# High Mountain Asia UCLA Daily Snow Reanalysis, Version 1

## USER GUIDE

**How to Cite These Data**

As a condition of using these data, you must include a citation:

Liu, Y., Y. Fang, and S. A. Margulis. 2021. *High Mountain Asia UCLA Daily Snow Reanalysis, Version 1*. [Indicate subset used]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/HNAUGJQXSCVU. [Date Accessed].

FOR QUESTIONS ABOUT THESE DATA, CONTACT NSIDC@NSIDC.ORG

FOR CURRENT INFORMATION, VISIT https://nsidc.org/data/HMA_SR_D

National Snow and Ice Data Center

# TABLE OF CONTENTS

# 1 DATA DESCRIPTION

## 1.1 Parameters

This data set is a snow reanalysis over High Mountain Asia (HMA) derived by assimilating satellite retrieved fractional snow covered area (fSCA) observations from the Landsat and Moderate Resolution Imaging Spectroradiometer (MODIS) platforms over the joint Landsat-MODIS record between Water Year (WY) 2000 and 2017 (Margulis et al., 2019). The data set contains daily estimates of posterior snow estimates (snow water equivalent (SWE), fSCA, snow depth (SD), and snow albedo) and posterior forcings. Ensemble statistics (e.g. mean, median, and spread) are given for the key posterior snow estimates (SWE, fSCA, and SD), while ensemble median values are given for other states and flux variables. Data quality and some cautions of using this dataset are indicated in a classification mask and a non-seasonal snow/ice mask.

## 1.2 File Information

### 1.2.1 Format

Data are provided in NetCDF files.

### 1.2.2 File Contents

Each data set granule consists of four files, one for each main parameter. Table 1 lists all the main parameters, the corresponding file name extensions, and a reference to where to find more file parameter details. The detailed parameter description for each file type can be found in
Table 2 to
Table 5. Additionally, a mask file is provided for each granule. Details on the mask file can be found in Table 6 and section 2.4 Quality, Errors, and Limitations. The dimensions (225 x 225 x 5 x 366) for SWE, fSCA, and SD correspond to latitude by longitude by number or ensemble statistics by day of water year. The ensemble statistics are given in the order of: ensemble mean, ensemble standard deviation, ensemble median, 25th percentile, and 75th percentile. The dimensions (225 x 225 x 366) for snow albedo and posterior forcing variables, correspond to latitude by longitude by day. The dimensions (225 x 225 x 18) for the masks correspond to latitude by longitude by water year (classification mask), or latitude by longitude (non-seasonal snow/ice mask).

The water year refers to October 1st - September 30th, where the year number corresponds to the ending year., i.e. WY 2000 = 01 October 1999 – 30 September 2000.

Table 1. Parameter Details

| File name extensions | Parameters | Reference to Parameter Details |
|---|---|---|
| *SWE_SCA_POST* | Posterior snow water equivalent (SWE) and posterior fractional snow covered area (fSCA) | Table 2 |
| *SD_POST* | Posterior snow depth (SD) | Table 3 |
| *SNOW_ALBEDO* | Posterior snow albedo | Table 4 |
| *FORCING_POST* | Posterior forcing variables | Table 5 |
| *MASK* | Classification mask and non-seasonal snow/ice mask | Table 6 |

Table 2. Parameter Details for *SWE_SCA_POST* files

| Parameter | Description | Unit | Dimension |
|---|---|---|---|
| SCA_Post | posterior fSCA | - | 225 x 225 x 5 x 366 |
| SWE_Post | posterior snow water equivalent | m | 225 x 225 x 5 x 366 |
| Latitude | latitude | degrees N | 225 x 1 |
| Longitude | longitude | degrees E | 225 x 1 |

Table 3. Parameter Details for *SD_POST* files

| Parameter | Description | Unit | Dimension |
|---|---|---|---|
| SD_Post | posterior snow depth | m | 225 x 225 x 5 x 366 |
| Latitude | latitude | degrees N | 225 x 1 |
| Longitude | longitude | degrees E | 225 x 1 |

Table 4. Parameter Details for `*SNOW_ALBEDO*` files

| Parameter | Description | Unit | Dimension |
|---|---|---|---|
| Snow_Albedo_Post | posterior snow albedo | - | 225 x 225 x 366 |
| Latitude | latitude | degrees N | 225 x 1 |
| Longitude | longitude | degrees E | 225 x 1 |

