

Interpreting transition and emission probabilities from a Hidden Markov Model of remotely sensed snow cover in a Himalayan Basin

Sean Minhui Tashi Chua^{a,b}, Dave Penton^b and Albert Van Dijk^a

^a*Fenner School of Environment and Society, Australian National University, Canberra*

^b*CSIRO Land and Water, Canberra*

Email: sean.m.t.chua@gmail.com

Abstract: Remote sensing is often used to monitor snow cover in areas without sufficient in-situ observations. However, a common problem with remotely sensed snow cover products is the mis-classification of snow cover due to obscuration of the land surface by cloud, along with the similar spectral characteristics of snow and cloud. In a two part preliminary investigation we assessed the ability of a Hidden Markov Model (HMM) to reduce classification errors in optical snow cover mapping along with the transition and emission probabilities output from the model, to our knowledge for the first time documented. This research focuses on the latter and was conducted within the Sustainable Development Investment Portfolio (SDIP) at CSIRO.

A Hidden Markov model can utilise a series of input observations and calculate the probability that they represent the ground state. These probabilities are then used to model the most likely series of states, this effectively provides a dynamic filter that can mitigate the problems faced when remotely sensing snow in mountainous areas. As part of a larger project, we applied this approach to snow cover mapping over a single sub-basin in the Himalayas in Eastern Nepal based on imagery from the (MODIS) instruments on the Terra and Aqua satellites. This study analysed spatially mapped transition and emission probabilities extracted from a two state Hidden Markov Model.

The ability of a Hidden Markov model to employ a dynamic filter that can utilise entire sequences of observations will likely offer improved accuracy when compared to other time-series filtering methods. This is because it is able to utilise a much larger number of observations without compounding losses in accuracy or causing reductions in temporal resolution. Our probability analysis shows the potential for a HMM approach to provide a robust and flexible method for processing ‘noisy’ data such as remotely sensed snow cover measurements. The improvement of spatio-temporal snow cover measurements has broader implications for hydrological modelling, particularly in countries dependent on snow melt for subsistence agriculture and hydroelectric facilities such as Nepal. Improvements also benefit the long-term analysis of snow cover trends which are an important proxy for assessing the impacts of climate change in sensitive mountain areas. Further studies should apply this method to multiple study sites and quantitatively compare it to other cloud-cover reduction techniques for snow cover imagery.

Keywords: *Hidden Markov Model (HMM), machine learning, remote sensing, snow, cryosphere*

1. INTRODUCTION

The remote sensing of snow cover typically uses visible and near-infrared satellite-borne sensors such as LANDSAT or the MODerate Resolution Imaging Spectroradiometer (MODIS). The Normalised difference snow index (NDSI) allows snow to be discriminated from other land surfaces through its high reflectance in visible wavelengths (Hall *et al.*, 2010). Due to their similar spectral signature clouds are the primary barrier to accurate measurements of snow cover using optical sensors. This is a significant problem in areas with persistent cloud-cover such as the Himalayas. Most solutions utilise a temporal filter on a pixel-by-pixel basis that prioritises observations other than cloud over a set window of time (Gafurov and Bárdossy, 2009). Whilst this can result in a reduction of classification errors, this type of filter treats sequential observations using static, conditional rules that may impact the temporal resolution of observations and overly simplify the complexities of the snow-cloud-sensor relationship. In this study we implement a Hidden Markov Model (HMM) to MODIS snow cover observations as an alternative filtering method and investigate the transition and emission probabilities output by this model. The Himalayas provide a challenging study site for the remote sensing of snow cover due to the monsoonal cloud cover they experience thus providing a suitable study site for this research (Immerzeel *et al.*, 2009). Furthermore, the regions vulnerability to climate change, dependence on snow melt for drinking and irrigation and poverty provide strong motivations for research in this area.

