https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

Open Access Earth System Science Data Discussions

A new snow depth data set over northern China by developing a comprehensive framework using the complex GNSS station networks (GSnow-CHINA v1.0, 2013-2020)

Wei Wan[1], Jie Zhang[2], Liyun Dai[3], Hong Liang[4], Baojian Liu[1], Zhizhou Guo[1], Heng Hu[4], Ting Yang[5], Limin Zhao[6]

[1] Institute of Remote Sensing and GIS, School of Earth and Space Sciences, Peking University, Beijing 100871, China

[2] College of Oceanography and Space Informatics, China University of Petroleum (East China), Qingdao 266580, China

[3] Key Laboratory of Remote Sensing of Gansu Province, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China

[4] Meteorological Observation Center, China Meteorological Administration, Beijing 100081, China

[5] Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China

[6] Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China

Correspondences to: Wei Wan (w.wan@pku.edu.cn) and Hong Liang (liangh@cma.gov.cn)

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

Open Access Earth System Science Data Discussions

Abstract: The currently available long-term snow depth data sets are either from point-scale ground measurements or from gridded satellite/modeled/reanalysis data with coarse spatial resolution, which limits the applications in climate model, hydrological model, and regional snow disaster monitoring. Benefit from its unique advantages of cost-effective and high spatial-temporal resolution (~ 1000 m^2, hourly in theory), snow depth retrieval using the Global Navigation Satellite System Interferometric Reflectometry (GNSS-IR) technique has become a popular topic in recent years. However, due to complex environmental and observation conditions, developing robust and operational technology to produce long-term snow depth data sets using observations from various GNSS station networks is still challenging. The two objectives of this study are 1) to propose a comprehensive framework using raw data of the complex GNSS station networks to retrieve snow depth and to control its quality automatically; and 2) to produce a long-term snow depth data set over northern China (i.e., GSnow-CHINA v1.0, 12h/24h, 2013-2020) using the proposed framework and historical data from 80 stations. The data set has high internal consistency with regards to different GNSS systems (mean r = 0.97 & RMSD = 1.93 cm), different frequency bands (mean r = 0.96 & RMSD = 2.73 cm), and different GNSS receivers (mean r = 0.88). The data set also has high external consistency with the in-situ measurements and the passive microwave (PMW) product, with a consistent illustration of the interannual snow depth variability. The results also show the good potential of GNSS to derive hourly snow depth observations for better monitoring snow disasters. The proposed framework to develop the data set provides comprehensive and supportive information for users to process raw data of ground GNSS stations with complex environmental conditions and various observation conditions. The resulting GSnow-CHINA v1.0 data set is distinguished from the current point-scale in-situ data or coarse-gridded data, which can be used as an independent data source for validation purposes. The data set is also useful for regional climate research and other meteorological and hydrological applications. The algorithm and the data files will be maintained and updated as more years of data become available in the future. The GSnow-CHINA v1.0 data set is available at https://doi.org/10.11888/Cryos.tpdc.271839 (Wan et al. 2021).

Keywords: Snow depth, Global Navigation Satellite System Interferometric Reflectometry (GNSS-IR), GNSS station networks, northern China

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

1 Introduction

Snow cover is one of the most active elements in the cryosphere, and the maximum snow area during winter nearly occupies 50% of the total land surface area of the Northern Hemisphere (Frei and Robinson, 1999; Armstrong and Brodzik, 2001; Robinson et al., 1993). The snow change plays a significant role in the hydrological, ecological, and climatic systems (Henderson et al., 2018). Therefore, accurately estimating snow cover and snow depth and their variations is essential for studies on climate and hydrology.

Currently, snow cover products derived from optical remote sensing data present high accuracy, but snow depth products show significant uncertainties. Snow depth can be measured at point-scale using ground-based ultrasonic snow depth sensors or laser snow depth sensors, and mainly include observations from meteorological stations, snow surveys, and hydrological stations (Kinar and Pomeroy, 2015). Large-scale snow depth can be retrieved from optical, passive microwave, and active remote sensing observations (Shi and Dozier, 2000; Guerreiro et al., 2016; Leinss et al., 2014; Che et al., 2016), yet currently operational observations have shortcomings. Optical remote sensing is affected by solar radiation and cloud (Dai et al., 2017). Passive microwave remote sensing is with coarse spatial footprints (> 25 km), and the observations saturate in deep snow (> 0.8 m) (Lievens et al., 2019). Active microwave remote sensing is with a long revisiting period (> 20 days) and with a high cost (Lievens et al., 2019).

The available global/hemispheric/regional snow depth data sets are mainly derived from ground observations, microwave remote sensing, model simulations, and reanalysis (Xiao et al., 2020). Representative snow depth data sets include: 1) In-situ measurements from ground networks such as SCAN and SNOTEL (point-scale, hourly/daily/7-day/monthly; http://www.wcc.nrcs.usda.gov), 2) Data sets derived from satellite passive microwave brightness temperatures, e.g., the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and its follow-on the Advanced Microwave Scanning Radiometer-2 (AMSR2) (25 km, daily, global/regional, 2002-, https://nsidc.org/), and the Global Snow Monitoring for Climate Research (GlobSnow) data set produced from the data assimilation of microwave radiometer data and meteorological station data (25 km, daily, hemispheric, 1979-, https://www.globsnow.info/), 3) Snow depth data set simulated using model such as snow modules in the Global Land Data Assimilation System (GLDAS-2.0, 1948, 0.25° x 0.67°, 3-hourly and monthly; https://ldas.gsfc.nasa.gov/gldas), and 4) Reanalysis snow depth data sets from the

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

ERA-Interim (1979-, 0.75°, 6-hourly/daily/monthly; http://www.ecmwf.int/) and the Modern-Era Retrospective Analysis for Research and Applications as well as their series data sets ((MERRA/MERRA-2/MERRA-Land, 1979, 0.5° x 0.67°; https://gmao.gsfc.nasa.gov/reanalysis/).

The aforementioned long-term snow depth data sets are either point-scale or gridded data with coarse spatial resolution. Previous studies also demonstrated that current snow depth data sets show significant inconsistencies and uncertainties, which limit their applications in climate change projections and hydrological processes simulations (Xiao et al., 2020; Zhang et al., 2021). Due to the complex spatial-temporal variability and the limitations of the current observation approaches, it is still challenging to derive snow depth data set of long-term with high spatial-temporal resolution. In particular, it lacks detailed observations of snow depth on a regional scale, which limits the applications in climate models, hydrological models, and snow disaster monitoring.

Estimating snow depth using the Global Navigation Satellite System Interferometric Reflectometry (GNSS-IR) technique has become a popular topic in recent years, ever since the principle was proposed by (Larson et al., 2009). Snow depth is determined by calculating the relative change of the effective multipath reflector height (i.e., the snow surface) to the snow-free surface. This technique is cost-effective because it does not require an additional transmitter, and instead, it continuously receives L-band microwave signals transmitted by the GNSS satellites. The temporal resolution for snow sensing is expected to be hourly, along with the increasing number of GNSS satellites in orbit (Tabibi et al., 2017). The spatial footprint of the GNSS-derived snow depth is recognized as ~ 1000 m^2, which is a scale between point-scale and satellite-scale (i.e., from tens of meters to tens of kilometers) (Larson and Nievinski, 2013). Therefore, GNSS-IR could provide new snow depth data sets which could be supplementary to the current in-situ and satellite data sets. However, developing robust and operational technology to produce long-term snow depth data sets using data from various GNSS station networks is still challenging due to complex environmental and observation conditions.

This study, taking advantage of 80 sites from the GNSS station networks over northern China, develops a comprehensive framework to process raw data from various stations and subsequently develops a new GNSS-IR snow depth data set (GSnow-CHINA v1.0, 12h/24h, 2013-2020). Northern China has a wide-distributed snow cover from October to April next year. China's annual mean snow extent is greater than 9’000’000 km^2, with a stable snow-covered area of ~ 4’200’000 km^2. This region is the main snow-covered area in China, which also

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

plays a vital role in the climate research of the Northern Hemisphere and the cryosphere. The unique characteristics of the GSnow-CHINA v1.0 and the framework to develop it are as follows:

(1) GSnow-CHINA v1.0 is a snow depth data set developed using GNSS data source, independent from the current satellite, modeled, reanalysis, and in-situ data sets. The spatial resolution of this data set is between the in-situ point-scale and the coarse-gridded data, which makes it a new data set suitable for validation purposes.

(2) GSnow-CHINA v1.0 is a long-term snow depth data set over China with high temporal and spatial resolution, which provides a new data source for regional and global climate research. The data set is also helpful for monitoring local snow disasters and water resource management.

(3) The proposed framework to develop the data set provides comprehensive and supportive information for users to process raw data of ground GNSS stations with complex environmental conditions and various observation conditions. The technique has the potential to provide finer resolution snow depth product (e.g., one to two hours) with adequate observations from multiple GNSS systems.

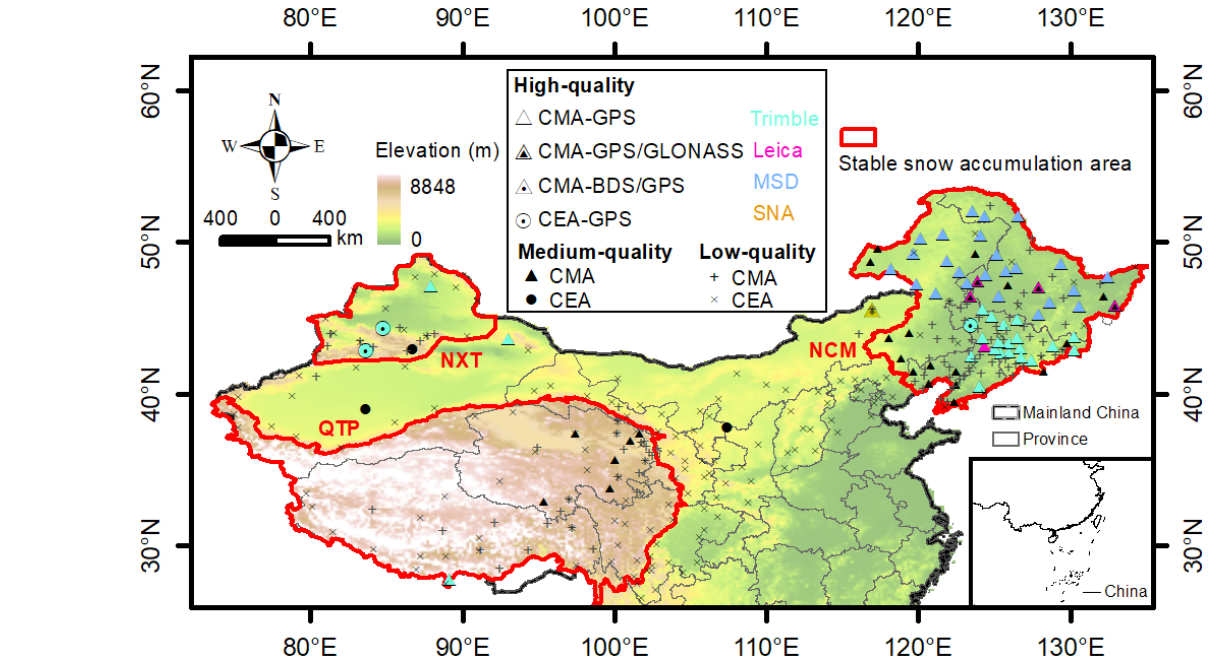
2 Study area and data

2.1 Study area

Northern China lies between latitudes 25°N and 55°N and longitudes 70°E and 140°E and include humid, semi-humid, semi-arid, and arid zones. Snow is the primary freshwater resource over this area and the main natural disaster in winter for the pastoral areas. The study area includes the three main stable snow accumulation areas over China, i.e., Northeast China and Inner Mongolia (NCM), North Xinjiang and Tianshan mountain (NXT), and Qinghai-Tibet Plateau (QTP) (Figure 1).