Table 5. Parameter Details for `*FORCING_POST*` files

| Parameter | Description | Unit | Dimension |
|---|---|---|---|
| PPT_Post | posterior surface precipitation | mm | 225 x 225 x 366 |
| Ps_Post | posterior surface pressure | mbar | 225 x 225 x 366 |
| q_Post | posterior reference-level specific humidity | kg/kg | 225 x 225 x 366 |
| Rl_Post | posterior downwelling longwave surface radiation | $W/m^2$ | 225 x 225 x 366 |
| Rs_Post | posterior surface downwelling solar radiation | $W/m^2$ | 225 x 225 x 366 |
| Ta_Post | posterior reference-level air temperature | K | 225 x 225 x 366 |
| Latitude | latitude | degrees N | 225 x 1 |
| Longitude | longitude | degrees E | 225 x 1 |

Table 6. Parameter Details for `*MASK*` files

| Parameter | Description | Unit | Dimension |
|---|---|---|---|
| Classification_mask | indicators of whether prior simulation or posterior update was performed within a certain pixel/year | - | 225 x 225 x 18 |
| Non_seasonal_snow_mask | indicators of whether a pixel was classified as non-seasonal snow/ice | - | 225 x 225 |
| Latitude | latitude | degrees N | 225 x 1 |
| Longitude | longitude | degrees E | 225 x 1 |
| Water_Year | water year corresponds to the classification mask | WY | 18 x 1 |

### 1.2.3 Naming Convention

The data files are named according to the following convention, which is described in Table 7 below:

```
HMA_SR_D_v[nn]_N[latitude]E[longitude]_agg_16_[parameter]_WY[YYYY_YY]_[vv].nc
```

Table 7. File Naming Convention

| File Designator | Description |
|---|---|
| HMA_SR_D | Data set ID. |
| v[nn] | Data set version number. |
| N[latitude] | N for north, followed by a 3-digit latitude, e.g. N34_0 for 34° N indicating the latitude of the lower-left corner of the 1° latitude by 1° longitude file. |
| E[longitude] | E for east, followed by a 3- or 4-digit longitude, e.g. E66_0 for 66° E and E103_0 for 103° E indicating the longitude of the lower-left corner of the 1° latitude by 1° longitude file. |
| agg_16 | This refers to a spatial aggregation factor of 16 from the original resolution of the DEM (30 m) to the model resolution (480 m). |
| [parameter] | Main data parameter. Options are SWE_SCA_POST, SD_POST, SNOW_ALBEDO, FORCING_POST, and MASK. More details on each individual file type can be found in Table 1 to Table 6. |
| WY[YYYY_YY] | WY is short for water year, followed by the starting year and the last two digits of the ending year. E.g. WY1999_00 refers to data for the period of 01 October 1999 to 30 September 2000. |
| .nc | File extension indicating this is a NetCDF file. |

**Example file names:**

- N34_0E66_0_agg_16_SWE_SCA_POST_WY1999_00.nc
- N34_0E66_0_agg_16_SD_POST_WY1999_00.nc
- N34_0E66_0_agg_16_SNOW_ALBEDO_WY1999_00.nc
- N34_0E66_0_agg_16_FORCING_POST_WY1999_00.nc
- N34_0E66_0_agg_16_MASK_WY1999_00.nc

### 1.2.4 Browse File

A .png browse file is provided for each data granule showing the SWE of the day of water year 90 (DOWY 90). Figure 1 shows an example browse file.

