MetOp/AVHRR Snow Detection Method for Meteorological Applications

NIILO SILJAMO,^a OTTO HYVÄRINEN,^a AKU RIIHELÄ,^a AND MARKKU SUOMALAINEN^b

^a *Finnish Meteorological Institute, Helsinki, Finland*

^b *Center of Ubiquitous Computing, Faculty of Information Technology and Electrical Engineering, University of Oulu, Oulu, Finland*

(Manuscript received 17 February 2020, in final form 28 September 2020)

ABSTRACT: Snow cover plays a significant role in the weather and climate system by affecting the energy and mass transfer between the surface and the atmosphere. It also has far-reaching effects on ecosystems of snow-covered areas. Therefore, global snow-cover observations in a timely manner are needed. Satellite-based instruments can be utilized to produce snow-cover information that is suitable for these needs. Highly variable surface and snow-cover features suggest that operational snow extent algorithms may benefit from at least a partly empirical approach that is based on carefully analyzed training data. Here, a new two-phase snow-cover algorithm utilizing data from the Advanced Very High Resolution Radiometer (AVHRR) on board the MetOp satellites of the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) is introduced and evaluated. This algorithm is used to produce the MetOp/AVHRR H32 snow extent product for the Satellite Application Facility on Support to Operational Hydrology and Water Management (H SAF). The algorithm aims at direct detection of snow-covered and snow-free pixels without preceding cloud masking. Pixels that cannot be classified reliably to snow or snow-free, because of clouds or other reasons, are set as unclassified. This reduces the coverage but increases the accuracy of the algorithm. More than four years of snow-depth and state-of-the-ground observations from weather stations were used to validate the product. Validation results show that the algorithm produces high-quality snow coverage data that may be suitable for numerical weather prediction, hydrological modeling, and other applications.

KEYWORDS: Snow cover; Remote sensing; Satellite observations

1. Introduction

Snow has far-reaching effects on climate and ecosystems. In weather forecasting and climate research, the effects of snow have been widely considered, but a recent study by Niittynen et al. (2018) reiterates that snow may have a strong impact on ecosystems as well. Snow is also an important factor in hydrology, as discussed by Thirel et al. (2012), who emphasized the role of forests in snow products. Pullen et al. (2010) discuss the importance of snow data in weather models and forecasting.

Current remote sensing satellites used for snow detection are either in polar or geostationary orbits, which have their advantages and disadvantages. Most of the seasonal snow is in high latitudes, which are poorly covered by geostationary satellites. Whereas instruments aboard geostationary satellites provide excellent temporal resolution, polar satellite instruments have a better spatial resolution and a better polar coverage, making them often a better option in snow detection. However, due to their orbital characteristics, only a few

observations per day may be available, making them more susceptible to, for example, cloudiness preventing surface observation. Other challenges, such as topography, surface properties, weather, and snow-cover evolution are present in the satellite snow product development for both orbit types.

This paper introduces a new global two-phase snow-cover algorithm and product for Advanced Very High Resolution Radiometer (AVHRR) on board the first generation MetOp satellites. This product is a part of the product portfolio of the Satellite Application Facility (SAF) on Support to Operational Hydrology and Water Management (H SAF). Earlier product and algorithm for SEVIRI on board the Meteosat Second Generation (MSG) satellites provides snow data on limited geographical coverage on MSG/SEVIRI disk, mainly Europe and Africa (Siljamo and Hyvärinen 2011). This new product for MetOp/AVHRR provides truly global coverage and a better spatial resolution in polar regions.

The AVHRR on board polar-orbiting satellites is a well-known imager instrument with a long history in remote sensing. The visual and IR ranges of the electromagnetic spectrum are covered by six channels (five in simultaneous use). Snow extent is usually provided either in binary (snow/no snow) or fractional (percentage) format. The product presented in this paper resembles binary products, with an additional class for a partial snow cover.

There are, of course, many previous snow extent or coverage algorithms and products based on those algorithms. Many AVHRR processing packages include a snow product (e.g., Dybbroe et al. 2005). Hüsler et al. (2012) suggest another method for snow detection over the European Alps using

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JAMC-D-20-0032.s1>.

Corresponding author: Niilo Siljamo, niilo.siljamo@fmi.fi

DOI: 10.1175/JAMC-D-20-0032.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

AVHRR. There are well-known global snow analysis products for the more recent Moderate Resolution Imaging Spectroradiometer (MODIS) instrument (e.g., [Miller et al. 2005](#)). [Notarnicola et al. \(2013a,b\)](#) present another MODIS snow detection algorithm and validation results with a resolution of 250 m. In the paper by [Hori et al. \(2017\)](#), an algorithm using both AVHRR and MODIS and product covering Northern Hemisphere for years 1978–2015 is described, along with validation results. The CryoClim project provides cryospheric products using various satellite instruments, including MODIS-based snow-cover extent ([Solberg et al. 2009](#)).

Even though the AVHRR and MODIS instruments are perhaps the most well-known and most used polar-orbiting instruments for meteorological and hydrological applications, other instruments can be utilized. For example, [Selkowitz and Forster \(2015\)](#) present a method for automatic mapping of persistent ice and snow for Landsat TM and ETM+.

The *Aqua* and *Terra* satellites, which have the MODIS instrument on board, are nearing the end of their lifetime. Their successor, the Visible Infrared Imaging Radiometer Suite (VIIRS) on board the National Polar-Orbiting Partnership (NPP) and the Joint Polar Satellite System (JPSS) satellites, is an instrument with similar channels and snow detection capabilities as MODIS ([Miller et al. 2006](#)). [Key et al. \(2013\)](#) describe snow and ice products for the *Suomi NPP*/VIIRS. [Riggs et al. \(2017\)](#) describe both MODIS and VIIRS snow-cover products.

Likewise, in the early 2020s, the MetOp satellites with the AVHRR instrument on-board will be superseded by the next generation of EUMETSAT polar satellites (MetOp SG) with the new Meteorological Imager (METImage). A similar snow product as the one presented here will be developed for the METImage instrument.

In the geostationary orbit, there are several satellites [such as GOES, Meteorological Satellite (Meteosat), *FY-2*, and Himawari] and instruments that can be used to provide snow products for different regions. GOES data are used for snow fraction detection ([Romanov et al. 2003](#)) and to detect snow and clouds ([Li et al. 2007](#)). For MSG/SEVIRI there are several snow extent products, such as H SAF H31 based on the algorithm presented in [Siljamo and Hyvärinen \(2011\)](#). A fractional snow-cover product obtained by the *FY-2* Visible and Infrared Spin Scan Radiometer (VISSR) instrument is described by [Wang et al. \(2017\)](#).

Instead of relying on one instrument or satellite, the National Oceanic and Atmospheric Administration/National Environmental Satellite, Data, and Information Service (NOAA/NESDIS) Interactive Multisensor Snow and Ice Mapping System (IMS) ([Helfrich et al. 2007](#); [Ramsay 1998](#)) provides high-resolution multisensor snow product. Rather than being fully automatic, the production employs human analysts who merge data from different sources. Validation results for IMS are presented, for example, by [Chen et al. \(2012\)](#).

The pros and cons of different snow products have been studied by many authors. [Frei et al. \(2012\)](#) compare three different snow products (AMSR-E, IMS; MODIS) and discuss the differences of products and future directions in their applications. The merits of the normalized difference snow index

(NDSI) are discussed by [Härer et al. \(2018\)](#), who show that improvements are needed at the local scale. A review by The European Cooperation in Science and Technology (COST) Action ES1404 “HarmoSnow” summarizes different approaches used for snow data assimilation in different fields, such as numerical weather prediction (NWP) and hydrology ([Helmert et al. 2018](#)).

While optical instruments employing visual and IR bands provide high-quality high-resolution data, they have weaknesses. The most important weakness of using optical channels for snow detection is the requirement of cloud-free conditions. Also, during nights the optical instruments have only limited applicability. The high temporal resolution of the instruments on the geostationary orbit helps to mitigate this in midlatitudes, as it is more probable that there are cloud-free moments during the day, but for polar orbiters, night and cloud cover are a serious hindrance. There are regions that can be cloud covered for several days and vast areas may be too dark for snow detection for several weeks during the polar night. However, on favorable conditions, instruments on polar satellites provide excellent spatial resolution.

The challenges of cloud cover in snow detection are discussed for example in [Parajka and Blöschl \(2008\)](#). They use three different approaches to merge *Aqua* and *Terra* MODIS snow products in space and/or time to improve product coverage. They show that more progressive merging decreases cloud covered area, but with reduced accuracy of the snow detection.