The NCM region has various geomorphic types. Mountains and hills surround the east, west, and north of this region, and the middle of this region is plain. The mean minimum air temperature in January is below -30 °C. The annual mean snow depth is greater than 5 cm with a maximum value of greater than 30 cm. The mean snow density of this area is ~0.15 g. cm^{-3}. The NXT region has abundant seasonal snow water resources, vital to local irrigation and animal husbandry. The mean air temperature is -4 ~ 9 ℃ with a long winter period. The QTP region is the core region of "The Third Pole" with a mean altitude of ~4378 m. Rainfall of the QTP is concentrated chiefly from May to September, while snowfall usually starts from September to April of the following year.

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021



130 **Figure 1.** Distributions of the GNSS sites over northern China. The symbols are colored by the GNSS receiver type, such as Trimble, Leica, MinShiDa (MSD), and SiNan (SNA).

2.2 Data

Observations from the GNSS station networks over northern China are the primary data source to produce
135 the snow depth data set. The networks include two separate categories constructed by two organizations, i.e., the network constructed by the China Meteorological Administration (CMA) and the Crustal Movement Observation Network of China constructed by the China Earthquake Administration (CEA). China started to construct ground GNSS stations in 2009, and the stations have turned into a certain amount since 2012. The CMA stations were built to observe precipitable water vapor, while the CEA stations were built to monitor crustal deformation.

140 As shown in Figure 1, raw data from all 174 CMA sites and 171 CEA sites are acquired from the Center of Meteorological Observation, CMA, to initially evaluate the capability to retrieve snow depth site by site. The sites are divided into three categories, i.e., high quality, medium quality, and low quality, following the recognition rule used for site quality determination. The rule will be introduced in the following "Methods" Section. Overall, there are 55 high-quality sites (52 for CMA and 3 for CEA) and 25 medium-quality sites (22

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

for CMA and 3 for CEA). The high-quality CMA sites are composed of various types regarding the received data of different GNSS systems, i.e., 47 GPS-only, 4 GPS/GLONASS compatible, and 1 GPS/BDS compatible. The CEA sites are GPS-only sites. Most of the high-quality sites are located in the NCM region, while a few are located in the NXT and QTP regions.

Figure 2 shows the periods of the high-quality and medium-quality GNSS sites used for snow depth retrieval. For CMA, despite the possible raw data missing for some sites, the majority time spans for the high-quality sites are 2013-2020, 2015-2020, and 2016-2020, and that for the medium-quality sites are 2015-2020. For CEA, the high-quality sites are merely from 2018/2019-2020, with one medium-quality site having the earliest record from the year 2010. Due to the irreplaceable value of each site because of its unique natural environment and characteristic of snow, we preserve the high-quality and medium-quality sites as much as possible during the production of this snow depth data set.

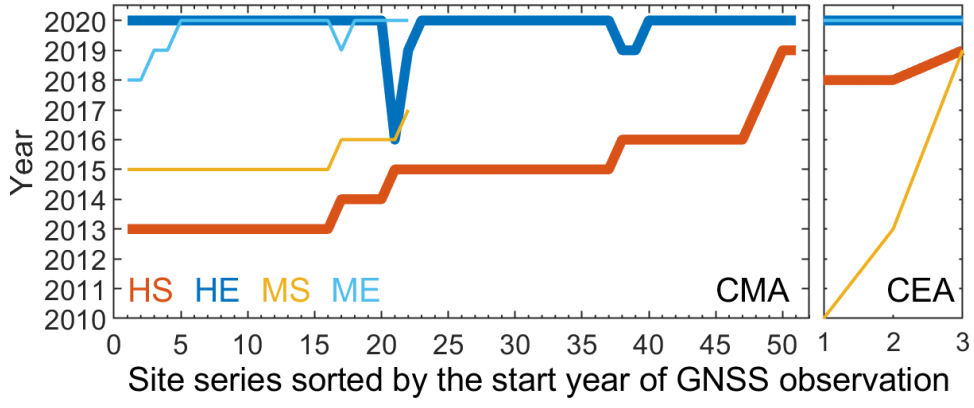


Figure 2. Periods of the GNSS sites used for snow depth retrieval. HS: start year of the high-quality site; HE: end year of the high-quality site; MS: start year of the medium-quality site; ME: end year of the medium-quality site. CMA: China Meteorological Administration; CEA: China Earthquake administration.

The broadcast ephemeris data were downloaded from the International GNSS Service (IGS) website and used to calculate each GNSS satellite's position. For CMA and CEA sites, the minimum elevation angle of the GNSS satellite is set to be 10° in the raw Receiver Independent Exchange Format (RINEX) file.

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021


Open Access Earth System Science Data Discussions

The Soil Moisture Active and Passive (SMAP) L3 36 km soil moisture data are used to estimate the penetration depth of GNSS signals to the soil layer. It is a quality control step to derive a more accurate reflector height of the non-snow surface. The Moderate Resolution Imaging Spectroradiometer (MODIS) 1 km Normalized Difference Vegetation Index (NDVI) data are used to identify the vegetation effects on snow depth retrieval. Two independent snow depth data products are used to analyze the quality of the data set produced in this study. One is the 1979-2020 snow depth product using passive microwave remote sensing produced by (Che and Dai, 2015; Che et al., 2008; Dai et al., 2015) (daily, 25 km), named PMW hereafter for short. The snow depth of this product is derived using the SMMR and SSMI/S microwave brightness temperature processed by the National Snow and Ice Data Center (NSIDC). The other is the daily in-situ snow depth measurements using laser snow depth sensors provided by the Meteorological Observation Center, CMA.

3 Methods

The flowchart to produce and validate the GSnow-CHINA data set is shown in Figure 3. The raw GNSS data used for snow depth retrieval is the daily RINEX data derived directly from individual CMA/CEA GNSS sites. Significant steps to produce the data set are described as follows:

1) The observables for snow depth retrieval, i.e., satellite Pseudorandom Noise (PRN) numbers, observation time, satellite elevation angle, satellite azimuth angle, pseudorange, carrier phase (CP), Signal-to-Noise Ratio (SNR), are extracted or calculated from the raw data.
2) The Lomb-Scargle periodogram analysis is executed on several non-snow days to determine the mean reflector heights for each quadrant, used as reference height when calculating snow depth.
3) A comprehensive evaluation of the quality of all the GNSS sites is done based on the data quality of the non-snow surface reflector height, and the sites are divided into high-, medium-, and low-quality accordingly.
4) For high- and medium-quality sites, the model for deriving daily reflector height is established, and the raw snow depth for each GNSS satellite, each quadrant, and each GNSS frequency is subsequently calculated as the difference value of the referenced height in Step 3) and the height of this step.
5) Several quality-control strategies are used to further improve the quality of the raw snow depth during the previous step, such as considering the penetration depth of soil, considering the vegetation effects,

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021


Open Access Earth System Science Data Discussions

filtering of outliers, adding valid flags such as the standard error (STE) of snow depth, the number of PRNs used to calculate a specific snow depth value, mean Peak-to-Noise Rate (PNR) of the Lomb-Scargle spectrum.

6) Daily 24-hour and sub-daily 12-hour snow depth are derived for general high- and medium- GNSS sites, and snow depths of finer resolution are additionally derived for several GPS/GLONASS compatible sites.

7) The GSnow-CHINA data set is evaluated using the PMW product and the in-situ measurements. The advantages and limitations of the produced data set are further analyzed to provide supportive information for future method improvement or data set extension.

The following sections introduce detailed descriptions of the solutions of several key steps in the processing framework.

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

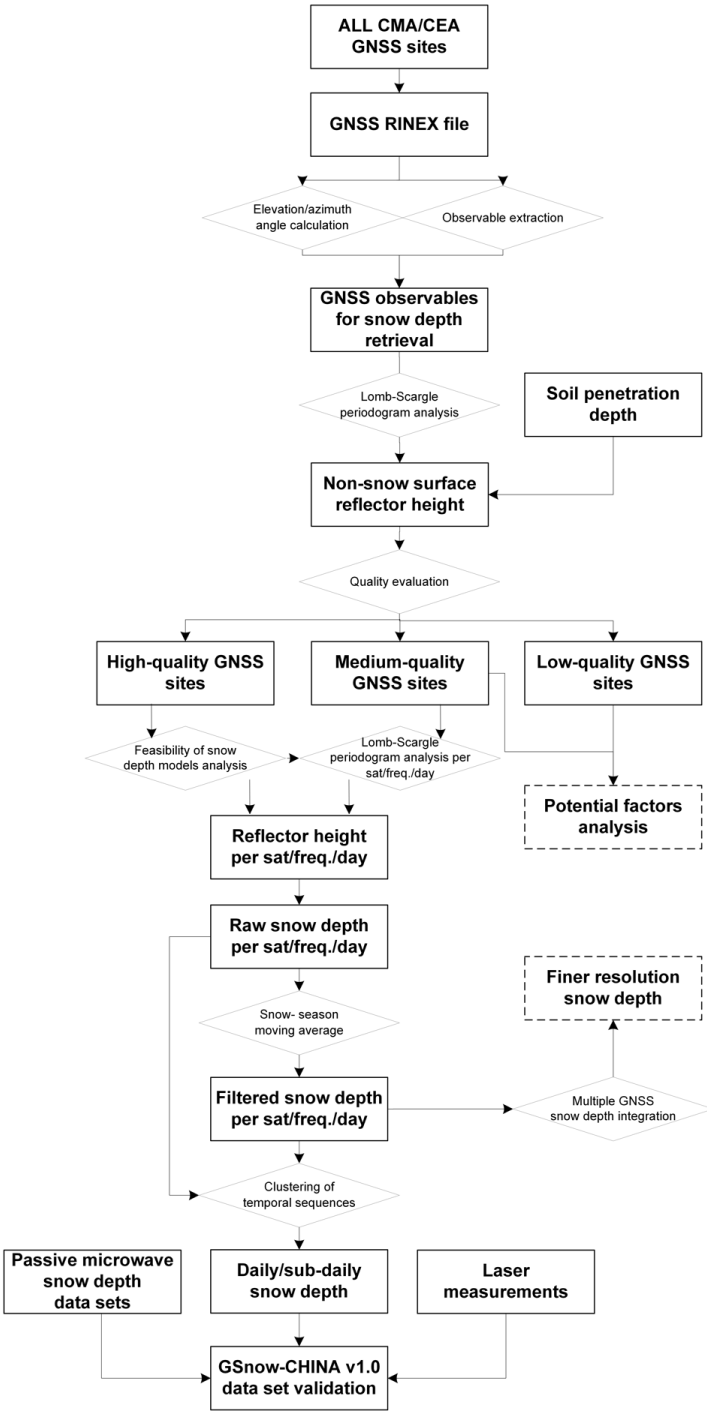


Figure 3. Flowchart showing the production and validation of the GSnow-CHINA v1.0 data set.