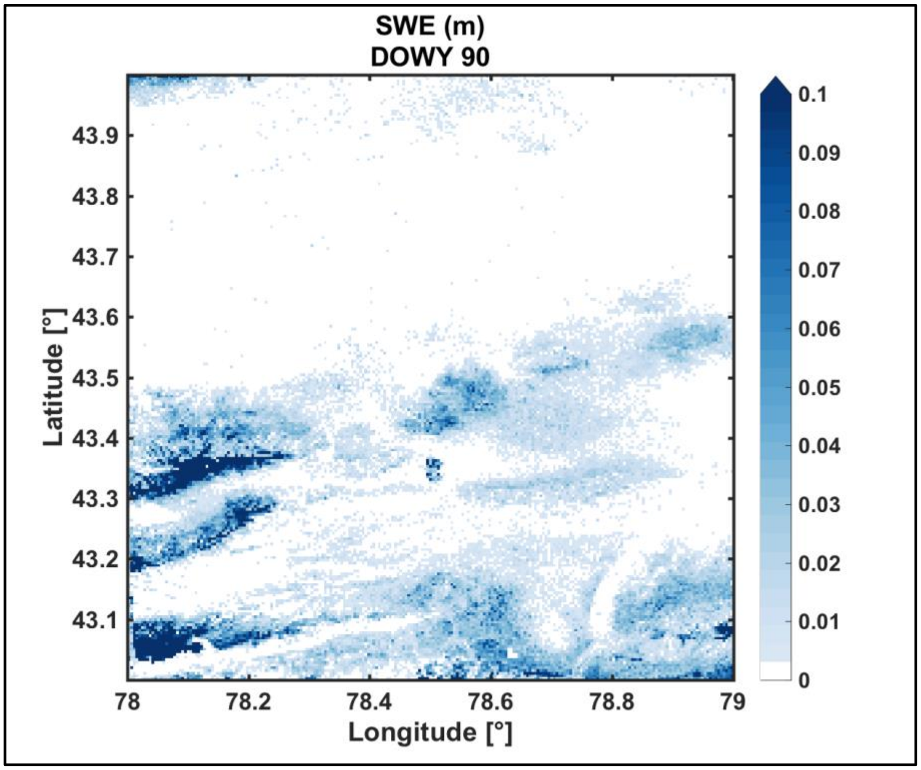


Figure 1. Sample browse file:

HMA_SR_D_v01_N43_0E78_0_agg_16_SWE_SCA_POST_WY2000_01_browse.png

## 1.3 Spatial Information

### 1.3.1 Coverage

Northernmost latitude: 45° N
Southernmost latitude 27° N
Easternmost longitude: 105° E
Westernmost longitude: 60° E

### 1.3.2 Resolution

16 arc-seconds (~500 m)

### 1.3.3 Geolocation

The following table provides information for geolocating this data set:

Table 8. Geolocation Details

| | |
|---|---|
| **Geographic coordinate system** | WGS 84 |
| **EPSG code** | 4326 |
| **PROJ4 string** | +proj=longlat +datum=WGS84 +no_defs |

| Reference | https://epsg.io/4326 |
|---|---|

## 1.4 Temporal Information

### 1.4.1 Coverage

01 October 1999 to 30 September 2017

### 1.4.2 Resolution

Daily

# 2 DATA ACQUISITION AND PROCESSING

## 2.1 Background

Snowpack plays a significant role in the hydrologic cycle over HMA. As a vital water resource, the distribution of snowpack volume also impacts the water availability for downstream populations. To assess the regional water balance, it is important to characterize the spatio-temporal distribution of water storage in the HMA snowpack.

## 2.2 Acquisition

Input data for the snow reanalysis are fSCA from the NASA Landsat satellites and Snow Covered Area and Grain size (MODSCAG; Painter et al., 2009) from the MODIS sensor on NASA's Terra satellite. Specifically, data was downloaded from USGS EarthExplorer for Landsat 5 Thematic Mapper (TM), Landsat 7 Enhanced Thematic Mapper (ETM+), and Landsat 8 Operational Land Imager (OLI). The Landsat fSCA was retrieved using a spectral unmixing algorithm (Painter et al., 2003; Cortés et al., 2014). The MODSCAG product was obtained from the JPL snow server.

Other input data to derive this snow reanalysis data set were topographic data, land cover data, and metrological data:

- The digital elevation model (DEM) obtained from the Shuttle Radar Topography Mission (SRTM) with 1 arc-second resolution served as topographic data. Gaps were filled with the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM, version 2) product with 1 arc-second resolution.
- Land cover data was obtained from the Advanced Very High Resolution Radiometer (AVHRR) global land cover classification (Hansen et al., 2000). Forest cover was obtained from the Tree Canopy Cover (TCC) product containing the Landsat Vegetation Continuous

Fields (Sexton et al., 2013). Glacier data was obtained from Global Land Ice Measurements from Space (GLIMS; Raup et al., 2007).

- Meteorological forcing data was obtained from the Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA-2; Gelaro et al., 2017).