A Markovian approach to modelling provides a dynamic alternative to traditional time-series filtering by employing a state based model that can interpret sequences of observations to try and accurately describe the stochastic processes behind them (Rabiner, 1989). These differ to the static temporal filters described above because they provide a flexible method of processing observations that can adapt to changes in the system. There are several key components that make up a Markov model. The order refers to the number of states taken into consideration when modelling the following state. A first-order Markov chain infers that the next state depends only on the current state and no previous states, this can be expanded to second-order chains and so on. Setting a finite number of previous states for the model to consider avoids creating an intractable problem where the length of time required to calculate a solution is too long to be useful (Ghahramani, 2001). A relevant example of a three-state Markov model applied to the weather as outlined by Rabiner (1989) lists three possible states which can be directly observed; rainy, cloudy or sunny and a corresponding matrix outlining the transition probabilities between each of them. A transition matrix outlines the probabilities associated with possible state changes within a Markov model. By processing the sequence of states one can determine the probability of state transitions within the model, this includes the probability of a state persisting over a given time.

In the above example the state of the system can be directly observed, a HMM becomes relevant when the state cannot be directly observed. In these cases only the sequence of stochastic outputs that relate to the hidden states can be observed (Rabiner and Juang, 1986). There are many different permutations of a HMM, these can include various combinations of states, orders, filters and algorithms. The following is a broad description of the general components in the context of this study. A HMM can interpret a noisy sequence of observations, associate these to non-directly observable states via probability distributions and then model the most likely series of states based on this information. In a HMM only the output, dependent on the hidden state, is observable and thus probability distributions must be based on these visible observations. Thus at a single time step in a two-state HMM one can infer the probability of transitioning between either state and the probability of a state's emission of an observation, the latter are known as emission probabilities. A conceptual representation of a two-state HMM is shown in Figure 1 in which the relationship between the states and observations of a model can be viewed through the possible transition and emission probability pathways. In the case of remote sensing the series of radiometric measurements are observable via satellite, these measurements are related to the chain of states that are not directly observable, for example; a series of hidden states. In this case these could be whether the ground is covered by vegetation or bare soil regardless of whether the returned radiometric measurement is that of cloud cover. The premise of a HMM's relevance to remote sensing is that there is a series of discretely timed and valued radiometric observations from satellites that represent states which cannot be directly observed from space (Viovy and Saint, 1994).

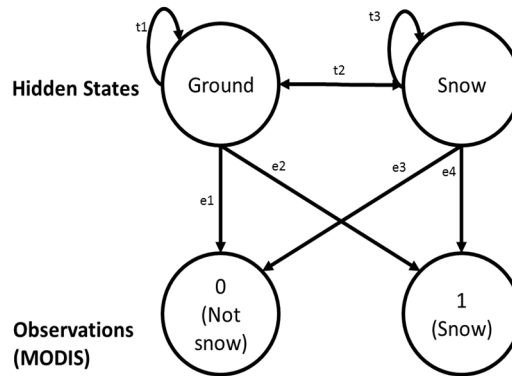


Figure 1. Conceptual diagram of a HMM (t^x = state transition probability, e^x = observation emission probability)

2. METHODOLOGY

2.1. Study area

The study site is the Dudh Koshi, a sub-basin of the Koshi river basin in the Eastern Himalayas (Figure 2). The Koshi river basin is the largest in Nepal, draining around 70,000 km² that includes the Eastern Nepalese Himalaya as well as part of Tibet and northern India (Agarwal *et al.*, 2014, Dixit *et al.*, 2009). Elevation ranges from 65 to 8848 m meters above mean sea level, with elevation increasing towards the north (Agarwal *et al.*, 2014, Shrestha *et al.*, 2015). The majority of precipitation in the basin occurs in the monsoon between June and September and its spatial distribution is strongly influenced by the mountain ranges in the north (Shrestha *et al.*, 2015).