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

3.1 Snow depth retrieval model

The state-of-the-art GNSS-IR snow depth retrieving models can be divided into two categories according to the two types of observables (i.e., the SNR and the carrier phase (CP)). The principle of the SNR model is to establish a linear relationship between the oscillation frequency of the SNR observation sequence of the reflected signal and the height of the reflection surface (Larson et al., 2009). This model was later derived into several variants: e.g., the triple-frequency SNR combination model (SNR_COM) (Zhou et al., 2019), the SNR model based on raw SNR sequences (Peng et al., 2016), the SNR model based on horizontal polarization antenna (Chen et al., 2014), the SNR model considering the influence of construction facilities (Vey et al., 2016), and the SNR model considering the influence of terrain (Zhang et al., 2017). The carrier phase combination model was initially proposed to estimate snow depth when there were no SNR data in the raw GNSS observation file (Ozeki and Heki, 2012). The initial form of this model was using the geometry-free linear combinations of the phase measurements (L4), and (Yu et al., 2015; Yu et al., 2018) extended the model to use triple-frequency carrier phase observations (F3) as well as the combination of pseudorange and carrier phase of dual-frequency signals (F2C).

The main formulas and applicability of the five models mentioned above to the data of GNSS sites in this study are analyzed in Table 1, and Table 2 further shows the meanings of variables for the models in Table 1. The SNR, L4, and F2C models are suitable for all sites because the observables used as inputs for these models are available in the GNSS raw data. Even though, the SNR model has been verified to have higher accuracy than the L4 and F2C models (Liu et al., 2021). The applicability of the SNR_COM and F3 models is limited because most of the GNSS sites do not contain three SNR or CP observables in a single raw data file. Considering both the applicability and the accuracy, the SNR model is determined as the primary model used to produce the snow depth data set.

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

Open Access Earth System Science Data Discussions

Table 1 Snow depth models and their corresponding formulas

Model	Main formulas	Applicability
SNR (Larson et al., 2009)	$\mathrm{SNR}^2 = A_c^2 = A_d^2 + A_m^2 + 2A_dA_m \cos Q$ $A_m = A\cos\left(\frac{4\pi h}{\lambda}\sin E + \varphi\right)$ $f = \frac{2h}{\lambda}$	**Suitable for all sites**
SNR_COM (Zhou et al., 2019)	$SNR_{com,i} = [SNR_{1,i}SNR_{2,i}SNR_{3,i}]$	Only suitable for several BDS sites (no triple SNR observations)
L4 (Ozeki and Heki, 2012)	$L_1 = \rho + I(f_1) + T + M_{L1} + \mathrm{noise}_1$ $L_2 = \rho + I(f_2) + T + M_{L2} + \mathrm{noise}_2$ $L_4 = L_1 - L_2 = I(f_1) - I(f_2) + M_{L1} - M_{L2} + \mathrm{noise}_1 - \mathrm{noise}_2$	Suitable for all sites but with relatively lower accuracy
F3 (Yu et al., 2015)	$L_3 = \rho + I(f_3) + T + M_{L3} + \mathrm{noise}_3$ $f_3 = \lambda_3^2(L_1 - L_2) - \lambda_2^2(L_1 - L_3) + \lambda_1^2(L_2 - L_3)$	Suitable for one GPS/GLONASS site
F2C (Yu et al., 2018)	$c_1 = \rho + I(f_1) + T + M_{c1}$ $f_{2c} = \frac{\lambda_1^2 + \lambda_2^2}{\lambda_1^2 - \lambda_2^2}(c_1 - L_1) - \frac{2\lambda_1^2}{\lambda_1^2 - \lambda_2^2}(c_1 - L_2)$	Suitable for all sites but with relatively lower accuracy

Table 2 Meanings of variables for the models in Table 1

Variables	Meanings
A_d	Amplitudes of the direct signal
A_m	Amplitudes of the reflected signal
A_c	Amplitudes of the synthetic signal
$\cos Q$	Cosine value of the angle between the direct signal and the reflected signal
λ	Carrier wavelength
E	Satellite elevation angle
h	Vertical reflection distance
f	Frequency of GNSS multipath reflection signal
φ	Phase values less than an entire period
$SNR_{com,i}$	SNR observation values of triple-frequency
λ_i	Wavelength
ρ	The true geometric range between the satellite and receiver
T	Tropospheric delay
$I(f_i)$	Ionospheric delay for L_i the signal
M_{Li}	Multipath error for L_i signal
$noise_i$	Integer ambiguities for L_i signal
L_4	Multipath error sequence of L4
f_3	Multipath error sequence of F3
f_{2c}	Multipath error sequence of F2C

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

The geometry and principle of the SNR model are shown in Figure 4. As shown in Figure 4a, the snow depth (h_{snow}) is calculated using a simple equation:

$$h_{snow} = h_0 - h \quad (1)$$

Where h_0 is the reflector height of the non-snow surface, and h is the reflector height of the snow-covered surface. The approaches to derive h_0 and h are similar, with Figure 4 b1~b3 showing the general technical process. First, the time series of GNSS SNR observation are shown as a function of sine (elevation angle), and the direct signal is removed using the polynomial fitting method, and the remaining is treated to be the contribution of the reflected signal from the land surface. Second, the reflected signal is converted from dB-Hz to Volts/Volts. Third, a Lomb-Scargle analysis is executed to the reflected signal curve to find out the dominant frequency of the transformation. The h_0 or h can be calculated by (Larson et al., 2009):

$$h = \lambda f / 2 \quad (2)$$

Where λ is the wavelength of the GNSS signal and f is the dominant frequency.

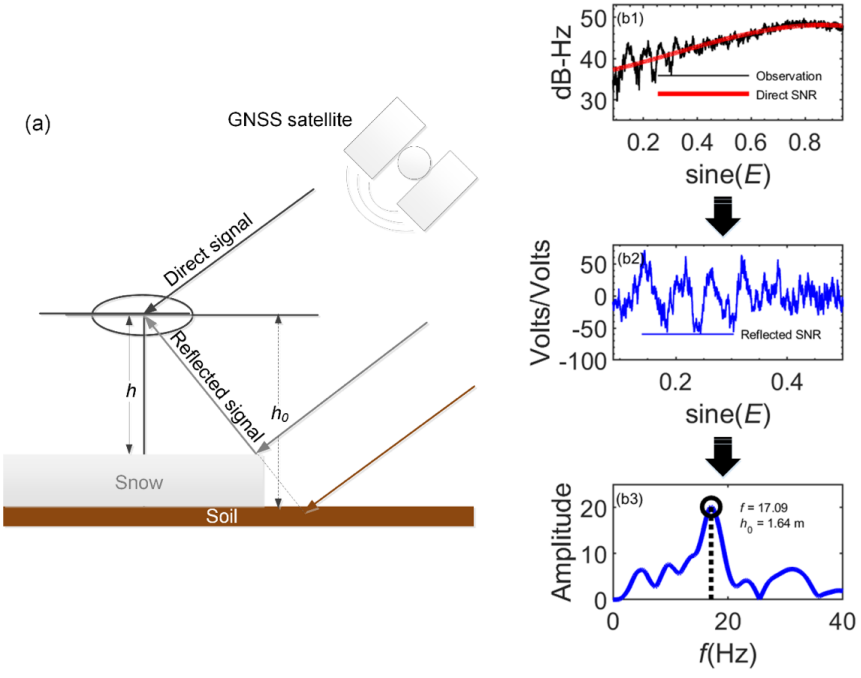


Figure 4. Geometry and principle of the SNR model.

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021


Open Access Earth System Science Data Discussions

245 **3.2 Determination of the non-snow surface reflector height**

For each site, ~ ten days of data with no snow on the ground are used to calculate the non-snow surface reflector height. According to the data availability, days of the year (DOYs) 110~119 or DOYs 274~283 are generally selected since these days have no snow according to historical in-situ data. The reflector height for each GNSS satellite, quadrant, and GNSS frequency band is calculated using the Lomb-Scargle spectrum, and it is 250 just the initial height being used for the quality evaluation of the GNSS sites. Due to the complex natural environment for various sites, it is not clear whether one site is suitable for snow depth retrieval. For sites suitable for snow depth retrieval, the finalized non-snow surface reflector height will be determined as the mean value of heights of the ten days. It is worth mentioning that, for GPS and GLONASS satellite, the reflector height is given per satellite, quadrant, and frequency band, while for BDS satellite, the reflector height is given by quadrant only 255 because the BDS MEO satellite changes its trajectory day by day.

3.3 Quality evaluation of the GNSS sites

The CMA and CEA sites are built under various natural and manual environmental conditions. Figure 5 shows several photos of typical CMA/CEA sites. The CMA sites are mainly built on the ground with antenna height ranging from 1.5 m to 5 m. Some sites are located in relatively flat and open land, while others are in yards 260 and are surrounded by buildings or other artificial objects. The majority of the CEA antennas are settled down on a standard rooftop, with the GNSS receivers being put in the accompanying small house. It explains why most of the CEA sites are not suitable for snow depth retrieval.

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

Open Access Earth System Science Data Discussions

Figure 5. Photos of typical GNSS sites. "bumz" and "bgfc" are two CMA sites, and "qhdl" is a CEA site.

265

A rigorous rule is defined to evaluate the quality of all the GNSS sites. Figure 6 shows the rule being applied to six individual sites with various surroundings, i.e., bumz, bfhr, bgfc, uqwl, qhdl, and qhbm. The top panel of each subfigure shows the environmental conditions around the station on Google Map, with different colors indicating the footprints for elevation angles of 10°, 15°, 20°, 25°, and 30°, respectively. The bottom panel of 270 each subfigure shows the sorted 10-day reflector heights of non-snow surface (i.e., h_0). The plots clearly show the differences in the heights for different sites. The first two sites, i.e., bumz and bfhr, show relatively long and stable h_0 values for all the GNSS satellites, quadrants, and frequency bands during the entire observation period. It indicates that these sites are flat enough for all the orientations and are ideal for determining the initial range of the non-snow surface reflector height, i.e., 2.5 ~ 2.8 m for bumz and 2.8 ~ 3.1 m for bfhr. Unlike these two 275 sites, the bgfc site has relatively stable h_0 values only in specific orientation whose natural condition is open and flat. This phenomenon can be verified from the photo of the bgfc site in Figure 5. This site is also good enough to determine the initial range of the non-snow surface reflector height, i.e., 3.6 ~ 4.1 m for bgfc. On the contrary,

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021


Open Access Earth System Science Data Discussions

the three sites at the bottom of Figure 6, i.e., uqwl, qhdl, and qhbm, show continuously changed h_0 values. It indicates that it is unreliable to determine a true h_0 by the Lomb-Scargle spectrum due to complex environmental conditions.