Active and passive microwave instruments (radars and radiometers) have advantages in cloud-covered and night conditions. Unfortunately, these methods have also associated restrictions, such as lower resolution (microwave radiometers) or very narrow swath widths (radars). The GlobSnow project ([Metsämäki et al. 2015](#)) provides both snow extent and snow water equivalent (SWE) products for long time periods, but operational near-real-time data do not seem to be available. Many SWE products need ancillary data such as snow-depth observations in the product generation. Such dependence on ancillary data is a limiting factor, for example, in NWP, where independent data are required or preferred.

Even the best satellite-based snow products are useless if the users do not have any indication of the reliability and accuracy of the product. Ideally, satellite snow extent products are validated using daily in situ snow coverage measurements with fine resolution of at least 10%. However, such measurements are not available on operational basis. Regional or local measurement campaigns do not allow continuous global validation.

Fortunately, synoptic weather stations provide in situ snow depths and the state-of-the-ground observations that can be used for satellite snow product validation. While the weather station network provides global coverage in general, there are regions where the network is sparse. When using weather station data for satellite product validation, the representativeness of the observations should be considered. At the moment, weather station observations seem to be the best in situ option for large-scale operational validation of snow products.

There are still limitations in the way the weather stations report snow-cover measurements. Many stations report the snow observations only when snow is present, others do not

provide snow measurements at all. Therefore, a missing snow observation cannot be interpreted as lack of snow at the station. Automatic weather stations can measure snow depth, but many commonly used snow-depth instruments do not provide reliable snow-depth observations of thin (less than 2.5 cm) snow layers.

This lack of snow coverage observations has stimulated creativity and new innovative methods for snow product validation have been described by many authors. For example, [Salvatori et al. \(2011\)](#) suggest the use of fixed webcam photographs for estimation of snow coverage. [Piazzì et al. \(2019\)](#) discuss the use of high-resolution Sentinel-2 imagery to validate medium-resolution snow products (H SAF H10 and H12). They also evaluate the consistency of Sentinel-2 observations based on in situ observations and webcam photographs. Even though webcams and high-resolution imagery can be used for validation, both methods are better suited for regional validation or case studies. [Hyvärinen and Saltikoff \(2010\)](#) study the possibilities to use social media as a source of observations.

The new snow detection algorithm for MetOp/AVHRR presented in this paper is used to produce the first daily operational global snow extent product (H SAF H32) for EUMETSAT. Earlier, an operational snow extent algorithm for MSG/SEVIRI and the corresponding product (currently known as H SAF H31) with limited coverage was published ([Siljamo and Hyvärinen 2011](#)). Both products aim specifically to fill the needs of NWP and hydrological modeling as discussed later in the paper. Extensive trials of MSG/SEVIRI H31 snow extent product in snow analysis have been performed at the Met Office successfully ([Pullen et al. 2018, 2019](#)). Similar work based on MetOp/AVHRR H32 snow extent has been started at the Finnish Meteorological Institute (FMI).

During a reorganization of the snow products in the SAFs, the development of the H31 and H32 snow extent products was transferred to H SAF, but the processing and production of these snow products remain in the Satellite Application Facility on Land Surface Analysis (LSA SAF). The LSA SAF, in general, is described in [Trigo et al. \(2011\)](#). Both the MSG/SEVIRI H31 snow product and the new MetOp/AVHRR H32 snow product based on the algorithm presented in this paper are available via LSA SAF website. The data are publicly available, but data retrieval requires registration for the LSA SAF website. Single product example files can be retrieved from product description pages.

The algorithm version presented in this paper uses the pre-processed data available in the LSA SAF production system, but the use of the algorithm is not limited to the LSA SAF system, as it can be modified to use other AVHRR data sources and similar auxiliary data.

Even though the algorithm was developed for operational use, it can be used to process archived data to produce snow extent datasets covering longer time spans that are needed in reanalysis and similar applications.

2. Development of the MetOp/AVHRR snow extent algorithm

The natural high variability of snow reflectance, caused by the subpixel variability of the surface, makes the development

of a general snow extent algorithm challenging. The resolution of the weather satellite instruments (about 1 km for MetOp/AVHRR) is rather coarse considering the existing variability at scales smaller than satellite resolution (e.g., [Cortés et al. 2014](#); [Dozier et al. 2009](#); [Salminen et al. 2009](#); [Wiscombe and Warren 1980a,b](#)). While snow cover itself may vary considerably inside one satellite pixel, there are also other surface features that must be taken in account.

Vegetation type and density have a significant impact on snow detection. The vegetation can vary from sparse and small (e.g., deserts) to thick and large (dense evergreen forests). There may be small-scale topography and water bodies of different sizes and shapes. Another source of variability is the snow on the canopy, which can vary from thin sprinkled snow to thick crown-covering snow causing damage to the trees. Finally, the snow cover itself can be thin and patchy (melting season, new snow) or thick enough to cover small surface features.

While the properties of snow, vegetation, and surface features cause a significant part of the variability, one must account for the viewing angle, which can have large effect. In nadir, trees may cover the surface below, but at the edge of the satellite scene the large viewing angle means that the obscuring effect of the canopy is considerably larger. In dense forests, there may be several trees between the surface and the satellite. However, deciduous forests are far less affected because the leaf canopy is absent during the snow-covered season.

Another challenge caused by trees and other large objects on the surface are shadows. In high latitudes, where the sun elevation is low during the snow season, shadows can cover large parts of the surface, especially in forests. The fraction of the shadowed surface is not constant and is related to the sun elevation and the type and size of the objects causing the shadows. Also, the effects of the atmosphere and clouds should be taken into account.

Considering all sources of variability in landscape and snow cover inside one satellite pixel, along with the technical limitations of the satellite imagery, development of a purely physics-based snow detection algorithm would be challenging. A further difficulty is caused by the unavailability of detailed satellite pixel-based information regarding these surface features. Thus, the development of an operational snow extent algorithm will benefit from a partly empirical approach based on a high-quality dataset of carefully analyzed satellite images.

Moreover, for an operational product, it is better to use an algorithm where the decisions can be backtracked to a single rule. This restricts using currently popular black-box image classification algorithms, such as deep learning.

The snow-cover extent algorithm developed for the H SAF MetOp/AVHRR snow extent is a sequence of classifiers. The classifiers operate on hand-crafted features, for example, algebraic combinations of results from two channels, which are chosen based on the reflection properties of AVHRR channel wavelengths from various land and cloud classes. These features are crafted such that a simple threshold on a feature can be used to identify a class. These combinations of features and thresholds are called rules, which are based on an

empirical manually classified dataset representing different surface properties as seen from the satellite.

As the aim was to develop a global daily snow extent product, the algorithm had to be divided in two phases (see Fig. 1). The first phase classifies single AVHRR images and creates intermediate single image snow extent product (SC1) on the satellite grid. After the last satellite scene of the day is processed, phase 2 reads all the SC1 products of that day in acquisition time order from the oldest to the newest. These SC1 snow products are then reprojected and merged into the global $0.01^\circ \times 0.01^\circ$ grid. The data in this grid are then tuned and smoothed to create the final daily snow extent product (SC2).

a. The physical basis of snow detection

The differences in reflectances of different soil and cloud types can be used for classification of pixels. Typical solar reflectance spectra of soil, vegetation, and different snow types are shown in Fig. 2, along with an example of the water cloud reflectance spectrum. An example of snow detection based on reflective and absorbing bands is the NDSI. Similarly, the so-called red edge (Horler et al. 1983), which refers to rapid change region in the reflectance spectrum of vegetation, can be used for vegetation detection.

For more accurate detection, the ratios between AVHRR channels were used. Figure 3a shows the ratio of radiances from AVHRR channels 3 and 2 for various surface types, and similarly Fig. 3b for AVHRR channels 3 and 1. The ratio of channels 3 and 2 discriminates snow pixels from snow-free land and cloud pixels, while the ratio of channels 3 and 1 is better at discriminating snow-free land from other classes. From these figures it can be observed that snow and snow-free regions are quite well separated along the axis representing the ratios—thus, these ratios serve well as features for classifying the pixels. Many such ratios, based on the reflection physics shown in Fig. 2, were researched to find the conditions that make up the final list of rules of the algorithm.

b. Algorithm development principles

During the development of the earlier MSG/SEVIRI snow extent algorithm Siljamo and Hyvärinen (2011) the driving philosophy was to avoid incorrect classifications, even if it leads to a smaller number of classified pixels and a larger number of unclassified pixels. This approach was discussed with and recommended by the NWP experts at FMI (C. Fortelius and L. Rontu, in several discussions since 2006) and the Met Office (S. Pullen, in several discussions since 2013). Initialization of snow observations for NWP is discussed, for example, in de Rosnay et al. (2014).

There is slightly different emphasis between NWP and hydrological modeling. In hydrological modeling, thin snow layers (~ 1 cm or less) are often irrelevant as the water content in them is insignificant. However, in weather models, the presence of a thin snow layer, which alters the radiative properties of the surface, is often important.

