Report on validation of VIIRS-FSC products against in-situ observations

Issue 1.0

SEN3APP

Processing Lines And Operational Services Combining Sentinel And In-Situ Data For Terrestrial Cryosphere And Boreal Forest Zone

FP7 Grant agreement No 607052

Document History

Date	Update	Contributor	Organization
25/09/2016	Initial document	Sari Metsämäki	SYKE

Table of Contents

Introduction

1. Document Identifier

SEN3APP_SYKE_VIIRS-FSC_VR_V1.0

2. Title

Report on validation VIIRS-FSC-products against in-situ observations

3. Authority

Finnish Environment Institute

4. Abstract

This document describes the employed datasets, protocol and results of the validation of SEN3App VIIRS snow near-real-time products (featuring Fractional snow Cover, FSC) over Europe.

The in-situ validation datasets are described in general level, i.e. the data source and contents. The quality of the observational data remains not discussed, as the quality information is available from the data provider.

Validation protocol is described in more detail; however, the applied methodology for evaluation of products performance is quite well established. Therefore, the benefits/drawbacks of the methodologies are discussed only briefly.

5. Keywords

Snow, Fractional Snow Cover, optical, Satellite, VIIRS, Europe, Snow Depth, Weather stations, Validation, Binary metrics

6. Key terminology

Key terminology concerns mostly the commonly used term describing the in-situ data and the methodology to accomplish the validation and to present the results.

Fractional Snow Cover refers to the percentage of the areal unit (typically a pixel) that is covered by snow. It is typically expresses as a percentage [0-100%) or alternatively as a real number [0-1].

Snow Depth is the depth (typically expressed as centimetres) of the snowpack that is measured at an observation site.

Binary snow data describes the snow conditions simply as two alternative classes: Snow-covered or snow-free (we use expression 'snow/no-snow' here). When using binary information it is not specified how large portion of the unit area is covered by snow neither how deep the snowpack is. However the binary data is the commonest information provided by the different snow products providers internationally while Fractional snow is more rare although gives more detailed information on the snow conditions.

Binary metrics are metrics (numerical values) developed to describe the success of the binary snow product to classify the product pixels correctly when compared with in-situ observations (also binary). For example, how many true (at in-situ site) observed 'snow' pixels are classified as 'snow' in the product. There are several metrics developed and commonly used (e.g. Painter, 2009; Rittger, 2013); we use the commonest ones decribed later in 11.1.

7. Background, Context and Scope

The product was developed under EU FP7 SEN3APP project. Where an exhaustive satellite data product portfolio and development tracks for these products were presented to potential users. Development work was continued in close collaboration with selected users, but also with the core products of the participating institutes. The fractional snow cover using the SCAMod -algorithm (Metsämäki et al, 2005 and 2012) is one of SYKEs core satellite data products. Here the validation is targeted to the data produced from NPP Suomi/VIIRS satellite sensor, which acts as a backup sensor for FSC -products from Sentinel-3, in case of satellite failure.

The purpose of the document is to give the user enough information to evaluate the suitability of the satellite data product and to be able to back track the procedure how the validation results were derived.

8. Outcomes

The gained results from the validation indicate that the Sen3app VIIRS FSC-product does a good job in mapping snow coverage throughout the year, and that its performance is a bit higher over (potential) snow season from October to June. There are some False snow commissions (true 'no-snow' cases falsely interpreted as 'snow' but the ratio of false snow commissions out of all truly 'non-snow cases is very low is very low, < 0.03 on the average. The most important validation metrics – F-score (see below in Section 11) – ranges between 0.79-0.83. The VIIRS products capability to identify existing true snow is high, more than 0.86 at its best (depending on the applied methodology for handling the snow data, see Sections 11 and 12.