For each site, the h_0 plot is visually checked carefully using the rule mentioned above and determined whether it is suitable for snow depth retrieval. For those sites that are suitable for usage, a range of h_0 is given manually to narrow the good h_0 values. The difference of the minimum and maximum value of the range is set to be no more than 0.5 m. Afterward, the finalized non-snow surface reflector heights are determined as the mean value of heights of the ten days. It should be noted that during this processing step, it can only eliminate those sites with poor data quality for snow depth retrieval rather than distinguishing high- and medium-sites. There are no apparent differences for the high- and medium-quality sites in terms of the natural environment. Instead, the medium-quality site is defined using two simple rules, i.e., one is the site has good-quality data, but there is no snow for almost all the years. The other is the site's lack of data for most of the years.

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

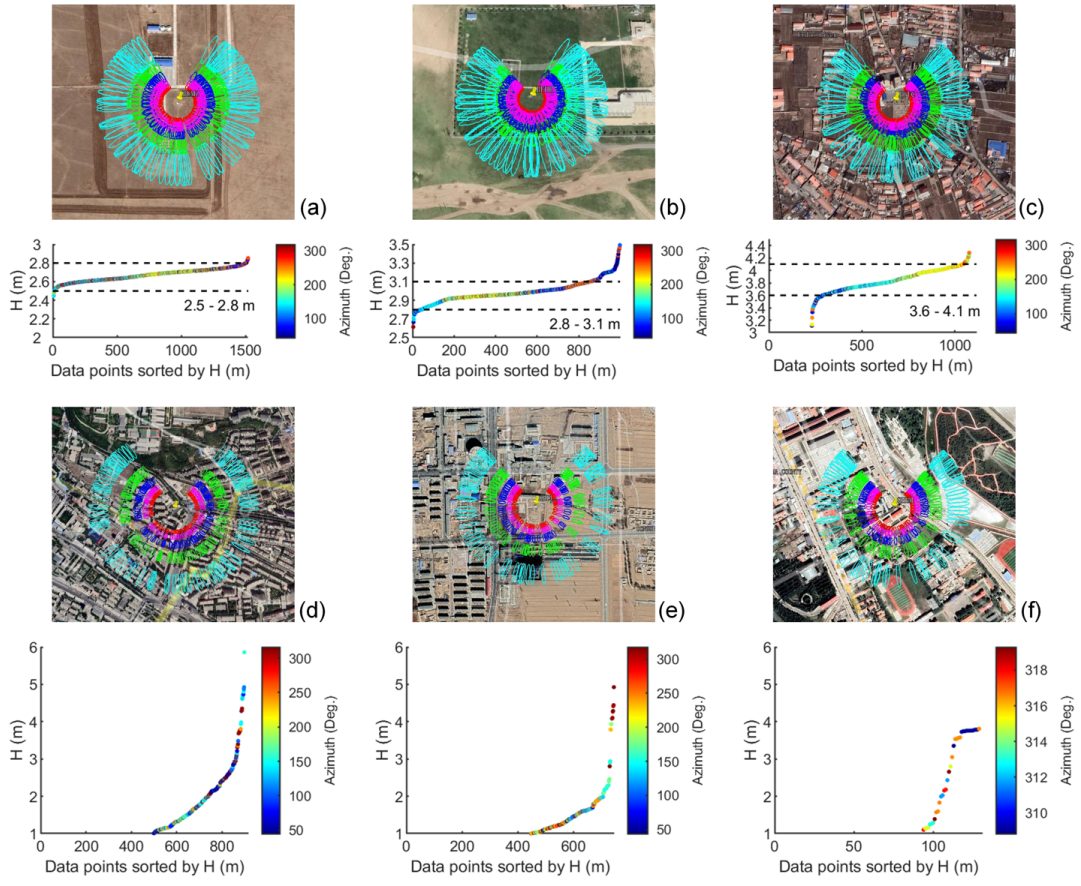


Figure 6. Examples show the high/medium-quality sites and the low-quality sites. High/medium-quality sites: (a) bumz, 2017; (b) bfhr, 2019; (c) bgfc, 2019; Low-quality sites: (d) uqwl, 2019; (e) qhdl, 2020; (f) qhbm, 2018. The top image in each subfigure shows the footprint of the observation for elevation angles of 10°, 15°, 20°, 25°, and 30°, respectively. The bottom image in each subfigure shows the distribution of the reflector heights for non-snow surfaces calculated from 10-days of observations using the SNR model. The background of this figure is from Google Earth (https://earth.google.com/web/) © Google Earth 2021.

3.4 Deriving snow depth of finer resolution

The default temporal resolution of the snow depth data set is 24-hour. However, some sites have adequate satellite observations that make it possible to produce finer resolution snow depth data. We have two different solutions to produce snow depth of finer temporal resolution. For most sites with only GPS observations, we try

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

to produce 12-hour snow depth if there are no less than five valid observations from 0 ~ 12 am or 12 ~ 24 pm within one specific day. The snow depth value for each 12-hour is defined as the mean of all the observations during this time window. For a few sites with GPS/GLONASS compatible observations, we use the exact processing solutions like the previous GPS-only sites and combine all the observations from the GPS and GLONASS systems to derive finer temporal resolution snow depth. Unlike the previous 12-hour maximum resolution, a 2-hour resolution can be achieved using compatible observations.

3.5 Quality control of the snow depth data set

Several postprocessing steps are executed to accomplish the quality control of the raw snow depth data set. This section gives detailed information on these steps as follows:

(1) Moving average filtering

For each site, as shown in Figure 7, the raw snow depth values over a snow season, i.e., from October 1st this year to April 30th the following year, are gathered together, and the moving average algorithm is executed to filter out the snow depth outliers. This moving average method is a traditional way to reject outliers (Wang et al., 2020; Tabibi et al., 2017; Nievinski and Larson, 2014). Snow depth values out of the 95% confidence interval are smoothed over a sliding window of length ten across neighboring elements. In the finalized GSnow-China data set, we also provide the original data set without filtering to allow users to check the initial form of the data. The following analyses in Sections 4 and 5 are based on the filtered data.

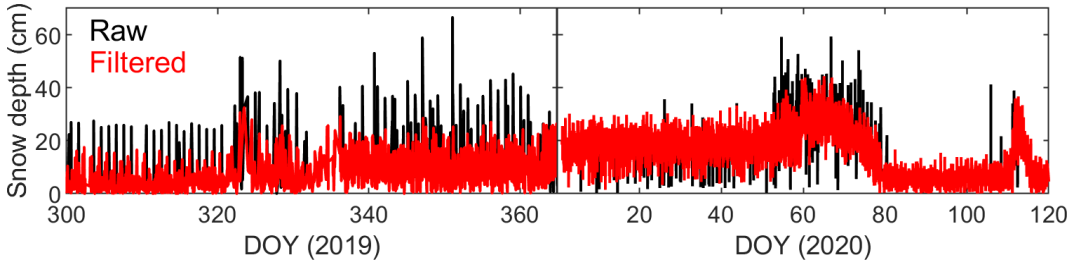


Figure 7. Examples showing the moving-average filtering of the snow depth results over one snow season.

(2) Modifying the system errors caused by the penetration depth of soil

The penetration depth of GNSS signal (h_p) through soil is dependent on the soil moisture and soil components. It directly influences the determination of the reflector height of the non-snow surface. Figure 8 (a)

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

325 (b) shows the relationship between penetration depth and soil moisture/soil components calculated using parameters provided in (Hallikainen et al., 1985). The penetration depth is deeper than 10 cm when soil is very dry (i.e., volumetric soil moisture (VSM) < 0.1 $cm^3.cm^{-3}$). The penetration depth is around or shallower than 5 cm under normal soil moisture conditions. Soil components also affect the penetration depth slightly. In this study, for each site, since the soil components data is unavailable, the VSM value is used as the basis to define 330 the error of the penetration depth for this site. The average VSM is calculated as the multiple-year mean value of the SMAP VSM. To balance the effects from both the soil moisture and the soil components, we define a simple and quantified rule to distinguish the error caused by soil penetration, i.e., h_p = 10 cm for VSM between 0 ~ 0.1 $cm^3.cm^{-3}$, h_p = 5 cm for VSM between 0.1 ~ 0.2 $cm^3.cm^{-3}$, and h_p = 2.5 cm for VSM higher than 0.2 $cm^3.cm^{-3}$. The h_0 is modified as $h_0 - h_p$ for the final production of the snow depth data set. Figure 8 (b) shows statistics 335 of the number of GNSS sites categorized by the modification value of the soil penetration depth. The majority has shallow penetration depth of 2.5 cm or 5 cm, with only a few has deep penetration depth of 10 cm.

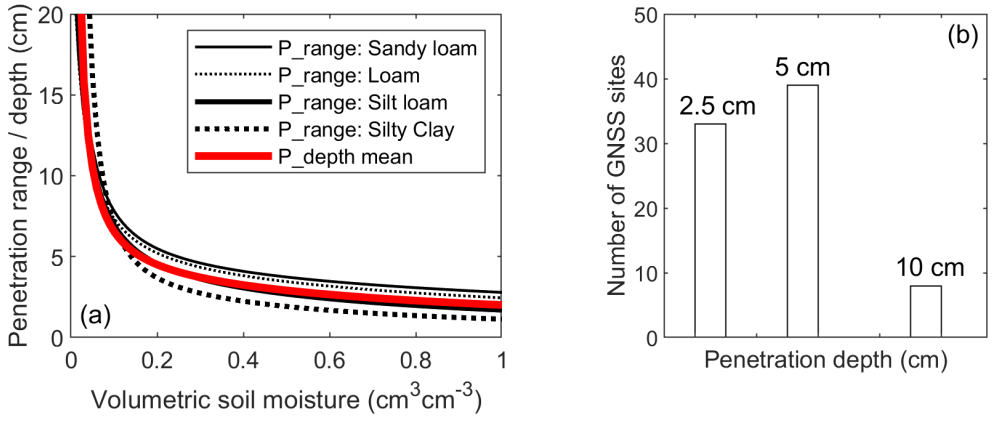


Figure 8. (a) The penetration depth of GNSS signals over the soil layer, taking GPS L1 band (wavelength = 19 cm) as an example. The red line indicates the mean penetration depth for various soil types; (b) Statistics of the number 340 of GNSS sites categorized by the modification value of the soil penetration depth.

(3) Eliminating the vegetation effects

For densely vegetated surfaces, particularly in Autumn, vegetation height is usually calculated as "fake snow depth" due to similar responses on the Lomb-Scargle spectrum. However, it is difficult to identify whether it is

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

345 vegetation or snow. As for northern China, this phenomenon occurs mainly in October and early November. In this study, for each site over from October 1st to November 15th, if there are snow depth records from the GNSS data, we use the NDVI from MODIS data and the historical weather report to determine whether it is actual snow or not. After this round of checking, to ensure the reliability of the snow depth, for 10 sites that probably have "fake snow depth" records, DOYs 270 ~ 300 are masked out from the data set.