Further discussions with the members of the NWP community using different weather models in different institutions

FIG. 1. General flowchart of the MetOp/AVHRR snow extent algorithm. The algorithm has two separate phases. Unit 1 processes a single satellite scene (in this case one PDU) and produces an intermediate SC1 snow product. After the last scene of the day has been received and processed, unit 2 merges all SC1 snow products to the global grid in acquisition order. After smoothing based on the 3×3 pixels around each pixel, the final daily snow product (SC2) is generated.

enhanced the view that the main points of the snow product targeting the meteorological community and especially NWP are as follows:

Directness: The number of preliminary steps before actual snow detection should be as small as possible. As the aim is to detect snow and snow-free pixels, there is no reason to use cloud masking as a preliminary step as it does not add value and could introduce errors.

Accuracy: Accuracy is preferred to coverage. While in many other applications large coverage is preferred, in NWP, missing data are easier to handle than misclassifications.

Single-source data: Only satellite data from a single instrument will be used. Limited use of static data (such as land-cover classification) and other products based on the same instrument and processed in the same system (such as LSA SAF LST) is possible.

Availability: The NWP community prefers operational products as there is at least some certainty of data availability in the future.

FIG. 2. Qualitative comparison of different ground types and the water cloud reflectances with the solar AVHRR channels. The reflectances of ground types are based on Baldridge et al. (2009), and the reflectance of a water cloud is based on an arbitrary image taken on 0208 UTC 12 Jun 2010 near Barbourville, Kentucky, using the hyperspectral Hyperion instrument on board the *Earth Observing-1 (EO-1)* satellite. The effective particle sizes of fine, medium, and coarse snow are 24, 82, and 178 μm , respectively. Note that, because ground type and cloud reflectances are based on different sources, only qualitative comparison is possible.

These points form the current development philosophy behind the new MetOp/AVHRR snow-cover algorithm. The new MetOp/AVHRR algorithm does not use third-party or other independent cloud masking, does not try to force classification of all pixels, is based on only MetOp/AVHRR data and has reached operational status.

c. Data sources for the algorithm

The inputs used in the algorithm are presented in Table 1. The algorithm utilizes the top-of-atmosphere radiances of three AVHRR channels (1, 2, and 3A) and brightness temperatures of two channels (4 and 5). These data are preprocessed in the LSA SAF processing system and then delivered for further processing in product dissemination units (PDU).

In addition to the radiances and brightness temperatures, there is preprocessed auxiliary data in the LSA SAF production system, such as pixel coordinates, elevation, sun and satellite angles, land-cover type, and water mask (see online documentation available in the LSA SAF website for details). To separate forests and open areas, the algorithm employs the International Geosphere–Biosphere Programme (IGBP) land-cover type (Loveland and Belward 1997) by the U.S. Geological Survey (USGS), which is also preprocessed and readily available in the operational LSA SAF production system (see Table 2). Other land-cover datasets can be used as well. Land surface temperature (LST) provided by the LSA SAF production system is employed to remove some misclassifications.

FIG. 3. The tentative discrimination of (a) “snow” and (b) “no snow” pixels using only the ratios of the solar channels of the AVHRR instrument. The densities of two land classes “snow” and “no snow” and two cloud classes “ice cloud” and “water cloud” are estimated using the subjective classification dataset.

d. The iterative algorithm development

Development of the algorithm began with subjective classification of selected areas in representative MetOp/AVHRR images to create a development dataset. Selected areas were classified to eight different classes (snow, snow-free, ice clouds, water clouds, mixed clouds, transparent clouds covering snow or snow-free surface, and water). Clouds were classified to several groups from snow detection perspective to help in the algorithm development. For this, 69 false-color red–green–blue (RGB) images were used from December 2007 to March 2010 to classify approximately 610 000 pixels. This subjectively classified dataset was used as a basis for the algorithm development.

The algorithm consists of rules based on channel differences and ratios, which aim to differentiate snow-covered and snow-free surfaces when possible. Any pixel that does not match one of the rules is left as unclassified. This group includes pixels where classification is too difficult or impossible for any reason, such as darkness (sun elevation angle), probable cloud cover, and difficulties deciding whether to classify pixels as snow-free or snow covered. As the algorithm aims to detect snow and snow-free surfaces, it is often not possible to determine the exact reason why a certain pixel is not classified.

Based on the subjectively classified dataset, the first set of empirical classification rules was created. The development environment provided tools for comparison of different channels, analysis of the behavior of different classification rules, and general analysis of relative importance of different rules. Once the initial set of classification rules was prepared, different satellite images from October 2014 to June 2015 were used to identify misclassifications and challenging cases around

TABLE 1. List of the inputs for the H SAF MetOp/AVHRR (H32) snow extent algorithm.

Input type	Code	Description
Satellite channels	R_1	Radiance in channel 1 (0.58–0.68 μm)
	R_2	Radiance in channel 2 (0.725–1.00 μm)
	R_3	Radiance in channel 3A (1.58–1.63 μm)
	R_4	Radiance in channel 4 (10.30–11.30 μm)
	R_5	Radiance in channel 5 (11.50–12.50 μm)
	T_{B4}	Brightness temperature in channel 4 (10.30–11.30 μm)
Angles	T_{B5}	Brightness temperature in channel 5 (11.50–12.50 μm)
	SAA	Sun azimuth angle
	SZA	Sun zenith angle
	VAA	Satellite azimuth angle
Other	VZA	Satellite zenith angle
	LC	Land-cover type
	LST	LSA SAF land surface temperature ($^{\circ}\text{C}$)
	ΔT_B	$T_{B4} - T_{B5}$

the world. Classification rules were updated to correct misclassifications and to find ways to extend the classification to initially unclassified areas.

In this development phase, nearly all of the land surface of the Earth on select days (e.g., all images of 19 February and 26 March 2015) was analyzed to find any suspicious classifications. External data sources, such as Google aerial and street view images, were used to get a better understanding of local conditions (e.g., surface type, vegetation). In addition, MODIS images were used to estimate the current surface status subjectively. The rules were adjusted in many cases so that challenging pixels were excluded from classification.

This first candidate for the final SC1 algorithm allowed development of the algorithm for the second phase (SC2), which reprojects and merges the results of the SC1 algorithm to a global daily product in required latitude–longitude grid. The SC1 products are processed in order from the oldest to the newest based on the acquisition time. Pixels in the SC1 product are reprojected to the nearest pixel in the global grid and the current classification in that pixel is updated either from default value (unclassified) or the previously set classification value (snow, no snow, partial snow). In practice, there are still small gaps and single-pixel misclassifications in the product after initial merging. Therefore, final smoothing based on 3×3 pixels around each pixel is used to generate the final product.

This global product candidate was checked visually against other data to find potential misclassifications, which were then studied in more detail using the SC1 product and external sources. Some rules in the algorithm were slightly adjusted.

e. The final algorithm

The final rules in the SC1 product, along with certain special conditions, are presented in Tables 3 and 4. As this product is targeting only land surfaces, a water mask provided by the LSA SAF production system is used. The notes column in Table 4 shortly describes the physical interpretation for each rule.

TABLE 2. IGBP Discover dataset land-cover classification system.

Class	Class name
1	Evergreen needleleaf forests
2	Evergreen broadleaf forests
3	Deciduous needleleaf forests
4	Deciduous broadleaf forests
5	Mixed forests
6	Closed shrublands
7	Open shrublands
8	Woody savannas
9	Savannas
10	Grasslands
11	Permanent wetlands
12	Cropland
13	Urban and built-up
14	Cropland/natural vegetation mosaics
15	Snow and ice
16	Barren
17	Water bodies

The rules used in the generation of the daily product are presented in Table 5.

After the algorithm was implemented in the LSA SAF production system and the product generation started, up-to-date snow-cover data were available from the LSA SAF system for testing purposes. For further validation, a longer time period was required. For that, LSA SAF processed MetOp/AVHRR data since 1 January 2015. After successful internal validation, the product reached operational status in early 2018.

3. Validation strategy and data

Quantitative product validation using surface observations was the final phase of the product development. This phase aims to confirm that the system produces valuable and reliable information, so that the product can be accepted for operational use. For this paper, the original internal 22-month validation period was extended to cover the period from January 2015 to March 2019 (51 months) to give more insight into the product and its properties.

a. Visual inspection of products

In the Fig. 4, an example of the product for 10 April 2017 is presented. At that time, northern snow cover is melting, but there is still snow in the northern parts of Eurasia and America. Snow is also present in the mountain ranges and other high elevation areas. Even though the snow cover is very well detected, the limitations of an optical algorithm are obvious as cloud cover creates patches of unclassified land. Darkness is the reason for the unclassified area, especially in the Antarctic. There are also some unclassified stripes crossing the equator, which are not covered by daytime satellite swaths. More product examples are presented in Figs. S1–S10 in the online supplemental material.