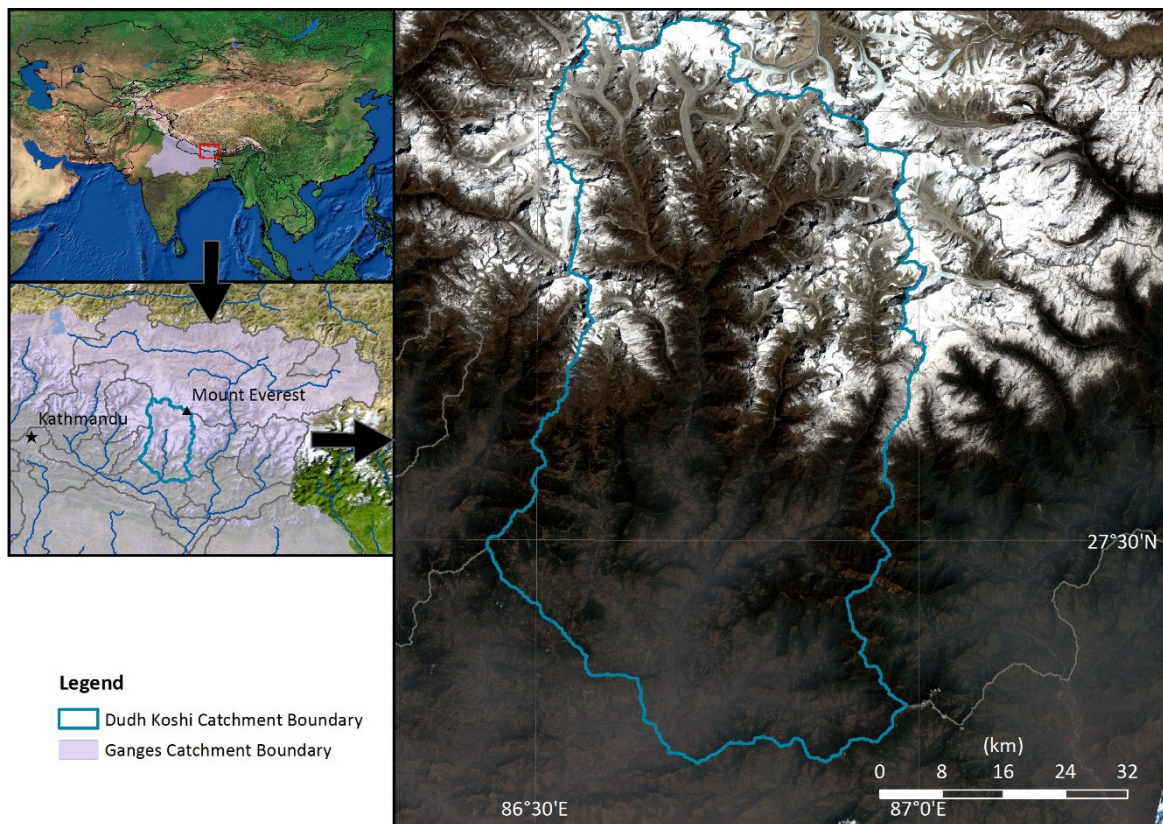


Figure 2. Study site, Koshi basin and sub-basins, Nepal (Summer). Blue outline denotes the catchment boundary over RGB optical image (adapted from Nepal *et al.*, 2015)

2.2. Datasets

Daily and 8-day aggregate snow cover data collected by MODIS (NSIDC, 2016) is available from the NSIDC (National Snow and Ice Data Center). Snow cover is calculated from raw spectral data by applying the NDSI (1) to the existing MODIS bands (band 4 = 545-565 nm, band 6 = 1628-1652 nm). Higher spatial resolution imagery is collected by LANDSAT however the temporal resolution is lower at 16 days (Goward *et al.*, 2001, Hall *et al.*, 2002). Images recorded daily have a greater chance of recording snow fall and melting events compared to the 8-day MODIS products and 16 day LANDSAT imagery (Gafurov and Bárdossy, 2009). Previous MODIS snow cover products recorded snow cover fractions on a scale of 0-100 however the latest version, collection 6, has shifted to binary snow cover measurements, thus this study focuses on the latter (NSIDC, 2016). Daily remotely sensed snow cover images are more susceptible to viewing angle problems and these represent a source of error. Version 6 of the products used incorporates both a solar zenith and short-wave infrared reflectance filter that reduce uncertainty in snow cover measurements related to high and low sun angles.