350 (4) Quality flags

The number of GNSS satellites used for this calculation is used as a quality flag for each snow depth data record. In this study, we set the threshold to be 5 to preserve as much data as possible. According to this quality flag, the users can decide whether to use a snow depth data record with a low number of observations. For each snow depth data record, the standard error (STE) of the snow depth observations is treated as another qualifying

355 flag. The users can also decide their own rules to filter the data according to this quality flag. The 8-day MODIS NDVI is also involved as a quality flag in the data set to show the vegetation conditions of the site initially. The 8-day values are combinations of the MODIS MOD13Q1 and MYD13Q1 products.

4 Validation of the data quality

4.1 Intra-comparisons of GNSS snow depth results

360 The intra-comparisons of the snow depths are executed from three aspects, i.e., comparison of different GNSS systems, comparison of different frequency bands, and comparison of different GNSS receivers. Figure 9 (a) (b) shows correlations of the snow depths between GPS and GLONASS for 24-hour and 12-hour respectively, using data from the four GPS/GLONASS compatible sites. Both show good agreement with the correlation coefficient r = 0.97 and with the Root Mean Square Difference (RMSD) of the 24-hour result

365 slightly better than the 12-hour result (i.e., 1.88 cm vs. 1.97 cm). The BDS results are not used for comparison due to the limited number of observations.

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

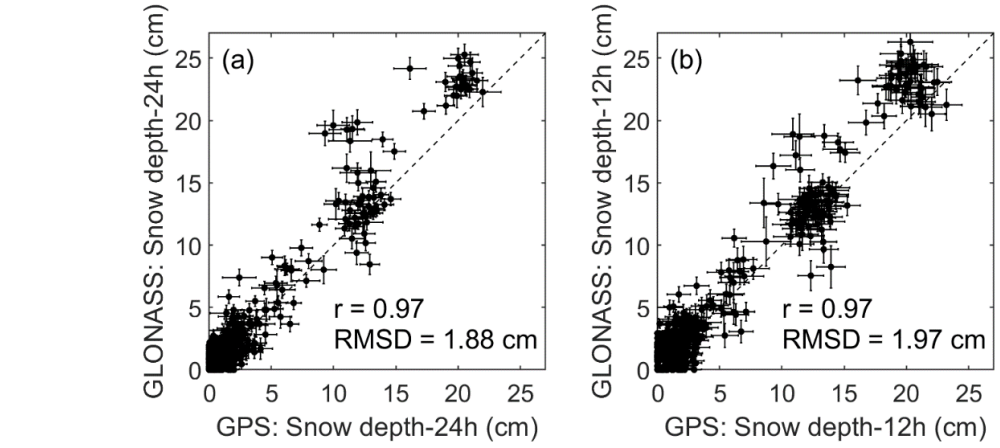


Figure 9. Correlations of 24 h/12 h snow depths from GPS and GLONASS observations. (a) 24 h; (b) 12 h. The error bar of each point is the standard error (STE) of the snow depths for all the observation records.

Figure 10 (a) (b) shows correlations between the snow depths between GPS L1 and L2 and between GLONASS L1 and L2, respectively, using data from the same four GPS/GLONASS compatible sites as in Figure 9. The results from different frequency bands show good consistency with each other, with the correlation coefficient r = 0.94 / RMSD = 2.53 cm for GPS and r = 0.97 / RMSD = 2.93 cm for GLONASS. The BDS results still are not used for comparison due to the limited number of observations.

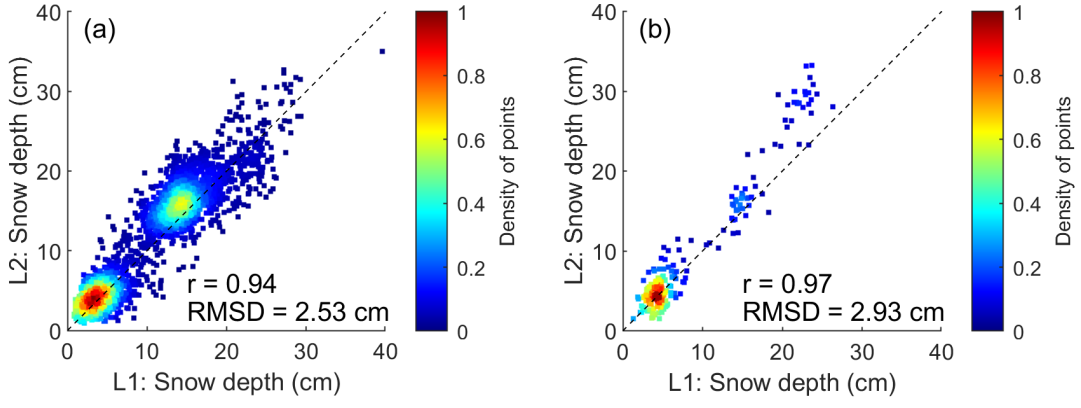


Figure 10. Correlations of snow depth from different GNSS frequencies. (a) GPS L1 vs. GPS L2; (b) GLONASS L1 vs. GLONASS L2. The color bar represents the density of points.

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

Open Access Earth System Science Data Discussions

The CMA and CEA sites are set up with various brands of GNSS receivers. Most of these receivers are from three brands, i.e., Trimble, Leica, and MinShiDa (MSD). Taking these three brands as examples, in order to evaluate the snow depth results from these three brands, Figure 11 (a1), (b1), & (c1) respectively show the differences of the snow depths derived from the three brands, taking the in-situ measurements as benchmarks. The results from the three brands show good consistency with r = 0.86, 0.90, and 0.89, respectively. Figure 11 (a2), (b2), & (c3) further show the histogram of the STEs of the snow depths from the three brands, and good consistency is also shown in these subfigures.

From the comprehensive intra-comparisons shown in Figures 9 ~ 11, we conclude that the snow depths derived from different GNSS systems, different frequency bands, and different GNSS receivers have overall good agreement, which makes it possible to combine all these results to produce the snow depth data set in this study.

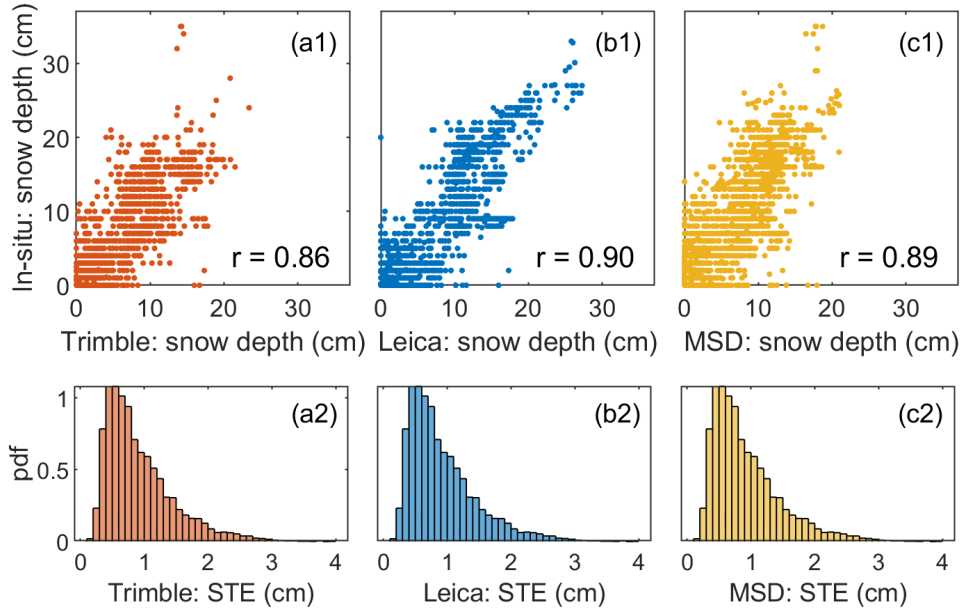


Figure 11. Comparisons of the GNSS-derived snow depth and the in-situ measurements from different types of GNSS receivers: (a1) Trimble; (b1) Leica; (c1) Minshida (MSD), and the histogram of the standard error (STE) of snow depths for different types of GNSS receivers: (a2) Trimble; (b2) Leica; (c2) Minshida (MSD)

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

Open Access Earth System Science Data Discussions

4.2 Comparison with in-situ measurements and the PMW products

The GNSS snow depth data set, the PMW data set, and the in-situ measurements are not consistent in terms of the spatial footprint. The GNSS and in-situ data are with closer footprint compared with the 25-km PMW data. The footprint of GNSS is approximately ~ 30 m x 30 m as illustrated in the following Figure 16. Due to the discrepancy of footprint, it is impractical to give factual accuracies when comparing these three data sets. Instead, we present the performance of the three data sets at daily scale, multi-year scale, and interannual variabilities.

Figure 12 shows an example of the comparisons of daily snow depth derived from GNSS, in-situ, and PMW. The data used in this figure is from 17 GNSS sites with the most extended temporal coverage (i.e., from 2013 to 2020), and the daily mean snow depth of the three data sets is calculated and shown in the figure. As expected, the GNSS and in-situ data have similar performance compared to the PMW data. The GNSS snow depths are generally lower than the in-situ measurements, particularly at those with snow depths higher than 10 cm. In addition, the peak of the PMW snow trend for each snow season moves to the right, which is due to the change of snow grain size (Dai et al., 2012).

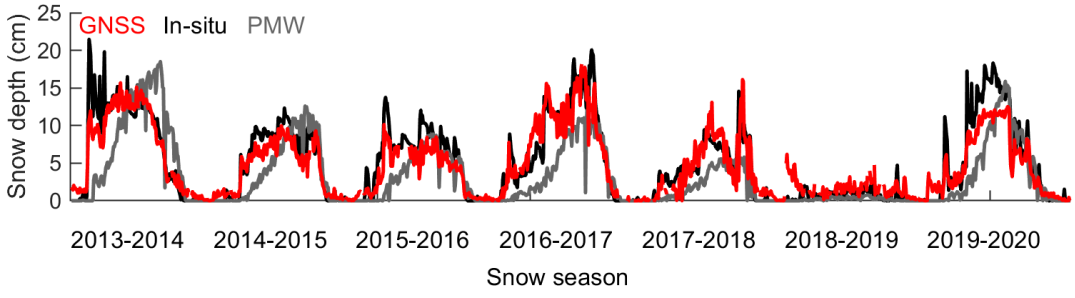


Figure 12. Comparisons of daily snow depth derived from GNSS, in-situ, and PMW.

The annual mean and maximum snow depths are significant indicators that can reflect the overall data quality and the variation trend over multiple years. Sixteen sites with the least missing daily snow depth values are used to compare the multi-year averages of the annual maximum/mean snow depth derived from GNSS, in-situ, and PMW. This comparison period is from 2016 to 2020 due to the data discontinuity for some sites at other periods. Coincidentally, all these sixteen sites are located in the NCM region, making it possible further to analyze the interannual variability of the multi-year maximum or mean snow depth. Figure 13 shows a site-by-site comparison of the five-year average of the annual maximum /mean snow depth derived from GNSS, in-situ, and PMW, respectively. Figure 13 (a1) & (b1) respectively show the spatial distribution of 16 sites marked by their corresponding values of

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

the average of the annual (a1) maximum and (b1) mean snow depth. The snow depth values are classified into five categories to show consistency and discrepancy better. It shows high consistency for the three data sets in general but with discrepancies for some sites. Figure 13 (a2) & (b2) respectively show the site-by-site comparison of the average of the annual (a2) maximum and (b2) mean snow depth.