## 2.3 Processing

This data set was derived using a previously developed snow reanalysis scheme (Margulis et al., 2019). A coupled land surface model from SSiB3 (Sun and Xue, 2001; Xue et al., 2003) and a snow depletion curve (Liston et al., 2004) was used to generate the ensemble prior estimates of the snow states and fluxes. A Bayesian update was performed to further constrain the ensemble prior estimates on satellite retrieved fSCA observations using a Particle Batch Smoother (Margulis et al., 2015) approach, where the snow states and fluxes, as well as other forcings and parameters, are updated simultaneously.

The HMA domain used in this data set is bounded by 27° N to 45° N and 60° E to 105° E. The basic processing units for the data set were tiles of size 1° latitude by 1° longitude. The data is made available in files corresponding to these 1° latitude by 1° longitude processing units. Only tiles likely to experience seasonal snow and with an average elevation above 1500 m were processed. The reanalysis data set is provided at a spatial resolution of 16 arc-seconds (~350 m - 500 m), and at a daily temporal resolution.

The individual processing steps are fully described in Margulis et al. (2019) and depicted in the snow reanalysis diagram below.

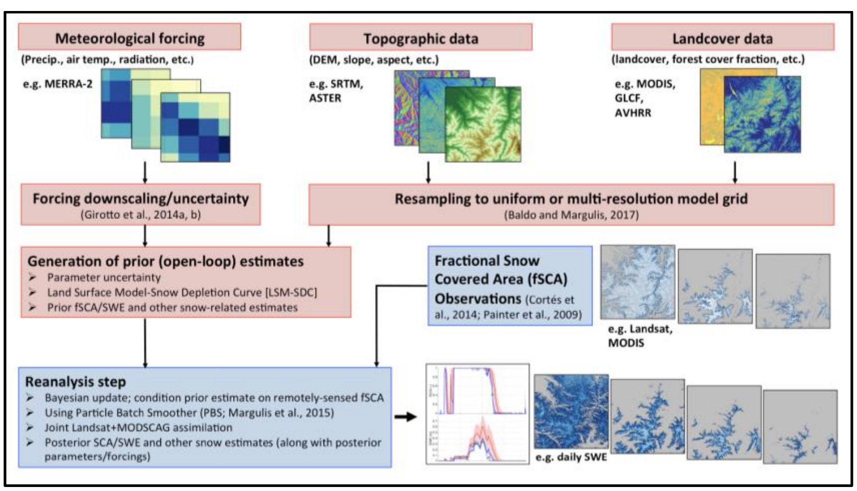


Figure 2. Schematic representation of the Bayesian snow reanalysis framework taken from Margulis et al. (2019).

## 2.4 Quality, Errors, and Limitations

This data set was specifically designed for seasonal snowpack estimation. When using the data consider the following caveats:

i. The posterior precipitation provided in the forcing files is derived from the scaled prior precipitation by fitting it to the fSCA data during snowmelt. Hence, the scaling factor is inferred primarily from the snow depletion. Therefore, while the snowfall portion of the posterior precipitation is likely reliable, the posterior rainfall should be viewed with significant caution. The rainfall portion in the posterior precipitation is unconstrained in the reanalysis and is therefore not recommended for use. Both the snowfall and rainfall portion in posterior precipitation can be estimated using the posterior air temperature provided in the files, i.e. snowfall equals precipitation under 2º C air temperature, and rainfall equals precipitation above 2ºC air temperature. This will introduce a small error as the actual partitioning is done at the hourly scale within the reanalysis while the outputs are in daily resolution.

ii. The classification mask provided in the mask file contains indicators of whether prior simulation or posterior update was performed within a certain pixel/year. Three types of classification were used:
    a. Type 1 for pixels without prior simulation. Pixels are identified as either water bodies or non-snowy pixels (with negligible snowfall in the nominal simulation). Prior simulations were skipped in those pixels and all output values equal to zero (e.g. SWE, fSCA, SD, forcings, etc.).
    b. Type 2 for pixels with prior simulation but no posterior update. These are pixels where the observed fSCA did not effectively change the prior simulation results during the data assimilation step. It is usually due to insufficient number of fSCA measurements or very intermittent snow; and therefore, their posterior outputs equal the prior.
    c. Type 3 for pixels with prior simulation and posterior update. They are identified as snowy pixels and received a substantial posterior update during the assimilation step.

