, X. J.; Sun, J; Chen, T. (2015): Study on spatiotemporal variation of temperature and precipitation in Qinghai-Tibetan Plateau from 1974 to 2013. *Journal of Chengdu University (Natural Science Edition)*, vol. 34, no. 2, pp. 191-195.

Riggs, G. A.; Hall, D. K. (2016): MODIS snow products collection 6 user guide. https://nsidc.org/sites/nsidc.org/files/files/MODIS-snow-user-guide-C6.pdf**.**

Rittger, K.; Painter, T. H.; Dozier, J. (2013): Assessment of methods for mapping snow cover from MODIS. *Advances in Water Resources*, vol. 51, pp. 367-380.

Rosenthal, W.; Dozier, J. (1996): Automated mapping of montane snow cover at subpixel resolution from the Landsat thematic mapper. *Water Resources Research*, vol. 32, no. 1, pp. 115-130.

Salakhutdinov, R.; Hinton, G. E. (2009): Deep boltzmann machines. *Journal of Machine Learning Research*, vol. 5, no. 2, pp. 1967-2006.

Sirguey, P.; Mathieu, R.; Arnaud, Y. (2009): Subpixel monitoring of the seasonal snow cover with MODIS at 250 m spatial resolution in the Southern Alps of New Zealand: Methodology and accuracy assessment. *Remote Sensing of Environment*, vol. 113, no. 1, pp. 160-181.

Tang, B.; Shrestha, B.; Li, Z.; Liu, G.; Ouyang, H. et al. (2013): Determination of

snow cover from MODIS data for the Tibetan Plateau region. *International Journal of Applied Earth Observation and Geoinformation*, vol. 21, pp. 356-365.

Thirel, G.; Salamon, P.; Burek, P.; Kalas, M. (2013): Assimilation of MODIS snow cover area data in a distributed hydrological model using the particle filter. *Remote Sensing*, vol. 5, no. 11, pp. 5825-5850.

Thompson, J. A.; Paull, D. J.; Lees, B. G. (2015): An improved liberal cloud-mask for addressing snow/cloud confusion with MODIS. *Photogrammetric Engineering & Remote Sensing*, vol. 81, no. 2, pp. 119-129.

Vincent, P.; Larochelle, H.; Lajoie, I.; Bengio, Y.; Manzagol, P. (2010): Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research*, vol. 11, no. 12, pp. 3371-3408.

Yang, J.; Jiang, L.; Ménard, C. B.; Luojus, K.; Lemmetyinen, J. et al. (2015): Evaluation of snow products over the Tibetan Plateau. *Hydrological Processes*, vol. 29, pp. 3247-3260.

Yang, J.; Jiang, L.; Shi, J.; Wu, S.; Sun, R. et al. (2014): Monitoring snow cover using Chinese meteorological satellite data over China. *Remote Sensing of Environment*, vol. 143, pp. 192-203.

Yang, K.; Wu, H.; Qin, J.; Lin, C.; Tang, W. et al. (2014): Recent climate changes over the Tibetan Plateau and their impacts on energy and water cycle: A review. *Global and Planetary Change*, vol. 112, no. 1, pp. 79-91.

Zhang, L.; Zhang, L.; Du, B. (2016): Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience & Remote Sensing Magazine*, vol. 4, no. 2, pp. 22-40.

Zhang, Y.; Cao, T.; Kan, X.; Wang, J.; Tian, W. (2016): Spatial and temporal variation analysis of snow cover using MODIS over Qinghai-Tibetan Plateau during 2003-2014. *Journal of the Indian Society of Remote Sensing*, vol. 45, no. 5, pp. 1-11.

Zhang, Y.; Kan, X.; Ren, W.; Cao, T.; Tian, W. et al. (2017): Snow cover monitoring in Qinghai-Tibetan Plateau based on Chinese Fengyun-3/VIRR data. *Journal of the Indian Society of Remote Sensing*, vol. 45, no. 2, pp. 271-283.

Zhu, L.; Xiao, P.; Feng, X.; Zhang, X.; Wang, Z. et al. (2014): Support vector machine-based decision tree for snow cover extraction in mountain areas using high spatial resolution remote sensing image. *Journal of Applied Remote Sensing*, vol. 8, no. 1, pp. 396-403.