For an initial subjective evaluation, false-color RGB images from MetOp/AVHRR and other satellite instruments, such as MODIS, were used to estimate the quality of the snow

TABLE 3. List of special conditions in the algorithm rules. The special condition is true when the condition in the middle column is true.

Name	When true	Notes
slc	LC = 2, 5–12, 14	IGBP; see Table 2
creg0	lat < −60° or lat > 60°	High latitudes
creg1	lat < −45° or lat > 58° or [lat > 45° and (lon < −30° or lon > 30°)]	Southern midlatitudes; northern high latitudes
cMid	$h \geq 1500$ m and (lat < −35° or lat > 35°)	Midlatitude non-Europe
cMo	$h \geq 3000$ m	Mid- or high-latitude highlands
creg4	month = 1–5 and {lat < −35° or lat > 60° or [lat > 35° and (lon < −30° or lon > 30°)]}	Mountains
creg	creg0 or creg1 or cMid or cMo	
Scold	[(creg0 or creg1 or cMid) and month = 1–5] or cMo	
tropic	$h \leq 3000$ m and lat < 20° and lat > −20°	
moderate	$h \leq 2500$ m and lat < 40° and lat > −40°	
for	LC = 1–6, 8, 14	Forest by IGBP; see Table 2
nonfor	LC = 7, 9–13, 15–	Nonforest by IGBP; see Table 2

product. Visual inspection of several examples did not reveal any obvious problems, although there were slight differences in the details. Figure 5 shows parts of Europe in 15 April 2018, in false-color (channels 3/6/7) MODIS RGB image, reprojected false-color MetOp/AVHRR (channels 1/2/3A) and reprojected H32 snow product for the same day. The limitations of the satellite algorithm reduce the number of classified pixels

especially in difficult conditions, but the speed of automatic snow detection balances this in practical applications where fast and reliable products are essential, such as NWP.

b. Surface observations for validation

The synoptic weather station observations were retrieved from the FMI observations database. It has an

TABLE 4. List of classification rules in the SC1 algorithm. If the condition is true, the snow-cover status is set to “value” (SN = snow, NS = no snow, PS = partial snow, WA = water, and UC = unclassified). These rules are applied sequentially from the top in the order presented, and the final snow-cover classification is the value in effect after the last rule. For definitions, see Tables 1 and 3. Logical AND is marked by $\wedge$, and logical OR is marked by $\vee$.

Rule condition	Value	Rule	Notes
Set default value	UC		
nonfor $\wedge R_2/R_1 < (-0.2T_{B5} + 57) \wedge R_3/R_1 < 0.002T_{B5} - 0.45 \wedge T_{B5} < 272.6 \wedge R_2/R_1 > -0.05T_{B5} + 15.5$	PS	R1	Cold surface, relatively bright
$T_{B4} > 290$	NS	R2	Too warm for snow
nonfor $\wedge R_3/R_1 > 0.134$	NS	R3	Open, near infrared vs visible too high for snow
creg $\wedge$ nonfor $\wedge R_2/R_3 > (-2T_{B4} + 585) \wedge T_{B4} < 277$	SN	R4	Cold, snow possible, R_2/R_3 indicates snow
Scold $\wedge$ nonfor $\wedge R_2/R_3 > (-2T_{B4} + 574) \wedge T_{B4} > 256.5 \wedge T_{B4} < 269.7$	SN	R5	Amends R4 in problem cases
Scold $\wedge$ for $\wedge R_2/R_1 > (-0.1T_{B5} + 29.5) \wedge R_2/R_1 < 2.86 \wedge T_{B5} < 280$	PS	R6	Cold forest, red edge less distinct for snow-free
$R_3/R_1 < 0.045 \wedge T_{B4} > 280$	NS	R7	Warm, reflectance indicates snow, set no snow
creg4 $\wedge (R_3 - R_2)/(R_3 + R_2) < -0.975 \wedge T_{B4} < 279 \wedge T_{B4} > 240$	SN	R8	Low R_3/R_2 with cold surface, no coldest clouds
for $\wedge R_3/R_1 > 0.135$	NS	R9	Forest, distinct red edge, snow highly unlikely
creg $\wedge R_2/R_3 > 120 \wedge T_{B4} < 276$	SN	R10	Similar to R8
creg $\wedge$ for $\wedge R_2/R_3 > 72 \wedge T_{B4} > 253$	SN	R11	North forest, no red edge, warm for ice clouds
Scold $\wedge$ for $\wedge R_2/R_3 > 45 \wedge T_{B4} > 263$	SN	R12	North forest, no red edge, warm for ice clouds
Scold $\wedge [(R_2/R_3 > 120 \wedge T_{B4} < 254) \vee (R_2/R_3 > 220 \wedge T_{B4} < 280) \vee (R_2/R_3 > 50 \wedge T_{B4} > 267 \wedge T_{B4} < 276 \wedge \Delta T_B < 1.5)]$	SN	R13	Snow, filter out clouds
$(T_{B5} > 280 \wedge R_2/R_1 > 2)$	NS	R14	Clear red edge, warm, snow very unlikely
$T_{B4} < 242 \wedge R_2/R_3 < 68.8$	UC	R15	Very cold surface with an indistinct R_2/R_3
$\Delta T_B > 4 \wedge R_3/R_1 > 0.09 \wedge R_3/R_1 < 0.11$	UC	R16	Remove misclassifications
VZA > 60	UC	R17	Viewing geometry is too challenging
SZA > 80	UC	R18	Illumination geometry is too challenging
tropic $\wedge$ slc $\wedge$ (SN $\vee$ PS)	UC	R19	No snow near-equatorial lowlands
moderate $\wedge (T_{B4} + T_{B5})/2 < 253 \wedge$ (SN $\vee$ PS)	UC	R20	Unreliable cold snow surface
LST $\geq 293.15 \wedge$ (SN $\vee$ PS)	NS	R21	Trust LST over snow retrieval
(SN $\vee$ PS) $\wedge R_1 < 1.2/\cos^2(\text{SZA}) \wedge R_2 < 1.2/\cos^2(\text{SZA}) \wedge R_3 < 0.02/\cos^2(\text{SZA})$	UC	R22	Unclassify too-dark snow pixels
If water pixel in water mask	WA	R23	No retrieval over water

TABLE 5. List of the smoothing rules used in the generation of the daily (SC2) product. The number of classifications in the 3×3 grid around a pixel is counted, and the values are used to define the final classification in each pixel. The F , W , U , S , P , and N are the numbers of nonprocessed, water, unclassified, snow-covered, partially snow-covered, and snow-free pixels, respectively. These rules are used one after the other from the top, and the final daily classification is the classification in effect after the last rule. Logical AND is marked by $\wedge$, and logical OR is marked by $\vee$.

Rule condition	Value set	Rule
Set default value	Original value	D1
$U + F > 4$	UC	D2
$S + P < 2 \wedge N < 2 \wedge U > 2$	UC	D3
$W > 3 \wedge S + N + P + U = 0$	WA	D4
$W + F \leq 3 \wedge$ $S + P = 0 \wedge N > 2$ $\wedge W + F = 0$	NS	D5
$S + P > 3 \wedge N = 0$	SN	D6
$S + P = 0 \wedge N > 2$	NS	D7
$S + P > 3 \wedge N > 2$	PS	D8

adequate global coverage and provides easy and fast access to the observations. For validation, stations that had over 20 snow-depth or state-of-the-ground observations between 1 January 2015 and 31 October 2016 were selected. Snow-depth, state-of-the-ground, and 2-m air temperature observations were retrieved from January 2015 to March 2019 from 4240 stations. Figure 6 shows the locations of the selected weather stations. Although there are considerable gaps in the global coverage, the regions of seasonal snow are very well covered.

The state-of-the-ground measurements (as defined in WMO 2015) are not widely available from weather stations, mainly because they are manual observations and an increasing number of weather stations have been converted to automatic operations. However, when this measurement is available, it is well

suited for snow extent product validation. Although it does not provide exact information about snow coverage, it provides an estimate (snow-free, less than half, over half, completely snow covered) that is better than estimated on/off snow coverage based on point observations of snow depth.

The total number of individual observations retrieved from the database was about 68 million. However, some of the stations reported snow-cover data only intermittently while many others provided hourly data. For validation, all observations from each station were merged to a single set of daily observations. The highest snow depth of the day or the largest coverage value of the state of the ground was selected as the daily observation of each station. After this, there were about 6.2 million daily observations of snow cover or air temperature from these stations. Of these observations, about 4.1 million included either snow-depth or state-of-the-ground observations or both and the rest were only temperature measurements.

As a part of the processing of observations, both snow-depth and state-of-the-ground observations were converted to three classes: snow, partial snow, and no snow. Both measurement types have different challenges that had to be taken into account in the conversion.