    The majority of pixels in this data set are classified as type 3, while the cases of type 1 or type 2 are relatively rare.

iii. The non-seasonal snow/ice mask provided in the mask files is suggested for use in combination with this data set to mask out pixels with non-seasonal snow features, such as glaciers or persistent snow/ice, etc. This is because the snow reanalysis scheme used to derive this data set is primarily developed for estimating SWE in seasonal snowpack; thus, the results in non-seasonal snow/ice pixels may be unreliable. The non-seasonal snow/ice pixels are denoted with 1 in the mask.

# 3 SOFTWARE AND TOOLS

The .nc data files can be opened using NetCDF visualization software such as Panoply or QGIS.

# 4 VERSION HISTORY

Table 9. Version History Summary

| Version | Release Date | Description of Changes |
|---|---|---|
| V1 | 22 April 2021 | Initial release |
| V1.1 | 20 December 2021 | SNOWMELT and SUBLIMATION files have been removed from this data set due to significant bias issues. Both are expected to be added back in a later version of this data set. |

# 5 RELATED DATA SETS

High Mountain Asia at NSIDC | Data Sets

# 6 RELATED WEBSITES

High Mountain Asia at NSIDC | Overview

NASA High Mountain Asia Project

NASA Research Announcement: Understanding Changes in High Mountain Asia

# 7 CONTACTS AND ACKNOWLEDGMENTS

**Dr. Steven A. Margulis**

University of California, Los Angeles (UCLA)

# 8 REFERENCES

Cortés, G., Girotto, M., & Margulis, S. A. (2014). Analysis of sub-pixel snow and ice extent over the extratropical Andes using spectral unmixing of historical Landsat imagery. Remote Sensing of Environment, 141, 64–78.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). Journal of Climate, 30(14), 5419–5454. http://doi.org/10.1175/JCLI-D-16-0758.1.

Hansen, M. C., Defries, R. S., Townshend J. R. G., & Sohlberg, R. (2000) Global land cover classification at 1 km spatial resolution using a classification tree approach, International Journal of Remote Sensing, 21:6-7, 1331-1364, https://doi.org/10.1080/014311600210209.

Liston, G. E. (2004). Representing Subgrid Snow Cover Heterogeneities in Regional and Global Models. Journal of Climate, 17(6), 1381–1397. http://doi.org/10.1175/1520-0442(2004)017<1381:RSSCHI>2.0.CO;2.

Margulis, S. A., Girotto, M., Cortés, G., & Durand, M. (2015). A Particle Batch Smoother Approach to Snow Water Equivalent Estimation. Journal of Hydrometeorology, 16(4), 1752–1772. http://doi.org/10.1175/JHM-D-14-0177.1.

Margulis, S. A., Liu, Y., & Baldo, E. (2019). A Joint Landsat and MODIS-Based Reanalysis Approach for Midlatitude Montane Seasonal Snow Characterization. Frontiers in Earth Science, 7, 4257. http://doi.org/10.3389/feart.2019.00272.

Painter, T. H., Dozier, J., Roberts, D. A., Davis, R. E., & Green, R. O. (2003). Retrieval of subpixel snow-covered area and grain size from imaging spectrometer data. Remote Sensing of Environment, 85(1), 64–77.

Painter, T. H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R. E., & Dozier, J. (2009). Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. Remote Sensing of Environment, 113(4), 868–879.

Raup, B., Racoviteanu, A., Khalsa, S. J. S., Helm, C., Armstrong, R., & Arnaud, Y. (2007). The GLIMS geospatial glacier database: A new tool for studying glacier change. Global and Planetary Change, 56(1–2), 101–110. https://doi.org/10.1016/j.gloplacha.2006.07.018.

Sexton, J. O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., Kim, D.-H., Collins, K. M., Channan, S., DiMiceli C., & Townshend, J. R. (2013) Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error, International Journal of Digital Earth, 6:5, 427-448, https://doi.org/10.1080/17538947.2013.786146.

Sun, S., & Xue, Y. (2001). Implementing a new snow scheme in Simplified Simple Biosphere Model. Advances in Atmospheric Sciences, 18(3), 335–354. https://doi.org/10.1007/bf02919314.

Xue, Y., Sun, S., Kahan, D. S., & Jiao, Y. (2003). Impact of parameterizations in snow physics and interface processes on the simulation of snow cover and runoff at several cold region sites. Journal of Geophysical Research: Atmospheres, 108(D22), 374. https://doi.org/10.1029/2002jd003174.

# 9 DOCUMENT INFORMATION

## 9.1 Publication Date

22 April 2021

## 9.2 Date Last Updated

26 January 2022