MODIS Terra images the Earth in descending orbit at 10:30 am and Aqua at 1:30 pm in an ascending orbit, these times are local time in Nepal. The derived products are available as $10^{\circ} \times 10^{\circ}$ (approximately 1200 km²) tiles at 500-meter resolution in sinusoidal projection. The Data processed was from 2002 until 2015 for both Terra and Aqua.

$$NDSI = \frac{band\ 4 - band\ 6}{band\ 4 + band\ 6} \quad (1)$$

2.3. Model implementation

Figure 3 depicts how the two-state HMM interacts with the MODIS data processed in this study. The model works on a pixel-by-pixel basis, processing the entire sequence of MODIS observations at each pixel to model the transition and emission probabilities for that pixel. There is a total of 17329 MODIS pixels within the study site. The states, corresponding symbols and initial starting probabilities of the model are shown in Table 1. A HMM model is comprised of two interacting algorithms outlined in Rabiner (1989). First, the Baum-Welch algorithm adjusts the initial state transition probabilities (Table 2) of the model using an input observation sequence extracted from the MODIS snow cover images, training the model. These are then used to calculate new emission probabilities (Table 3), this occurs for each pixel. This study focuses on interpreting the updated probabilities output by the Baum-Welch algorithm.

The Viterbi algorithm is a method of modelling the ‘hidden’ part, the non-directly observable states, within a HMM, this is effectively the most likely explanation of an observation sequence. This algorithm calculates the optimal sequence of states for the given sequence of observations using the optimised transition probabilities

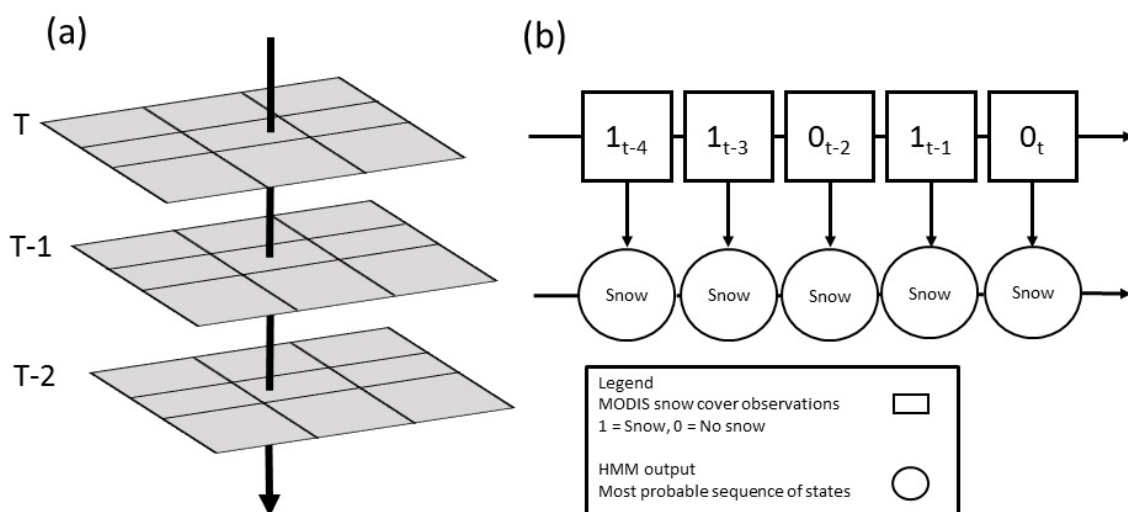


Figure 3. a) Conceptual diagram of the HMM analysis, T = MODIS snow cover image for a single day

(b) Conceptual diagram of a HMM processing a series of observations at a MODIS pixel, t = daily time step

and emission probabilities calculated by the Baum-Welch algorithm along with the sequence of observations. This part of the HMM is not addressed in this study.