Quite often there are no-state-of-the-ground observations. Snow depth is often reported only when snow is present. This makes the snow depth a practically useless indicator of a snow-free surface, because a missing snow-depth measurement can be either from a snow-free station or more often from a station that does not measure or report snow depth at all.

There are also different practices of reporting snow depth. Some stations report snow depth in meters, others in centimeters. Some stations use zero or negative values to indicate no snow, others use similar values for partial snow. The state-of-the-ground values are more straightforward to use, but the definitions itself required interpretation.

The state-of-the-ground code values 0–9 (“without snow or measurable ice cover”) are used as no snow observations; code

FIG. 4. An example of the H SAF H32 snow extent product for 10 Apr 2017. The algorithm catches the general features of the snow cover well, but the limitations of an optical algorithm can be seen as black irregular patches in cloud-covered areas. Unclassified areas near the poles are caused by low sun elevation and, in the Antarctic, polar night. There are also some unclassified stripes across the equator where the daytime satellite swaths do not cover all of the surface.

FIG. 5. A false-color MODIS RGB (channels 3/6/7), reprojected false-color MetOp/AVHRR RGB (channels 1/2/3A), and reprojected H SAF H32 snow extent for 15 Apr 2018. The MODIS image (by NASA Worldview) shows that there is snow (dark red). There are also different clouds (white or pink) in the area. MetOp/AVHRR shows snow in dark yellowish shades and clouds as light blue or light yellow. The general features of the snow cover are detected, but there are potential misclassifications, especially in regions where snow is presumably melting.

values 11,12, 15, and 16 are used as partial snow cover; and values 10, 13, 14, and 17–19 (“with snow or measurable ice cover”) represent snow-covered surfaces. Snow-depth values are classified as snow if snow depth is greater than zero and snow-free if snow depth is less than zero. Sometimes there are special values that indicate partial snow cover, and these have been converted accordingly. Zero values are converted to partial snow cover as the value should be used to report that

there is no snow at the measuring point but there is still some snow at the vicinity. Currently, there is some uncertainty as sometimes zero snow depth is used to report no snow.

Then, a list of daily observations in each station was created. Because the selected validation measures need binary data (no snow/snow), different options for treatment of partial snow-cover values were tested both in the satellite product and surface observations: converting partial snow

FIG. 6. The number of weather stations used in the validation in 1° × 2° grid cells.

to snow-free (“no snow”), converting partial snow to full snow cover (“snow”), and excluding partial snow (“off”).

The daily observations were then converted to a single value that represents the snow coverage at each station. If both snow-depth and state-of-the-ground values were available but conflicting, the observation was marked as such and excluded from validation. Two consistent observations or the only observation (snow depth or state of the ground) was used as the daily value for the station and included in the validation dataset (about 4.1 million daily observations).

In many cases, clouds or inadequate solar illumination prevent proper classification of the satellite pixels and therefore the number of classified pixels varies from day to day. No attempt was made to mitigate this and thus, if the pixel is not classified as snow-free, partially snow covered, or snow covered by the algorithm, it is not used in the validation.

c. Validation measures

For validation of the snow product with surface observations, the common validation measures computed from a 2×2 contingency table (Table 6) were used. Then, following the terminology of Hogan and Mason (2012), cases where the satellite detected snow are either hits a , when the satellite correctly detected snow, or false alarms b , when surface observation contradicts it. Similarly, cases where the satellite detected snow-free surface are either correct rejections d or misses c when surface observation shows the presence of snow.

However, the snow cover has a clear seasonal cycle, and during summer there are relatively few snow observations relative to no-snow observations ($d \gg a + b + c$). This complicates the validation of snow product because most common validation measures degenerate to trivial values when the number of cases in one category is very small compared with the other.

It is easy to see that the commonly used measure proportion correct (PC),

$$PC = \frac{a + d}{a + b + c + d}, \quad (1)$$

tends to 1 if d dominates. For other measures this might be not as self-evident, but it can be seen when the measures are shown as the function of two conditional probabilities hit rate H ,

$$H = \frac{a}{a + c}, \quad (2)$$

and false alarm rate F ,

$$F = \frac{b}{b + d}, \quad (3)$$

and the base rate s ,

$$s = \frac{a + c}{a + b + c + d}. \quad (4)$$

In the perfect analysis, H should be 1 and F should be 0.

Now PC is

TABLE 6. Contingency table of the comparison between two categorical snow analyses. The symbols a – d represent the number of cases in each group.

Analysis 1	Analysis 2 (baseline)	
	Snow	No snow
Snow	a (hit)	b (false alarm)
No snow	c (miss)	d (correct rejection)

$$PC = (1 - F)(1 - s) + Hs, \quad (5)$$

and dominating d implies $s \rightarrow 0$ and $F \rightarrow 0$, so PC tends to 1 as it should.

Slightly counterintuitively, the often-used replacement for PC, critical success index (CSI), which ignores correct rejections and therefore is used in cases in which d dominates, also degenerates. Its definition is

$$CSI = \frac{a}{a + b + c} = \frac{H}{1 + F(1 - s)/s}, \quad (6)$$

and when there are very few snow observations ($s \rightarrow 0$), CSI will tend to zero. Also the Heidke skill score (HSS; the PC corrected for random hits),

$$HSS = \frac{2(ad - bc)}{(a + c)(c + d) + (a + b)(b + d)}, \quad (7)$$

will tend to zero. On the other hand, the false alarm ratio (FAR),

$$FAR = \frac{b}{a + b} = \left[1 + \left(\frac{s}{1 - s} \right) \frac{H}{F} \right]^{-1}, \quad (8)$$

will tend to 1, while in the perfect analysis it should be 0.

A measure that does not degenerate is the symmetric extremal dependence index (SEDI),

$$SEDI = \frac{\ln F - \ln H + \ln(1 - H) - \ln(1 - F)}{\ln F + \ln H + \ln(1 - H) + \ln(1 - F)}, \quad (9)$$

which in the perfect analysis should be 1. The SEDI can be used to assess whether there is a real drop in quality of the snow product in summer or if it is due to the characteristics of the validation measures used. If either F or H is zero or unity, SEDI is not defined. When either F or H was zero, a very small number (0.0001) was added to the numerator and denominator in the calculation of H and F . This was done purely for visualization purposes so that all cases would be plotted in the figures below.

4. Validation results

In Fig. 7, pixel counts of each class are presented. There are days with some missing SC1 images and days without any data that can be seen in the figure as spikes and vertical stripes. In general, there are a significant number of pixels that can be classified as snow covered or snow-free.

The validation results for the full validation period (from January 2015 to March 2019) are presented in Table 7 for three

FIG. 7. Daily classification distributions in the MetOp/AVHRR snow extent product from January 2015 to April 2019. The main cause for vertical stripes and spikes is missing input data on those days. Seasonal snow of the large northern continents dominates the global snow extent.

different sets of stations. All stations are used in the global set of stations, and Europe excludes stations outside Europe. The third set of stations, called “variable,” includes stations that have reported both snow-free and snow-covered observations during the validation period, that is, these stations are probably in areas where snow is present occasionally.

Overall, the validation results are very good ($HSS > 0.90$ and $SEDI > 0.95$ for global validation). Thus, the MetOp/AVHRR snow extent algorithm is shown to produce realistic estimates of the snow cover. Full validation period hit rate (H) and false alarm rate F are very good as well as PC, HSS, and CSI, which all suggest that there are no large-scale systematic difficulties in the algorithm. Especially SEDI suggests that the algorithm produces good results during the summer, which could be rather challenging period due to small number of snow-covered pixels.

As expected, there are differences when the results of different treatments of partial snow are compared. Table 7 suggests that either excluding partial snow or converting these pixels to no snow produce the best scores, whereas conversion of partial to full snow cover deteriorates the validation score. This implies that partial snow cover is both difficult to measure realistically and to classify automatically. However, it seems that areas that are partially snow covered have more similarities with a snow-free surface than a snow-covered surface.

The validation of thin snow layers is especially challenging. Thin snow layers (< 2 cm) do change the spectral properties of the surface as can be easily seen by a human observer. By selecting cases where stations report no or thin snow depth with the state-of-the-ground observation available, algorithm capabilities with thin snow layers can be estimated. There are 7857 cases where snow depth is 1 or 2 cm. If we use these cases

TABLE 7. Global validation results for the period from January 2015 to March 2019. The global, Europe, and variable regions and all three partial snow treatments (no snow, snow, and off) are presented. Differences between versions are small, even though treating partial snow as snow seem to be slightly worse than the other two options. The number of cases is given in parentheses. The validation measures are defined in section 3c.