9. Inputs

The validation focuses on time period 01.10.2014-30.09.2016, so two snow periods and two summer periods are covered. The employed data are as follows:

1. VIIRS-based FSC-products covering Europe. Provided in collaboration of FMI and SYKE, available at FMI. Products are in netcdf-format or as geotiffs, depending on year.
2. In-situ observations made at Russian weather stations provided by RIHMI
3. ECMWF weather stations
4. In-situ observations made at ECA&D stations in Germany.
5. In-situ observations made at the Finnish Weather Stations (operated by the Finnish Meteorological Institute)

All in-situ datasets provide information on Snow Depth associated to time and location of the measurement. The station networks are shown in Figure 9.1.

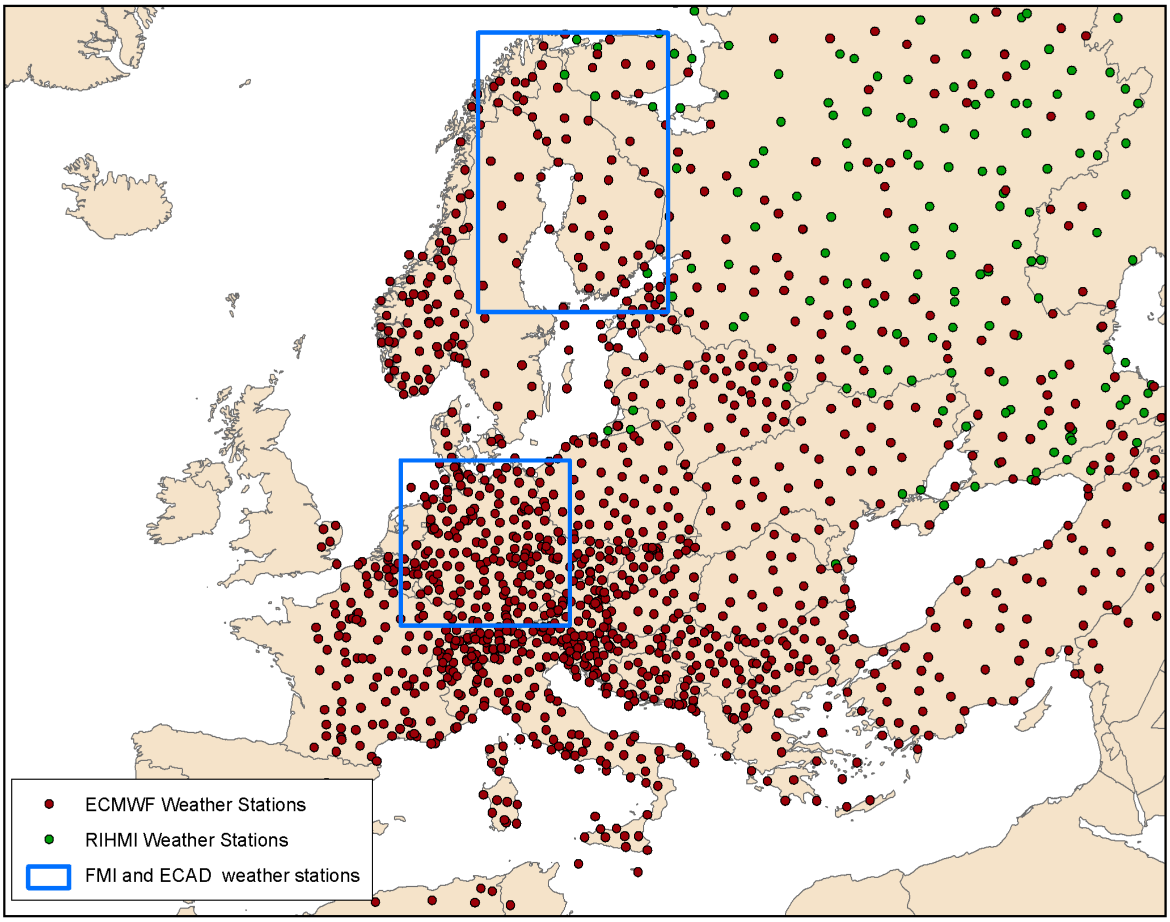


Figure 9.1. Map of locations of in-situ datasets used in the validation. The blue boxes outline the national weather station networks of Finland and Germany.