Table 1. HMM states, symbols and initial probabilities

States	snow	no snow / unknown
Symbols	1	0
Initial probabilities	0.5	0.5

Table 2. Transition probabilities (1 = snow, 0 = no snow/ unknown)

	1	0
1	0.8	0.1
0	0.1	0.8

Table 3. Emission probabilities (1 = snow, 0 = no snow / unknown)

	1	0
snow	0.50	0.50
no snow / unknown	0.51	0.49

3. RESULTS

Each HMM run implements the Baum-Welch algorithm, producing an updated set of transition and emission probabilities for each pixel within the study site based on the series of observations that are input for the pixel. The updated probabilities for each transition can be spatially represented across the study sites as shown in Figure 4 and Figure 5 where sets of four graphs for both transition and emission probabilities are visible. It should be noted that almost all snowfall in the sub-basin occurs in the northern, higher elevation parts. The transition probabilities represent the probability of a specific state transition occurring within the HMM for a specific pixel. The emission probabilities refer to the relationship between the hidden state in the model and the observations as provided by the input data. Emission probabilities are the probability of an emission (observation) accurately representing the internal hidden state of the model for that specific state transition.

Each of the four graphs within a set represents the probabilities for a specific transition or emission. Transition probability maps for all inputs are generally consistent across the study site showing high probabilities of remaining in either state (1 to 1, 0 to 0) and low probabilities for state changes (0 to 1, 1 to 0). Slight artefacts are visible in the north-east corner of the sub-basin for the 0 to 0 state transitions, this is likely due to a large shadowed valley surrounded by mountains effecting the sensors ability to record information (Figure 2), this in-turn influences the model.

Emission probabilities for the study's HMM results generally have spatial variation and differences between the state transitions (Figure 5). For the 1 to 1 transition the emission probabilities are mostly low with slightly higher probabilities across the high elevation areas, these likely represent the increased chance of snow cover at increasing elevations. The spatial distribution of emission probabilities in the 1 to 0 transition show that lower elevation areas have higher probabilities and the very high elevation peaks have low probabilities. The 0 to 1 transition shows higher emission probabilities in the high elevation areas of the sub-basin, this indicates that both MODIS daily snow cover products do a better job of detecting transition to snow values from no snow / unknown in these areas. Finally, the 0 to 0 state transition show lower probabilities in the higher elevation regions and higher probabilities in the lower, snow free areas.