	Global (4240 stations)			Europe (2021 stations)			Variable (2286 stations)		
	No snow (1 327 910)	Snow (1 337 570)	Off (1 328 561)	No snow (727 098)	Snow (729 781)	Off (725 434)	No snow (948 663)	Snow (955 943)	Off (949 833)
Bias	0.989	0.929	0.964	1.009	0.913	0.962	1.031	0.954	0.990
H	0.928	0.884	0.920	0.933	0.857	0.913	0.974	0.918	0.958
F	0.009	0.008	0.007	0.005	0.005	0.004	0.006	0.005	0.004
FAR	0.062	0.049	0.045	0.076	0.062	0.051	0.055	0.038	0.032
PC	0.983	0.977	0.983	0.991	0.985	0.991	0.992	0.987	0.992
CSI	0.875	0.845	0.882	0.866	0.811	0.870	0.921	0.886	0.929
HSS	0.924	0.903	0.928	0.923	0.888	0.925	0.955	0.932	0.959
SEDI	0.978	0.963	0.976	0.981	0.957	0.977	0.993	0.977	0.989

and convert partial snow cover to full snow cover before calculating validation measures, the SEDI ~ 0.65 , but in that case many thin snow-cover cases are excluded. Because many thin snow cases have been set to 0 cm in observation quality control, it may be advisable to include 0 cm snow depths with state-of-the-ground observations in the analysis even though that will add many patchy snow-cover cases to the analysis. In this case, the total number of thin snow observations is 354 075. The distribution of observations is strongly skewed, and validation result interpretation must be done with great care, but SEDI is still ~ 0.8 . These results suggest that the algorithm does detect thin snow layers reasonably well. Comparison of thin snow layer capabilities of different satellite snow products in a dedicated validation paper would be beneficial.

Further analysis of partial snow cover both in the algorithm and in the observations would benefit from more detailed observations, such as high-resolution imagery and observations of the temporal development of the snow cover. Unfortunately, such observations are not widely available. Also, more detailed analysis of the excluded cases of conflicting snow-depth and state-of-the-ground observations in the weather stations could improve the understanding of the weather station observations.

The results do not differ significantly when different regions are compared. They may be slightly better in the variable region, which includes only those stations that recorded both snow-covered and snow-free observations. This may imply that the algorithm does perform very well in the areas that have seasonal snow.

The daily validation measures were calculated to create time series that are presented in Figs. 8–10. These three figures differ in the way partial snow cover is treated. In Fig. 8, partial snow-cover observation and product classifications are all converted to snow-free (“no snow”). Figure 9 shows the results when partial snow is converted to full snow cover (“snow”) and Fig. 10 show the results when partial snow is excluded from validation (“off”).

In the panels of all three figures different validation measures are presented. On the top-left panel, also the number of snow pixels each day is presented. Each time series uses the same color coding, where the dark green data points mark the days on which $d \leq 20(a + b + c)$ (i.e., the proportion of correct snow-free observations is not too large), light green marks the days on which $d > 20(a + b + c)$, and orange marks the days on which $d > 200(a + b + c)$. This color coding is used to emphasize that many validation measures can be misleading when the distribution of correct observations is strongly skewed.

These strongly skewed distributions are common during the northern summer when seasonal snow cover in the well-lit regions is at its minimum. During summer, a small number of misclassifications caused by, for example, thunderstorms, peculiar surface features or unrepresentative surface observations, can change the results significantly even when practically all other classifications are correct. This can be seen very well in the PC and F , which are nearly perfect during the summers even though more sophisticated measures (HSS and SEDI) show high dispersion of values. At the beginning of the winter, when the snow-covered area grows and the number of hits

a grow and correct rejections d decrease, the validation measures improve greatly and stay at a high level most of the winter and spring.

The relatively rare misclassifications during the northern summer do show in some of the validation measures, but in general, the algorithm produces very good results throughout the year and excellent results during the northern winter and spring when snow cover has the largest impact on weather.

5. Discussion

a. H32 and other products

Many optical snow detection algorithms rely on detecting cloud-free regions before the actual snow-cover classification. The approach used in the algorithm presented in this paper bypasses cloud detection and associated potential misclassifications by trying to find directly snow-covered and snow-free pixels. When that fails, the pixel is considered unclassified without any further analysis for the reason.

The strengths and weaknesses of different satellite snow extent products can be compared. Although multisource products (such as IMS) provide much better coverage, the quality and accuracy of multisource products is often difficult to analyze, especially if the source of the snow-cover estimate in each pixel cannot be traced. Single-source products (such as H31 and H32) do have more data gaps, but when available the accuracy is well defined and consistent across the coverage area. This is beneficial for users who prefer consistent and predictable behavior of the product, such as NWP and reanalysis.

The accuracy of the algorithm presented here should be compared with the accuracy of other snow extent products available. Literature-based accuracy estimates of some operational snow extent products are presented in Table 8. A similar table has been presented earlier by Surer et al. (2014). The accuracy of the present product is comparable with accuracy other products. For most products, as for the present algorithm, the PC and the H are usually over 90%, and F less than 10%. However, as can be seen in Figs. 8–10, trying to compress the quality of a product in one number, or even a set of numbers, can be a difficult as the values of verification metrics vary with the season, terrain type, illumination conditions, and other physical properties of the surface. A more detailed comparison at the pixel or a gridpoint level would be most useful, but that is a task for another article. Even more useful, but also a very demanding, task would be to assimilate different products in NWP and see which product improves the forecasts most.

Another factor to keep in mind is the level of human interference in the product generation. A fully automatic product provides consistency and speed. An experienced human analyst may improve the product quality but with added uncertainty of the product behavior and error statistics, especially in areas that are not prioritized in the analysis. As manual analysis of the snow cover is time-consuming, it is unlikely that human analysis can be extended to full global coverage in operational products.

FIG. 8. MetOp validation measure time series, with partial classifications converted to snow-free. Each day is color coded to indicate whether the proportion of correct snow-free observations is so large that most of the validation measures degenerate (indicated by light green and, in the most extreme cases, orange). Most of the snow observations are from the northern seasonal snow region, which dominates in the time series. The results are excellent during the northern winter.

Global, Metop/AVHRR, partial = snow

FIG. 9. As in Fig. 8, but partial classifications are converted to full snow cover.

Global, Metop/AVHRR, partial = off

FIG. 10. As in Fig. 8, but partial classifications are excluded.

TABLE 8. Some operational snow products that are available currently, with published accuracy estimates (disk = Geostationary satellite detection disk, NH = Northern Hemisphere).

Product	Data source	Start year	Resolution	Time res.	Coverage	Reference/accuracy
H SAF H32	MetOp (AVHRR)	2015	0.01°	Daily	Global	Described in this paper
H SAF H31	MSG (SEVIRI)	2008	3 km (nadir)	Daily	Disk	The PC against IMS is ~99% calculated over the whole year, but somewhat lower during the NH winter only; the $H \approx 80\%$ and $F < 10\%$. (Siljamo and Hyvärinen 2011)
VIIRS snow products	<i>Suomi NPP</i> ; <i>NOAA-20</i>	2012	375 m/0.05°	Daily	Global	Comparable to MODIS, but VIIRS has the potential to map snow-cover area more accurately (Riggs et al. 2017; Thapa et al. 2019)
MODIS snow products	<i>Terra</i> (MODIS); <i>Aqua</i> (MODIS)	2000	500 m/0.05°	Daily; 8-day; monthly	Global	The PC against in situ snow observations in Collection 5 is about ~93%, but lower accuracy is found in forested areas and complex terrain and when snow is thin and ephemeral; very high PC, up to 99%, may be found in croplands and agricultural areas (Hall and Riggs 2007); note that Collection 6 has been published (Riggs et al. 2017)
NOAA/NESDIS IMS	Multisource	1998	4 km	Daily; 8-day	NH	Described in Helfrich et al. (2007) and Ramsay (1998); the PC against in situ snow observations is 80%–90%, $H \approx 95\%$, and $F 0\%$ –20% during the NH winter, but values vary with season (Brubaker et al. 2005; Chen et al. 2012)

b. Validation

Validation with snow depth and the state-of-the-ground observations from surface weather stations shows good agreement with the MetOp/AVHRR H32 snow product especially during winter and spring. Thus, the MetOp/AVHRR snow extent product can provide new and reliable data about snow extent. Similar benefits can be achieved with the previously published MSG/SEVIRI snow extent product with better temporal resolution, which helps to reduce the limitations caused by cloud cover and short daylength. However, the resolution of the MSG/SEVIRI product in the northern parts of Europe is limited. The MetOp/AVHRR product provides global coverage and much better resolution in polar regions. However, cloud cover and availability of daylight may prevent snow detection.