10. Standards and Traceability

The validation relies on the use of established metric for describing the success of the classified 'snow/non-snow' data. We use binary metric to describe the success of the product performance; these metrics are described e.g. in (Painter, 2009 and Rittger, 2013), where they were used for evaluating the performance of the authors' snow products. We also use Cumulative Distribution function (CDF) for describing the relationship between SnowDepth and Fractional Snow cover. This is basis mathematics and can be found in general mathematical literature.

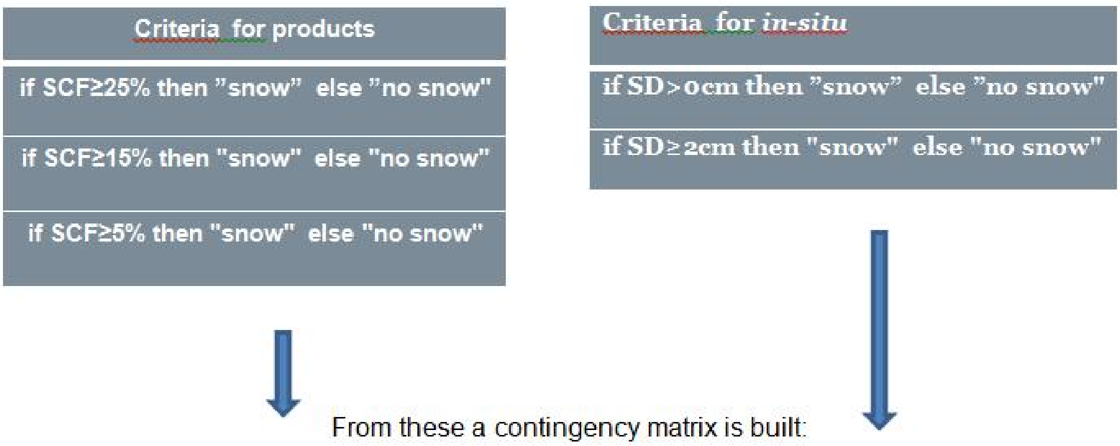
11. Methodology, Processing

Sen3app VIIRS product provide the Fractional snow Cover (FSC) but these are converted to binary data by thresholding. This is necessary as the available in-situ data do not provided information on FSC but only Snow Depth. The idea of converting Snow depth to binary information is very established, while there is no way to convert snow Depth to FSC. There is of course certain uncertainty present in SnowDepth to binary snow conversion, but this is the only feasible way at the present time.

We use three different alternative thresholds for FSC to convert the original FSC value to binary ‘snow/no-snow’ information, FSC=25%, FSC=15% and FSC=5%. For Snow Depth, two thresholds are used: 0cm and 2cm. These values are commonly used in validation studies.

The calculation of a Binary metric is based on the number of matching/non-matching classification of binary snow data: We have four classes:

1. True negatives (TN), when both product and in-situ data shows ‘no-snow’
2. True positives (TP), when both product and in-situ data shows ‘snow’
3. False negative (FN), when product shows ‘no-snow’ while in-situ indicates ‘snow’
4. False positives(FP), when products shows ‘snow’ while in-situ indicates ‘no-snow’



in-situ	Product	
	snow	No-snow
snow	True Positives (TP)	False Negatives (FN)
no-snow	False Positives (FP)	True Negatives (TN)

Figure 11.1 Scheme for creating a contingency matrix.