The maximum values are consistent for the three data sets without regard to the in-situ data having one outlier at Site jldg. This data point is an outlier because the historical weather reports showed no significant snowfall events before or after these dates. This result indicates that the laser measurements in operational meteorological observations are not always reliable. For the mean values shown in (b2), the GNSS and in-situ have a better agreement than the PMW because of the significant difference in their spatial footprint. Figure 13 (a3) & (b3) further show the correlation between the GNSS and in-situ or PMW. Accordingly, higher consistencies are achieved from GNSS vs. in-situ than GNSS vs. PMW, with $r = 0.71$ vs. 0.59 for the maximum and $r = 0.84$ vs. 0.76 for the mean. The outliers are not involved during the correlations.

The interannual variability of the multi-year average of the annual maximum (mean) snow depth using the same data in Figure 13 is further shown in Figure 14. The snow depth values in this figure are the mean values of all 16 sites. The maximum and mean achieve consistent interannual variabilities for all the three data sets, with the absolute maximums of the PMW being relatively higher than the other two. This result generally indicates that the GNSS data set in this study can be used as a new data source to monitor the interannual variability of snow depth.

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

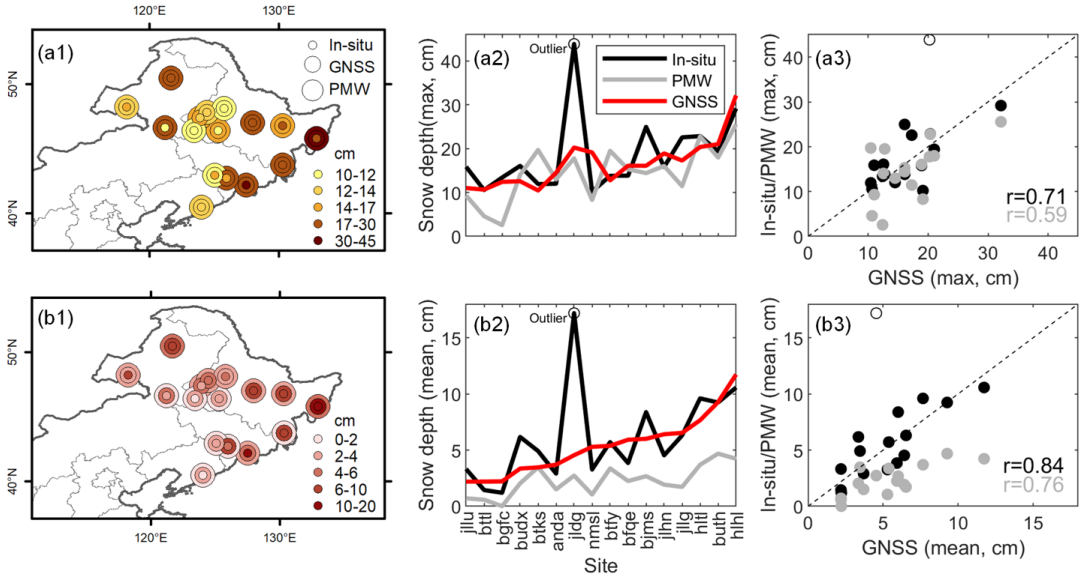


Figure 13. Site-by-site comparison of the five-year average of the annual maximum /mean snow depth derived from GNSS, in-situ, and PMW, respectively. (a1) The spatial distribution of the sites marked by their corresponding values of the five-year average of the annual maximum snow depth; (b1) Same as (a1) but the annual mean; (a2) The site-by-site comparison of the five-year average of the annual maximum snow depth; (b2) Same as (a2) but the annual mean; (a3) The correlation between the GNSS and in-situ/PMW for the five-year average of the annual maximum; (b3) Same as (a3) but the annual mean.

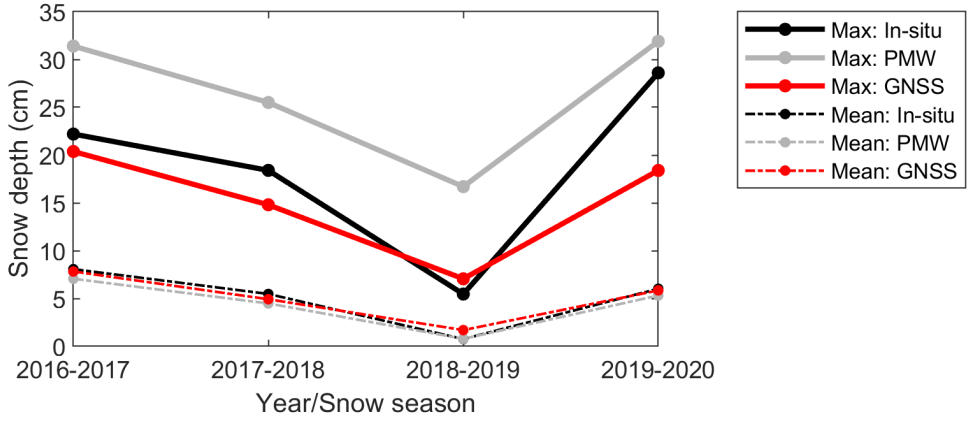


Figure 14. Interannual variability of the multi-year average of the annual maximum (mean) snow depth derived from GNSS, in-situ, and PMW.

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

4.3 Reflection on extreme snow event

Real-time and accurate monitoring of extreme snow events is of vital practical value. To test if this new GNSS data set can provide supportive information for this application, we use the extreme snow event that happened on February 21 ~ 22 in the year 2015 to analyze the performance of the GNSS, in-situ, and PMW data sets. The event is selected because we have overlapped GNSS data from two GPS/GLONASS compatible sites, i.e., bfqe and bttl, which can provide finer resolution snow depth observations. Figure 15 (a) shows the daily snow depth variations before and after the snow event. As expected, the GNSS and in-situ data have similar responses to the event, while the PMW data has a weak response. These two sites are located in the region with evergreen coniferous forest, which prevents the PMW data from acquiring reliable snow depth values due to its wider observation extent of 25 km. Figure 15 (b) further shows the response of the 2-hour GNSS snow depth data during the week of the event. It captures the evolution of the event in a more detailed way from DOY 51 than that of the other two data sets. It provides a potential to increase the monitoring frequency of extreme weather in a cheap and effective way in the future, even with a higher resolution of 1-hour or better, particularly for those sites that have compatible observations from more GNSS satellite systems such as GPS, GLONASS, BDS, and Galileo.

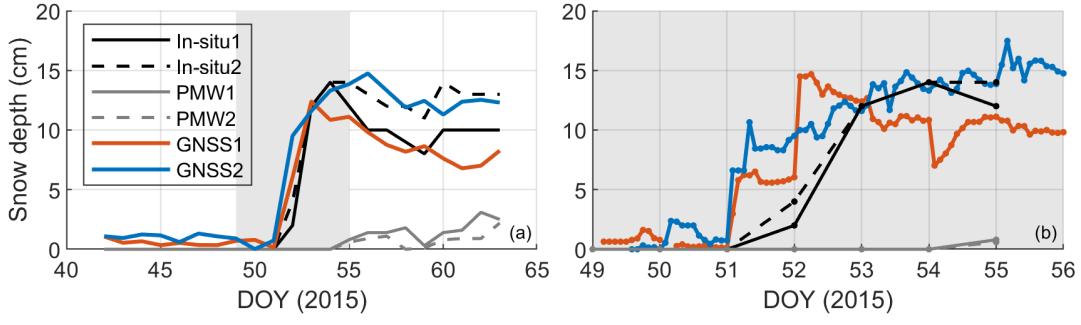


Figure 15. Performance of the GNSS snow depth on a snow event. (a) Daily data; (b) Two-hour data

4.4 Data set descriptions

The GSnow-CHINA Version 1.0 data set is developed using observations from the two GNSS networks constructed by the CMA and CEA. It is called Version 1.0 because we produce the data set using historical observations till the year 2020, and there is room for improvement of the algorithm (e.g., how to properly consider

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021


Open Access Earth System Science Data Discussions

the effects of vegetation). We will continue to maintain and update the algorithm and the data set as more years of data become available in the future. The data set includes snow depth of 24-hour, 12-hour, and 2-hour temporal resolutions if possible, for 80 sites from 2013 ~ 2020 over northern China (25° ~ 55° N, 70° ~ 140° E). The sites over southern China are not included because there is most probably no snow in that region. The high and medium sites are all preserved in the data set with multiple quality flags for users to apply the data.

There are two folders in the data set, i.e., the SITE_INFO and the SNOW_DEPTH. The SITE_INFO folder includes the general information of the 80 GNSS sites, with four separate sheets in one .XLS file corresponding to CMA high-quality, CMA medium-quality, CEA high-quality, and CEA medium-quality, respectively. The items in the file are listed as SITE_NAME, LAT (latitude), LON (longitude), ALT (altitude), RECEIVER_TYPE, GNSS_TYPE, ANTENNA_HEIGHT (in meter), and MEAN_VSM (Volumetric soil moisture in cm^3cm^{-3}; Mean value derived using SMAP soil moisture data of 2015-2020). The SNOW_DEPTH folder includes the snow depth values for all available sites. The folder is structured by ~/site/. For example, ~/hltl/stores the snow depth data of the Site hltl. There are four sub-folders in the folder of each site, i.e., raw0, filtered0, raw, and filtered. The “raw0” and “filtered0” folders store raw data and raw-but-filtered data for individual satellite/quadrant/frequency/time. The "raw" and "filtered" folders store 24-hour/12-hour data produced using raw data in the corresponding “raw0” and “filtered0” folders. The file names including *_24h.csv, *_12h.csv, and *_02h.csv represent the 24-hour, 12-hour, and 2-hour resolution data. Each CSV file gathers this specific snow season (e.g., the 2019 file stores values from October 1, 2019, to April 30, 2020). We recommend using the snow depth data in the “filtered” folder for validation/application purposes while using the snow depth data in the “raw” folder for algorithm testing purposes.

Three quality flags are included in each snow depth file, i.e., the STE, NUM_OF_PRNS, and NDVI, standing for the STE of snow estimations, the number of GNSS sites, and the MODIS NDVI value. These flags should be used to filter the data to balance the data volume and the snow depth accuracy. In addition, we do not recommend using the snow depth values of less than 5 cm in the data set, which is beyond the accuracy of the current GNSS-IR technology.

Figure 16 shows an example of the snow sensing footprint for a specific satellite track. For a 3-meter antenna height under regular 10° ~ 30° elevation angles, the footprint of a specific satellite track is defined as ellipses

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

characterized by the First Fresnel Zone (Larson and Nievinski, 2013), with the maximum length of ~ 30 m for one direction. The GNSS footprint can be recognized as a ~ 30 m x 30 m circle for all orientations. This footprint is between the point-scale of the in-situ measurements and the course 25-km resolution of PMW, which makes it an effective supplement data source for research, validation, and application purposes.