All data used during the development work was gathered prior to 2016. The days studied extensively during the development were not excluded from validation. However, the data from 2016 onward were used for validation only. The validation metrics, based on surface observations as the ground truth, shown in Figs. 8–10, remain consistent throughout the whole validation period of 2015–19. This result clears concerns of overfitting, which can result from limited temporal variance of the development data.

c. Artificial observations

Because the lack of snow cover is not always reported, the idea of using artificial snow-free surface observations based on temperature was tried. Unfortunately, cold temperatures are

common even when the surface is snow-free. For that reason, generation of artificial surface observations in the trial runs was limited to cases in which the daily minimum and maximum temperatures are high enough to ensure that the surface is truly snow-free ($T_{\min} > 5^{\circ}\text{C}$ and $T_{\max} > 10^{\circ}\text{C}$).

Validation measures were recalculated based on this artificial secondary dataset, but the results were essentially identical (differences less than 1%) with the results based on actual snow-cover observations. Thus, the idea of artificial snow-free surface observations was rejected, and only actual snow measurements were used in this study. However, the use of generated artificial snow observations based on other measurements could be beneficial in filling the gaps in snow observations and might be worth a further study.

d. Future

The current operational algorithm version would benefit from further analysis of the relative importance of individual classification rules. Some rules may need adjusting or may be redundant in the current form. Nearly all seasonal snow is in the Northern Hemisphere, but snow in the Southern Hemisphere should get more attention in further snow product development.

Continuous availability of near-real-time products, including snow products, is important for the NWP community. While current H SAF H32 and H31 products will be generated until the end of the MSG and MetOp programs, new products based on the same development principles are expected to be available for the next generation of EUMETSAT weather

satellites [Meteosat Third Generation (MTG) and MetOp-SG]. While the same development philosophy will be used, new tools such as machine learning may be used as an aid to speed up the development work. Even though black-box models, such as neural networks, are popular at the moment, they may not be the best idea for production due to the difficulty of backtracking solutions. However, machine learning tools providing visualizable solutions, such as random forests or support vector machines (Bishop 2006), may be useful in future snow-cover algorithm development.

6. Conclusions

In this article, a new MetOp/AVHRR-based snow extent algorithm and product are introduced. The algorithm is used operationally to produce daily global H SAF MetOp/AVHRR H32 snow extent product. The algorithm applies the same approach of avoiding preliminary cloud masking before actual snow-cover recognition, which was used successfully earlier in the former LSA SAF, current H SAF MSG/SEVIRI H31 snow extent product. The algorithm aims to be as independent as possible of any external datasets, algorithms, and products.

Validation results based on snow-depth and state-of-the-ground observations in weather stations are very good ($HSS > 0.90$; $SEDI > 0.95$) and suggest that the algorithm produces realistic estimates of snow cover especially during the northern winter and spring. Together with the H SAF MSG/SEVIRI snow extent product, which employs a similar algorithm (Siljamo and Hyvärinen 2011), these two products provide excellent snow coverage data of the world and especially Europe.

Acknowledgments. This work was financially supported by the H SAF project, co-funded by EUMETSAT. We are grateful to Drs. Terhikki Manninen, Elena Saltikoff, and Kati Anttila for their valuable comments during the writing of this paper. We also thank Drs. Carl Fortelius, Laura Rontu, and Kalle Eerola from the FMI and Dr. Samantha Pullen from the Met Office for the discussions we had about the needs of numerical weather prediction during the development of our snow extent products and Ari Aaltonen for his help in the data retrieval from the FMI data archives. We acknowledge the use of imagery from the NASA Worldview application (<https://worldview.earthdata.nasa.gov>), part of the NASA Earth Observing System Data and Information System (EOSDIS).

Data availability statement. The H SAF MetOp/AVHRR snow extent product described in this study is produced in the LSA SAF processing system and is freely available online via LSA SAF website (<https://landsaf.ipma.pt/en/>), which requires registration for data access. One example file is available from the product description page without registration. The dataset used is MetOp/AVHRR global daily snow extent (H32); the subset used is January 2015–March 2019. In this study, synoptic weather station observations as archived for internal use in the FMI were used, but, in general, weather station data are publicly available both

freely and for a fee. One such free dataset is published by Unidata/UCAR (2003): Historical Unidata Internet Data Distribution (IDD) Global Observational Data, <https://doi.org/10.5065/9235-WJ24>.

REFERENCES

- Baldrige, A., S. Hook, C. Grove, and G. Rivera, 2009: The ASTER spectral library version 2.0. *Remote Sens. Environ.*, **113**, 711–715, <https://doi.org/10.1016/j.rse.2008.11.007>.
- Bishop, C. M., 2006: *Pattern Recognition and Machine Learning*. Springer, 738 pp.
- Brubaker, K. L., R. T. Pinker, and E. Deviatova, 2005: Evaluation and comparison of MODIS and IMS snow-cover estimates for the continental United States using station data. *J. Hydrometeor.*, **6**, 1002–1017, <https://doi.org/10.1175/JHM447.1>.
- Chen, C., T. Lakhankar, P. Romanov, S. Helfrich, A. Powell, and R. Khanbilvardi, 2012: Validation of NOAA-Interactive Multisensor Snow and Ice Mapping System (IMS) by comparison with ground-based measurements over continental United States. *Remote Sens.*, **4**, 1134–1145, <https://doi.org/10.3390/rs4051134>.
- Cortés, G., M. Girotto, and S. A. Margulis, 2014: Analysis of sub-pixel snow and ice extent over the extratropical Andes using spectral unmixing of historical Landsat imagery. *Remote Sens. Environ.*, **141**, 64–78, <https://doi.org/10.1016/j.rse.2013.10.023>.
- de Rosnay, P., G. Balsamo, C. Albergel, J. Muños-Sabater, and L. Isaksen, 2014: Initialisation of land surface variables for numerical weather prediction. *Surv. Geophys.*, **35**, 607–621, <https://doi.org/10.1007/s10712-012-9207-x>.
- Dozier, J., R. O. Green, A. W. Nolin, and T. H. Painter, 2009: Interpretation of snow properties from imaging spectrometry. *Remote Sens. Environ.*, **113**, S25–S37, <https://doi.org/10.1016/j.rse.2007.07.029>.
- Dybbroe, A., K. Karlsson, and A. Thoss, 2005: NWCSAF AVHRR cloud detection and analysis using dynamic thresholds and radiative transfer modeling. Part I: Algorithm description. *J. Appl. Meteor.*, **44**, 39–54, <https://doi.org/10.1175/JAM-2188.1>.
- Frei, A., M. Tedesco, S. Lee, J. Foster, D. K. Hall, R. Kelly, and D. A. Robinson, 2012: A review of global satellite-derived snow products. *Adv. Space Res.*, **50**, 1007–1029, <https://doi.org/10.1016/j.asr.2011.12.021>.
- Hall, D. K., and G. A. Riggs, 2007: Accuracy assessment of the MODIS snow products. *Hydrol. Processes*, **21**, 1534–1547, <https://doi.org/10.1002/hyp.6715>.
- Härer, S., M. Bernhardt, M. Siebers, and K. Schulz, 2018: On the need for a time- and location-dependent estimation of the NDSI threshold value for reducing existing uncertainties in snow cover maps at different scales. *Cryosphere*, **12**, 1629–1642, <https://doi.org/10.5194/tc-12-1629-2018>.
- Helfrich, S. R., D. McNamara, B. H. Ramsay, T. Baldwin, and T. Kasheta, 2007: Enhancements to, and forthcoming developments in the Interactive Multisensor Snow and Ice Mapping System (IMS). *Hydrol. Processes*, **21**, 1576–1586, <https://doi.org/10.1002/hyp.6720>.
- Helmert, J., and Coauthors, 2018: Review of snow data assimilation methods for hydrological, land surface, meteorological and climate models: Results from a COST HarmoSnow Survey. *Geosciences*, **8**, 1–22, <https://doi.org/10.3390/geosciences8120489>.