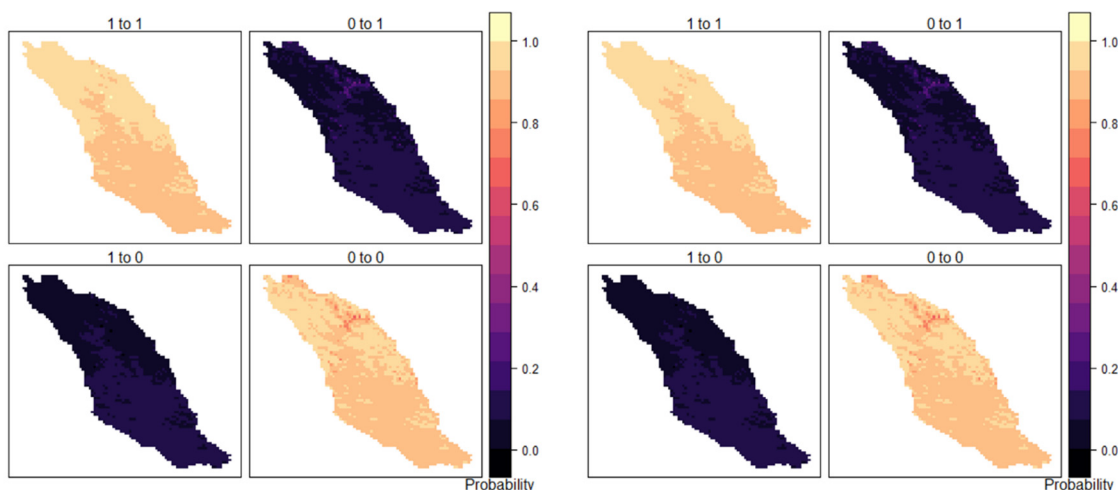


Figure 4. MOD10A1 (left) and MYD10A1 (right) transition probabilities for a two state HMM in the Dudh Koshi sub-basin (2002-2015) 1 = Snow, 0 = No snow / unknown

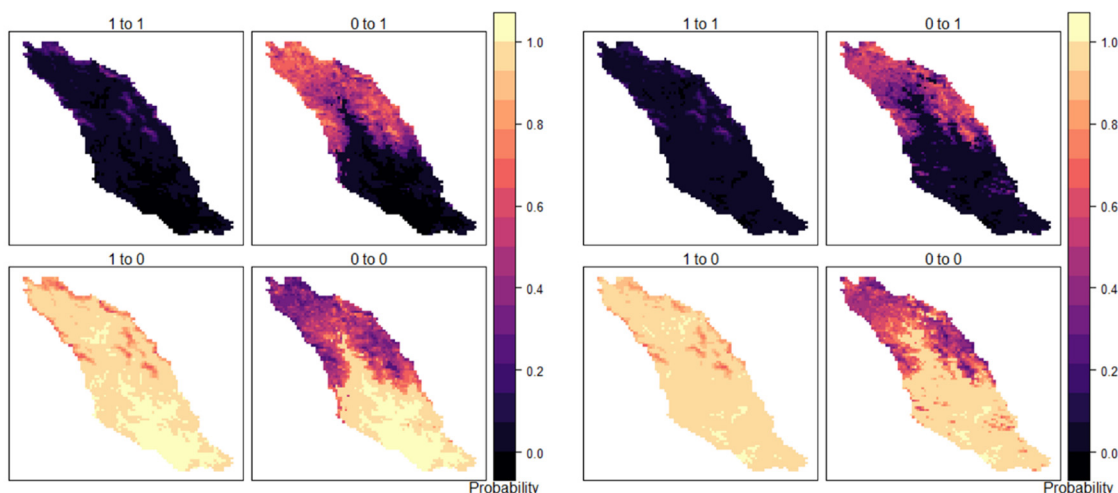


Figure 5. MOD10A1 (left) and MYD10A1 (right) emission probabilities for a two state HMM in the Dudh Koshi sub-basin (2002-2015) 1 = Snow, 0 = No snow / unknown

4. DISCUSSION

The transition probabilities in Figure 4 show that the model gives high probabilities for remaining in a snow covered or not snow-covered state and low probabilities for transitioning between snow or not snow. This is evidence of the smoothing ability of the HMM in that it reduces the amount of false transitions between snow and not snow values that typically occur due to cloud obscuration. The emission probability maps in Figure 5 represent the chance of a particular observation (MODIS snow cover products) occurring whilst in a certain state, either 1 (snow) or 0 (no snow / unknown). Overall they show that the MODIS snow cover products mostly agree with the HMM output for the 1 to 0 transition, except for very high elevation areas where snow cover is permanent. As snowfall does not occur in the southern part of the basin, the probability is almost 0 for both transitions to a snow state, 1 to 1 and 0 to 1. Observations tend to agree in the lower elevations for the transitions to a no snow / unknown state, represented as 0 to 0 and 1 to 0, as this part of the study site very rarely has snow observations. Spatially mapping HMM probabilities for remotely sensed data in this way has not been conducted before and this may provide a valuable new method of visualising a sensors relationship with a particular model and land surface type.