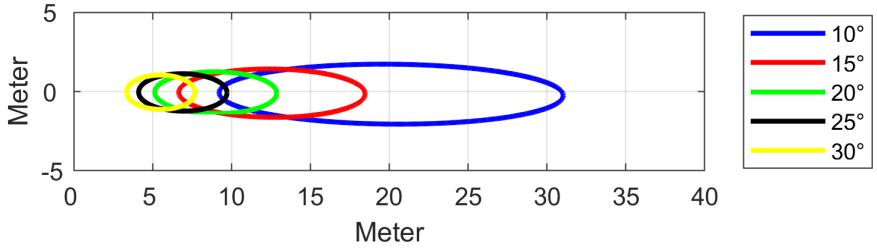


Figure 16. The footprint of the GNSS snow depth observation for a specific satellite track with different satellite elevation angles.

5 Extended analysis of the data set and method

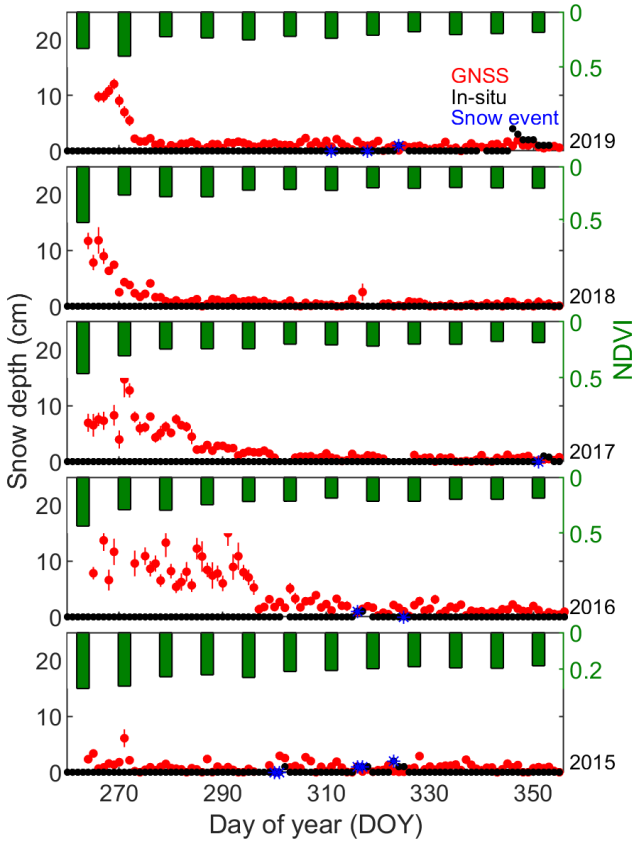
Although this study releases a data set using the current GNSS sites, which are suitable for snow depth retrieval. Those sites that are not suitable for this purpose still deserve an extended analysis to promote this research domain's development further. Also, although the method to retrieve snow depth used in this data set is determined as the SNR model due to data availability, it deserves an extended discussion of the selection of the method for interested readers who dedicates to developing their own data set. The following Sections 5.1 and 5.2 give an extended analysis of the two issues mentioned above.

5.1 Factors that affect the site quality for snow depth retrieval

(1) Natural surroundings. The natural environment within the footprint of the observations is the most significant factor that determines whether a specific GNSS site is suitable for snow depth retrieval or not. Open and flat ground with no vegetation is the ideal environment to set up a snow site, although some previous studies investigated methods to eliminate the influence of terrain (Zhang et al., 2017; Zhang et al., 2020). Vegetation is another factor that needs to be considered for accurate retrieval of snow depth. Figure 17 shows an example of Site bfxc, which has serious vegetation effects on snow depth retrieval before DOY 300 for 2015 ~ 2019. The

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

520 vegetation information is presented by the MODIS 1-km 8-day Normalized Difference Vegetation Index (NDVI) data. The period of the vegetation effects for different years are different, e.g., the year 2016 has the most extended period of ~ 30 days from DOY 270 to 300, while the year 2019 has the shortest period of only ~ ten days around DOY 270. The effect of vegetation is not strictly consistent with the variation of NDVI, which makes it impossible to build a model to qualify the vegetation effect using NDVI data.



525

Figure 17. Examples showing the vegetation effects on snow depth retrieval. The site presented here is bfxc (2015-2020)

Figure 18 shows a correlation between the GNSS snow depth and the in-situ measurement colored by NDVI. 530 Note that for those points on the x-axis with in-situ values equal to 0 but have various GNSS snow depth values, the NDVI values are generally higher than other data points. It illustrates that GNSS measures vegetation rather

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

than snow for these data points. A previous study suggested that it is practical to use the amplitude of the GNSS SNR data to retrieve vegetation height for observations of 1-second sampling. Therefore, for GNSS observations at the sampling intervals, it may be possible to use the SNR amplitude to build a model to qualify the vegetation effect on snow depth retrieval. However, this is not practical for the CMA or CEA sites used in this study because the sampling interval is 30 seconds which is impossible to model the SNR data series to derive the amplitude. Future research will consider using other vegetation indicators to identify this issue.

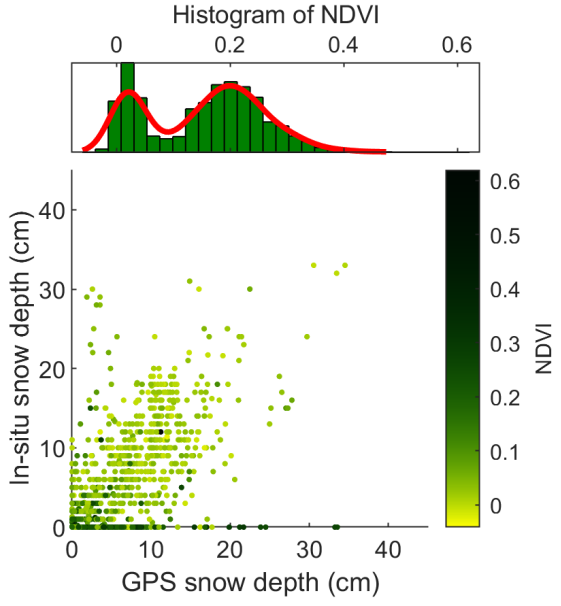


Figure 18. Statistics of the GNSS snow depth when the corresponding in-situ = 0.

(2) Quality of the observation data. The data quality is another critical factor that determines if a site is suitable for snow depth retrieval or not. First, the minimum elevation angle of GNSS satellites should be set to 5° ~ 15° to preserve the multipath effect as much as possible because only data with low elevation angles can show the surface reflection. Second, the observables used as inputs for the corresponding snow depth models should be stored in the raw RINEX file. If the stored observables satisfy conditions for multiple models, one can choose the model according to their accuracy or combine to use all the models during the calculation. This issue will be further discussed in the next Section 5.2. Third, the cycle slip of GNSS observation can severely reduce the data quantity available for snow depth retrieval. One can use the quality control file generated during the

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021


Open Access Earth System Science Data Discussions

observation to filter out these data. Finally, random errors, e.g., human activities at some point, may exist during the observation. Also, the snow depth results on snowy days could, to some extent, affect the accuracy.

5.2 Selection of snow depth models

Although there are many models to retrieve snow depth, as illustrated in Table 1, considering the availability of the observables and the accuracy of the models, not all models are applicable or optimal in practical application. Figure 19 shows an overall strategy of model determination for using GNSS data to retrieve snow depth. One should first consider if the SNR observable exists in the RINEX file since the carrier phase and pseudorange are observations that generally exist for positioning. If the observables satisfy all the snow depth models, the optimal model is selected according to the number of frequencies in the RINEX file. If the frequencies received by the receiver are less than 3, the SNR model is the best choice since it is simple and has reliable accuracy (Plan A in the figure). If the received frequencies are equal to or are greater than 3, the SNR_COM and F3 models can be used (Plan B in the figure). However, one can still use Plan A to replace Plan B in practical applications. If the SNR observable does not exist (Plan C), the F3 model is preferred when the number of CP is greater than 3, while the L4 or the F2C model is selected when the number of CP is less than 3. Nevertheless, the effects of the ionosphere delay on the L4 and F2C models are difficult to remove, which leads to the relatively low accuracy of these two models (Liu et al., 2021).

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

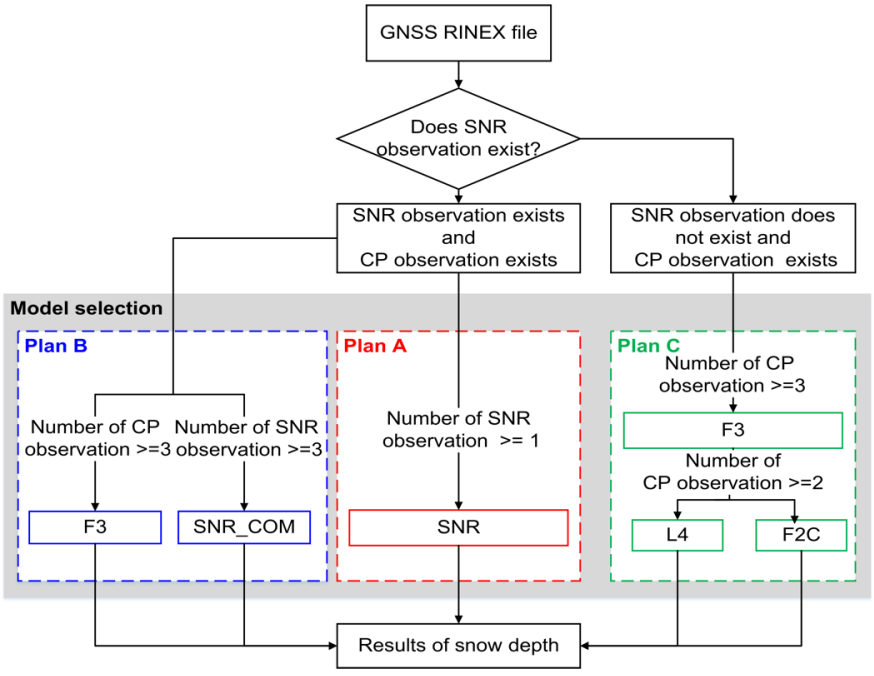


Fig. 19 The strategy of model selection for using GNSS data to retrieve snow depth. CP: carrier phase. Different solutions are represented as Plan A, B, and C.

6 Data availability

The GSnow-CHINA v1.0 data set is archived and available at https://doi.org/10.11888/Cryos.tpdc.271839 (Wan et al., 2021).