- Hogan, R. J., and I. B. Mason, 2012: Deterministic forecast of binary data. *Forecast Verification: A Practitioners Guide in Atmospheric Science*, 2nd ed. I. T. Jolliffe and D. B. Stephenson, Eds., John Wiley and Sons, 31–59.
- Hori, M., K. Sugiyama, K. Kobayashi, T. Aoki, T. Tanikawa, K. Kuchiki, M. Niwano, and H. Enomoto, 2017: A 38-year (1978–2015) Northern Hemisphere daily snow cover extent product derived using consistent objective criteria from satellite-borne optical sensors. *Remote Sens. Environ.*, **191**, 402–418, <https://doi.org/10.1016/j.rse.2017.01.023>.
- Horler, D. N. H., M. Dockray, and J. Barber, 1983: The red edge of plant leaf reflectance. *Int. J. Remote Sens.*, **4**, 273–288, <https://doi.org/10.1080/01431168308948546>.
- Hüsler, F., T. Jonas, S. Wunderle, and S. Albrecht, 2012: Validation of a modified snow cover retrieval algorithm from historical 1-km AVHRR data over the European Alps. *Remote Sens. Environ.*, **121**, 497–515, <https://doi.org/10.1016/j.rse.2012.02.018>.
- Hyvärinen, O., and E. Saltikoff, 2010: Social media as a source of meteorological observations. *Mon. Wea. Rev.*, **138**, 3175–3184, <https://doi.org/10.1175/2010MWR3270.1>.
- Key, J. R., and Coauthors, 2013: Snow and ice products from Suomi NPP VIIRS. *J. Geophys. Res. Atmos.*, **118**, 12 816–12 830, <https://doi.org/10.1002/2013JD020459>.
- Li, X., R. T. Pinker, M. M. Wonsick, and Y. Ma, 2007: Toward improved satellite estimates of short-wave radiative fluxes—Focus on cloud detection over snow: 1. Methodology. *J. Geophys. Res.*, **112**, D07208, <https://doi.org/10.1029/2005JD006698>.
- Loveland, T. R., and A. S. Belward, 1997: The IGBP-DIS global 1km land cover data set, DISCover: First results. *Int. J. Remote Sens.*, **18**, 3289–3295, <https://doi.org/10.1080/014311697217099>.
- Metsämäki, S., and Coauthors, 2015: Introduction to GlobSnow snow extent products with considerations for accuracy assessment. *Remote Sens. Environ.*, **156**, 96–108, <https://doi.org/10.1016/j.rse.2014.09.018>.
- Miller, S. D., T. F. Lee, and R. L. Fennimore, 2005: Satellite-based imagery techniques for daytime cloud/snow delineation from MODIS. *J. Appl. Meteor.*, **44**, 987–997, <https://doi.org/10.1175/JAM2252.1>.
- , and Coauthors, 2006: NexSat: Previewing NPOESS/VIIRS imagery capabilities. *Bull. Amer. Meteor. Soc.*, **87**, 433–446, <https://doi.org/10.1175/BAMS-87-4-433>.
- Niittynen, P., R. K. Heikkinen, and M. Luoto, 2018: Snow cover is a neglected driver of Arctic biodiversity loss. *Nat. Climate Change*, **8**, 997–1001, <https://doi.org/10.1038/s41558-018-0311-x>.
- Notarnicola, C., and Coauthors, 2013a: Snow cover maps from MODIS images at 250 m resolution, Part I: Algorithm description. *Remote Sens.*, **5**, 110–126, <https://doi.org/10.3390/rs5010110>.
- , and Coauthors, 2013b: Snow cover maps from MODIS images at 250 m resolution, Part II: Validation. *Remote Sens.*, **5**, 1568–1587, <https://doi.org/10.3390/rs5041568>.
- Parajka, J., and G. Blöschl, 2008: Spatio-temporal combination of MODIS images—Potential for snow cover mapping. *Water Resour. Res.*, **44**, W03406, <https://doi.org/10.1029/2007WR006204>.
- Piazzzi, G., and Coauthors, 2019: Cross-country assessment of H-SAF snow products by Sentinel-2 imagery validated against in-situ observations and webcam photography. *Geosciences*, **9**, 129, <https://doi.org/10.3390/geosciences9030129>.
- Pullen, S., C. Jones, and G. Rooney, 2010: Using satellite-derived snow cover data to implement a snow analysis in the Met Office global NWP model. *J. Appl. Meteor. Climatol.*, **50**, 958–973, <https://doi.org/10.1175/2010JAMC2527.1>.
- , B. Gomez, C. Charlton-Perez, R. Renshaw, and B. Candy, 2018: Land surface data assimilation at the Met Office and developments in snow depth analysis. *Eighth ISWG/LSA SAF Workshop*, Lisbon, Portugal, ISWG/LSA SAF, <https://landsaf.ipma.pt/GetDocument.do?id=709>.
- , C. Charlton-Perez, B. Gomez, and C. Harris, 2019: Snow data assimilation developments at the Met Office. *Third Int. Surface Working Group Workshop*, Montreal, QC, Canada, ISWG, http://cimss.ssec.wisc.edu/iswg/meetings/2019/presentations/WEDNESDAY/0930_Pullen.pdf.
- Ramsay, B., 1998: The interactive multisensor snow and ice mapping system. *Hydrol. Processes*, **12**, 1537–1546, [https://doi.org/10.1002/\(SICI\)1099-1085\(199808/09\)12:10<1537::AID-HYP679>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1099-1085(199808/09)12:10<1537::AID-HYP679>3.0.CO;2-A).
- Riggs, G. A., D. K. Hall, and M. O. Román, 2017: Overview of NASA's MODIS and Visible Infrared Imaging Radiometer Suite (VIIRS) snow-cover Earth System Data Records. *Earth Syst. Sci. Data*, **9**, 765–777, <https://doi.org/10.5194/essd-9-765-2017>.
- Romanov, P., D. Tarpley, G. Gutman, and T. Carroll, 2003: Mapping and monitoring of the snow cover fraction over North America. *J. Geophys. Res.*, **108**, 8619, <https://doi.org/10.1029/2002JD003142>.
- Salminen, M., J. Pulliainen, S. Metsämäki, A. Kontu, and H. Suokanerva, 2009: The behaviour of snow and snow-free surface reflectance in boreal forests: Implications to the performance of snow covered area monitoring. *Remote Sens. Environ.*, **113**, 907–918, <https://doi.org/10.1016/j.rse.2008.12.008>.
- Salvatori, R., and Coauthors, 2011: Snow cover monitoring with images from digital camera system. *Eur. J. Remote Sens.*, **43**, 137–145, <https://doi.org/10.5721/ItJRS201143211>.
- Selkowitz, D. J., and R. R. Forster, 2015: An automated approach for mapping persistent ice and snow cover over high latitude regions. *Remote Sens.*, **8**, 16, <https://doi.org/10.3390/rs8010016>.
- Siljamo, N., and O. Hyvärinen, 2011: New geostationary satellite-based snow-cover algorithm. *J. Appl. Meteor. Climatol.*, **50**, 1275–1290, <https://doi.org/10.1175/2010JAMC2568.1>.
- Solberg, R., Ø. Trier, L.-A. Breivik, Ø. Godøy, M. Killie, L. M. Andreassen, J. Hausberg, and O. Olsen, 2009: CryoClim—A new system for cryospheric climate monitoring. *Proc. 33rd Int. Symp. on Remote Sensing of Environment*, Stresa, Italy, International Center for Remote Sensing of Environment/Joint Research Centre of the European Commission, http://publications.nr.no/5200/Solberg_-_CryoClim_-_A_New_System_for_Cryospheric_Climate_Mo.pdf.
- Surer, S., J. Parajka, and Z. Akyurek, 2014: Validation of the operational MSG-SEVIRI snow cover product over Austria. *Hydrol. Earth Syst. Sci.*, **18**, 763–774, <https://doi.org/10.5194/hess-18-763-2014>.
- Thapa, S., P. K. Chhetri, and A. G. Klein, 2019: Cross-comparison between MODIS and VIIRS snow cover products for the 2016 hydrological year. *Climate*, **7**, 57, <https://doi.org/10.3390/cli7040057>.
- Thirel, G., and Coauthors, 2012: Assessing the quality of a real-time Snow Cover Area product for hydrological applications. *Remote Sens. Environ.*, **127**, 271–287, <https://doi.org/10.1016/j.rse.2012.09.006>.
- Trigo, I. F., and Coauthors, 2011: The satellite application facility for land surface analysis. *Int. J. Remote Sens.*, **32**, 2725–2744, <https://doi.org/10.1080/01431161003743199>.
- Unidata/UCAR, 2003: Historical Unidata Internet Data Distribution (IDD) Global Observational Data. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, accessed 24 September 2019, <https://doi.org/10.5065/9235-WJ24>.

- Wang, G., L. Jiang, S. Wu, J. Shi, S. Hao, and X. Liu, 2017: Fractional snow cover mapping from FY-2 VISSR imagery of China. *Remote Sens.*, **9**, 983, <https://doi.org/10.3390/rs9100983>.
- Wiscombe, W. J., and S. G. Warren, 1980a: A model for the spectral albedo of snow. I: Pure snow. *J. Atmos. Sci.*, **37**, 2712–2733, [https://doi.org/10.1175/1520-0469\(1980\)037<2712:AMFTSA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1980)037<2712:AMFTSA>2.0.CO;2).
- , and ——, 1980b: A model for the spectral albedo of snow. II: Snow containing atmospheric aerosols. *J. Atmos. Sci.*, **37**, 2734–2745, [https://doi.org/10.1175/1520-0469\(1980\)037<2734:AMFTSA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1980)037<2734:AMFTSA>2.0.CO;2).
- WMO, 2015: International codes. Vol. I.2, Manual on Codes, WMO Technical Regulations, WMO Doc. 306, 82 pp., https://www.wmo.int/pages/prog/www/WMOCodes/WMO306_vI2/Publications/2015editionUP2017/WMO306_vI2_en_ONLINE.pdf.