5. CONCLUSIONS AND RECOMMENDATIONS

As part of a broader project a HMM was applied to remotely sensed snow cover data of an Eastern Himalayan basin, this study analyses the spatially mapped emission and transition probabilities of the model. These analyses provide insight into the ‘thinking’ of machine learning models within the context of remote sensing

through the ability to spatially represent the state transition probabilities and their relationship with the input data over the study site. We found that the modelled transition and emission probabilities demonstrated that the model was more stable, i.e. there was less false variation in ground state, than the state transition probabilities extracted from the cloud contaminated data provided as input. As this was a preliminary investigation further research into remote sensing applications for HMM emission and transition probabilities should be conducted. HMMs could be used to improve other remotely sensed products that suffer from cloud contamination.

ACKNOWLEDGEMENTS

This work was supported by CSIRO and ANU. This publication contributes to the Sustainable Development Investment Portfolio and is supported by the Australian aid program (<http://research.csiro.au/SDIP>).

REFERENCES

- Agarwal, A., Babel, M. S. and Maskey, S. (2014). Analysis of future precipitation in the Koshi river basin, Nepal. *Journal of Hydrology*, 513, 422-434.
- Dixit, A., Upadhyaya, M., Dixit, K., Pokhrel, A. and Rai, D. R. (2009). Living with water stress in the hills of the Koshi Basin, Nepal. *Living with water stress in the hills of the Koshi Basin, Nepal*.
- Gafurov, A. and Bárdossy, A. (2009). Cloud removal methodology from MODIS snow cover product. *Hydrology and Earth System Sciences*, 13, 1361-1373.
- Ghahramani, Z. (2001). An introduction to hidden Markov models and Bayesian networks. *International Journal of Pattern Recognition and Artificial Intelligence*, 15, 9-42.
- Goward, S. N., Masek, J. G., Williams, D. L., Irons, J. R. and Thompson, R. (2001). The Landsat 7 mission: terrestrial research and applications for the 21st century. *Remote Sensing of Environment*, 78, 3-12.
- Hall, D. K., Riggs, G. A., Foster, J. L. and Kumar, S. V. (2010). Development and evaluation of a cloud-gap-filled MODIS daily snow-cover product. *Remote sensing of environment*, 114, 496-503.
- Hall, D. K., Riggs, G. A., Salomonson, V. V., Digirolamo, N. E. and Bayr, K. J. (2002). MODIS snow-cover products. *Remote sensing of Environment*, 83, 181-194.
- Immerzeel, W. W., Droogers, P., De Jong, S. and Bierkens, M. (2009). Large-scale monitoring of snow cover and runoff simulation in Himalayan river basins using remote sensing. *Remote sensing of Environment*, 113, 40-49.
- Nsidc. (2016). *MODIS data* [Online]. Available: https://nsidc.org/data/modis/data_versions.html [Accessed].
- Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77, 257-286.
- Rabiner, L. R. and Juang, B.-H. (1986). An introduction to hidden Markov models. *ASSP Magazine, IEEE*, 3, 4-16.
- Shrestha, B., Agrawal, K., Alfthan, B., Bajracharya, S., Marechal, J. and Van Oort, B. (2015). The Himalayan Climate and Water Atlas: Impact of climate change on water resources in five of Asia's major river basins. *ICIMOD, GRID-Arendal and CICERO*.
- Viovy, N. and Saint, G. (1994). Hidden Markov models applied to vegetation dynamics analysis using satellite remote sensing. *Geoscience and Remote Sensing, IEEE Transactions on*, 32, 906-917.