7 Conclusions

This study proposes a comprehensive framework using raw data of the complex GNSS station networks to automatically retrieve snow depth and control its quality. Based on this, this study further produces a long-term snow depth data set over northern China (i.e., GSnow-CHINA v1.0, 12h/24h, 2013-2020) using the proposed framework and historical data from 80 stations. The data set has high internal consistency with regards to different GNSS systems (mean r = 0.97 & RMSD = 1.93 cm), different frequency bands (mean r = 0.96 & RMSD = 2.73 cm), and different GNSS receivers (mean r = 0.88). The data set also has high external consistency with the in-situ measurements and the PMW products, with a consistent illustration of the interannual snow depth variability. The results also show the good potential of GNSS to derive hourly snow depth observations for better monitoring

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

snow disasters. The proposed framework to develop the data set provides comprehensive and supportive information for users to process raw data of ground GNSS stations with complex environmental conditions and various observation conditions. The resulting GSnow-CHINA v1.0 data set is distinguished from the current point-scale in-situ data or coarse-gridded data, which can be used as an independent data source for validation purposes. The data set is also useful for regional and global climate research and other meteorological and hydrological applications. Both the algorithm and the data set will be maintained and updated as more years of data become available in the future.

Author contributions: WW designed the study and wrote the manuscript. HL provided the GNSS raw data for the production of this data set and co-designed the study. LD provided supportive information for the validation using the PMW snow depth product. LZ provided supportive information for the data filtering. JZ, BL, ZG, HH, and TY contributed to the data/codes accumulation. data filtering All authors contributed to the writing and editing of this paper.

Competing interests: The authors declare that they have no conflict of interest.

Acknowledgments: The first author would like to thank team members from the Meteorological Observation Center, China Meteorological Administration for producing, maintaining, and providing the raw GNSS RINEX data and the in-situ data. The authors would like to thank the SMAP team, the MODIS team, and the PMW team for archiving and providing the data used in this study.

Financial support: This study is jointly supported by the National Key Research and Development Program of China (Grant No. 2019YFE0126600), the National Natural Science Foundation of China (NSFC) projects (Grant No. 41971377 and No. 41501360), The open fund of the National Earth Observation Data Center (No. NODAOP2021002), the ESA-MOST China Dragon5 Programme (ID.58070), and the observing experiment project of Meteorological Observation Center of China Meteorological Administration (No. SY2020005).

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

Open Access Earth System Science Data Discussions

Review statement: This paper was edited by xxx and reviewed by xxx anonymous referees.

610

References

Armstrong, R. L. and Brodzik, M. J.: Recent northern hemisphere snow extent: A comparison of data derived from visible and microwave satellite sensors, Geophysical Research Letters, 28, 3673-3676, https://doi.org/10.1029/2000GL012556, 2001.

615 Che, T. and Dai, L.: Long-term series of daily snow depth dataset in China (1979-2020), National Tibetan Plateau Data Center [dataset], 10.11888/Geogra.tpdc.270194, 2015.

Che, T., Dai, L., Zheng, X., Li, X., and Zhao, K.: Estimation of snow depth from passive microwave brightness temperature data in forest regions of northeast China, Remote Sensing of Environment, 183, 334-349, https://doi.org/10.1016/j.rse.2016.06.005, 2016.

620 Che, T., Li, X., Jin, R., Armstrong, R., and Zhang, T. J.: Snow depth derived from passive microwave remote-sensing data in China, Annals of Glaciology, 49, 145-154, 2008.

Chen, Q., Won, D., and Akos, D. M.: Snow depth sensing using the GPS L2C signal with a dipole antenna, EURASIP J. Adv. Signal Process., 2014, 1-10, 10.1186/1687-6180-2014-106, 2014.

Dai, L., Che, T., and Ding, Y.: Inter-calibrating SMMR, SSM/I and SSMI/S data to improve the consistency of snow-
625 depth products in China, Remote Sensing, 7, 7212-7230, 2015.

Dai, L., Che, T., Ding, Y., and Hao, X.: Evaluation of snow cover and snow depth on the Qinghai–Tibetan Plateau derived from passive microwave remote sensing, The Cryosphere, 11, 1933-1948, 10.5194/tc-11-1933-2017, 2017.

Dai, L., Che, T., Wang, J., and Zhang, P.: Snow depth and snow water equivalent estimation from AMSR-E data based
630 on a priori snow characteristics in Xinjiang, China, Remote Sensing of Environment, 127, 14-29, https://doi.org/10.1016/j.rse.2011.08.029, 2012.

Frei, A. and Robinson, D. A.: Northern Hemisphere snow extent: regional variability 1972–1994, International Journal of Climatology, 19, 1535-1560, https://doi.org/10.1002/(SICI)1097-0088, 1999.

Guerreiro, K., Fleury, S., Zakharova, E., Rémy, F., and Kouraev, A.: Potential for estimation of snow depth on Arctic
635 sea ice from CryoSat-2 and SARAL/AltiKa missions, Remote Sensing of Environment, 186, 339-349, https://doi.org/10.1016/j.rse.2016.07.013, 2016.

Hallikainen, M. T., Ulaby, F. T., Dobson, M. C., El-Rayes, M. A., and Lil-Kun, W.: Microwave Dielectric Behavior of Wet Soil-Part 1: Empirical Models and Experimental Observations, IEEE Transactions on Geoscience and Remote Sensing, GE-23, 25-34, 1985.

640 Henderson, G. R., Peings, Y., Furtado, J. C., and Kushner, P. J.: Snow–atmosphere coupling in the Northern Hemisphere, Nature Climate Change, 8, 954-963, 10.1038/s41558-018-0295-6, 2018.

Kinar, N. J. and Pomeroy, J. W.: Measurement of the physical properties of the snowpack, Reviews of Geophysics, 53, 481-544, https://doi.org/10.1002/2015RG000481, 2015.

Larson, K. M. and Nievinski, F. G.: GPS snow sensing: results from the EarthScope Plate Boundary Observatory, GPS
645 Solutions, 17, 41-52, 10.1007/s10291-012-0259-7, 2013.

Larson, K. M., Gutmann, E. D., Zavorotny, V., Braun, A., Williams, M. W., and Nievinski, F. G.: Can we measure snow depth with GPS receivers? Geophysical Research Letters, 36, L17502, 2009.

Leinss, S., Parrella, G., and Hajnsek, I.: Snow Height Determination by Polarimetric Phase Differences in X-Band SAR Data, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7, 3794-3810,
650 10.1109/JSTARS.2014.2323199, 2014.

Lievens, H., Demuzere, M., Marshall, H.-P., Reichle, R. H., Brucker, L., Brangers, I., de Rosnay, P., Dumont, M., Girotto, M., Immerzeel, W. W., Jonas, T., Kim, E. J., Koch, I., Marty, C., Saloranta, T., Schöber, J., and De Lannoy, G. J. M.: Snow depth variability in the Northern Hemisphere mountains observed from space, Nature Communications, 10, 4629, 10.1038/s41467-019-12566-y, 2019.

655 Liu, S., Zhang, J., Wan, W., Liang, H., Liu, B., and Guo, Z.: A comprehensive evaluation of utilizing BeiDou data to estimate snow depth from two ground-based stations, GPS Solutions, under review, 2021.

Nievinski, F. G. and Larson, K. M.: Inverse Modeling of GPS Multipath for Snow Depth Estimation—Part II: Application and Validation, IEEE Transactions on Geoscience and Remote Sensing, 52, 6564-6573, 10.1109/TGRS.2013.2297688, 2014.

660 Ozeki, M. and Heki, K.: GPS snow depth meter with geometry-free linear combinations of carrier phases, Journal of Geodesy, 86, 209-219, 10.1007/s00190-011-0511-x, 2012.

https://doi.org/10.5194/essd-2021-432
Preprint. Discussion started: 7 December 2021

Peng, X., Wan, W., and Chen, X.: Using GPS SNR data to estimate Soil Moisture variations: Proposing a new interference model, 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 10-15 July 2016, 4819-4822, 10.1109/IGARSS.2016.7730257,

Robinson, D. A., Dewey, K. F., and Heim, R. R.: Global Snow Cover Monitoring: An Update, Bulletin of the American Meteorological Society, 74, 1689-1696, 10.1175/1520-0477, 1993.

Shi, J. and Dozier, J.: Estimation of snow water equivalence using SIR-C/X-SAR. II. Inferring snow depth and particle size, IEEE Transactions on Geoscience and Remote Sensing, 38, 2475-2488, 10.1109/36.885196, 2000.

Tabibi, S., Geremia-Nievinski, F., and Dam, T. v.: Statistical Comparison and Combination of GPS, GLONASS, and Multi-GNSS Multipath Reflectometry Applied to Snow Depth Retrieval, IEEE Transactions on Geoscience and Remote Sensing, 55, 3773-3785, 10.1109/TGRS.2017.2679899, 2017.

Vey, S., Güntner, A., Wickert, J., Blume, T., Thoss, H., and Ramatschi, M.: Monitoring Snow Depth by GNSS Reflectometry in Built-up Areas: A Case Study for Wettzell, Germany, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9, 4809-4816, 10.1109/JSTARS.2016.2516041, 2016.

Wan, W., Zhang, J., Dai, L., Liang, H., Liu, B., Guo, Z., Hu, H., Yang, T., and Zhao, L.: A GNSS-based snow depth data set over northern China (GSnow-CHINA v1.0, 12h/24h, 2013-2020), National Tibetan Plateau Data Center [data set], https://doi.org/10.11888/Cryos.tpdc.271839, 2021.

Wang, X., Zhang, S., Wang, L., He, X., and Zhang, Q.: Analysis and combination of multi-GNSS snow depth retrievals in multipath reflectometry, GPS Solutions, 24, 1-13, 2020.

Xiao, L., Che, T., and Dai, L.: Evaluation of Remote Sensing and Reanalysis Snow Depth Datasets over the Northern Hemisphere during 1980–2016, Remote Sensing, 12, 10.3390/rs12193253, 2020.

Yu, K., Li, Y., and Chang, X.: Snow Depth Estimation Based on Combination of Pseudorange and Carrier Phase of GNSS Dual-Frequency Signals, IEEE Transactions on Geoscience and Remote Sensing, 1-12, 10.1109/TGRS.2018.2869284, 2018.

Yu, K., Ban, W., Zhang, X., and Yu, X.: Snow Depth Estimation Based on Multipath Phase Combination of GPS Triple-Frequency Signals, IEEE Transactions on Geoscience and Remote Sensing, 53, 5100-5109, 10.1109/TGRS.2015.2417214, 2015.

Zhang, H., Zhang, F., Che, T., Yan, W., and Ye, M.: Investigating the ability of multiple reanalysis datasets to simulate snow depth variability over mainland China from 1981 to 2018, Journal of Climate, 1-48, 10.1175/JCLI-D-20-0804.1, 2021.

Zhang, S., Wang, X., and Zhang, Q.: Avoiding errors attributable to topography in GPS-IR snow depth retrievals, Advances in Space Research, 59, 1663-1669, https://doi.org/10.1016/j.asr.2016.12.031, 2017.

Zhang, Z., Guo, F., and Zhang, X.: Triple-frequency multi-GNSS reflectometry snow depth retrieval by using clustering and normalization algorithm to compensate terrain variation, GPS Solutions, 24, 52, 10.1007/s10291-020-0966-4, 2020.

Zhou, W., Liu, L., Huang, L., Yao, Y., Chen, J., and Li, S.: A New GPS SNR-based Combination Approach for Land Surface Snow Depth Monitoring, Scientific Reports, 9, 3814, 10.1038/s41598-019-40456-2, 2019.