From these, a variety of metrics can be calculated, each describing the product’s performance from a different perspective and thus gives an good overview of the performance in general. For instance, a product may be good in detecting the true snow cases, but on the other hand it may classify as snow also true non-snow cases. These characteristics are well described by binary metrics, as discussed below. We use the following metrics in Sen3app VIIRS-validation (see also Table 11.1):

Recall describes the products capability to identify true snow cases:
Recall = Number of identified true snow cases / all true snow cases

False Alarm Rate (FAR) describes how large portion of true snow-free cases are falsely classified as snow: FAR = snow free cases classified as snow / total number of snow free cases

Precision describes how large portion of the classified snow cases really are snow:
Precision = correctly classified snow cases / all cases classified as snow

Recall as is a metric describing the product's ability to identify true (according to the in-situ reference) snow cases. Accordingly, provision the *False Alarm* Rate would be unreasonable if there were very few in-situ observations on 'no-snow' compared to the total number of comparison pairs: even a few false positives would increase the False Alarm Rate close to one (100%), which necessarily would not reflect the product's performance is general. The *Hit Rate* – a very commonly used metric - gives the proportion of matching cases to all cases. It is very sensitive to the possible (and here existing) imbalance between number of in-situ 'snow' and 'no-snow' cases. Instead, *F-score* is considered more representative and is given a higher importance when interpreting the results. As Hit Rate, F-score is a measure of accuracy, but it considers both the Recall and *Precision*.

Table 11.1 The applied binary metric for describing the snow product's performance

Metric	Description	Special considerations for Nref_snow = TP+FN; Nref_nosnow = TN+FP; Ntot = TP+TN+FN+FP;
Recall	TP / (TP+FN)	IF Nref_snow<20 THEN Recall = 'not-defined'
Precision	TP / (TP+FP)	IF Nref_snow<20 or Nref_nosnow<20 THEN Precision= 'not-defined'
False Alarm Rate	FP / (FP+TN)	IF Nref_nosnow / Ntot <0.10 THEN False Alarm Rate = 'not-defined'
F-score	2*TP / (2*TP + FP + FN)	IF Nref_snow<20 or Nref_nosnow<20 THEN F-score = 'not-defined'

12. Evaluation of Performance

One aspect of the success of the FSC-estimation and therefore of the binary classification depends on the success of the applied cloud screening. The international snow monitoring community recognizes the problem of snow/cloud discrimination. Tis implies that non-recognizes cloud may easily be interpreted as 'snow' while some products falsely classifies particularly partial snow cover as 'cloud'. Keeping this in mind, it is worth reasonable to made the analyses separately for whole years (all months) and only the other only for snow season. This is because it is particularly summertime where non-detected clouds are easily interpreted a snow (some product providers really on statistics and do not permit 'snow' in summer but SEN3app VIIRS product relies only on satellite data and does not use any outside decision rules.

12.1. Performance throughout the year

Since we are originally dealing with Fractional snow information (FSC %-units), it is interesting to see how the FSC is distributed at the in-situ observed Snow Depths. this would provide quantitative information on the success of the FSC-estimates, since no classification by thresholding would not be applied. This approach dos not give the binary metrics easily expressed as numerical values but would give some indication of the product performance without interference by the effect of thresholding method (Converting or FSC and SnowDepth to binary 'snow/no-data' is always a bit debatable) Figure 12.1 provides a Cumulative Distribution Function (CDF) for different Snow Depths (SD). It can be seen from the figure that for very thin snow layer (<1 cm) the probability for the FSC<20% is ~0.97 so very high. So thins snow means low snow fraction. The fact that FSC can be also higher than 20% (see that probability of FSC<40 is ~0.99) can well be true. This typically happens when there is a fresh fallen new snow after a snow clearance – The ground may be covered with snow but it's only a thin layer. On the other side; with thick now layer 50-100 cm, the probability for the FSC<80% is very low - ~0.25. This means that the deeper the snowpack, the more probably VIIRS-based FSC is high, which naturally should be the case. Also for the other snow depths, the result is reasonable and indicates a good performance of the VIIRS snow product.

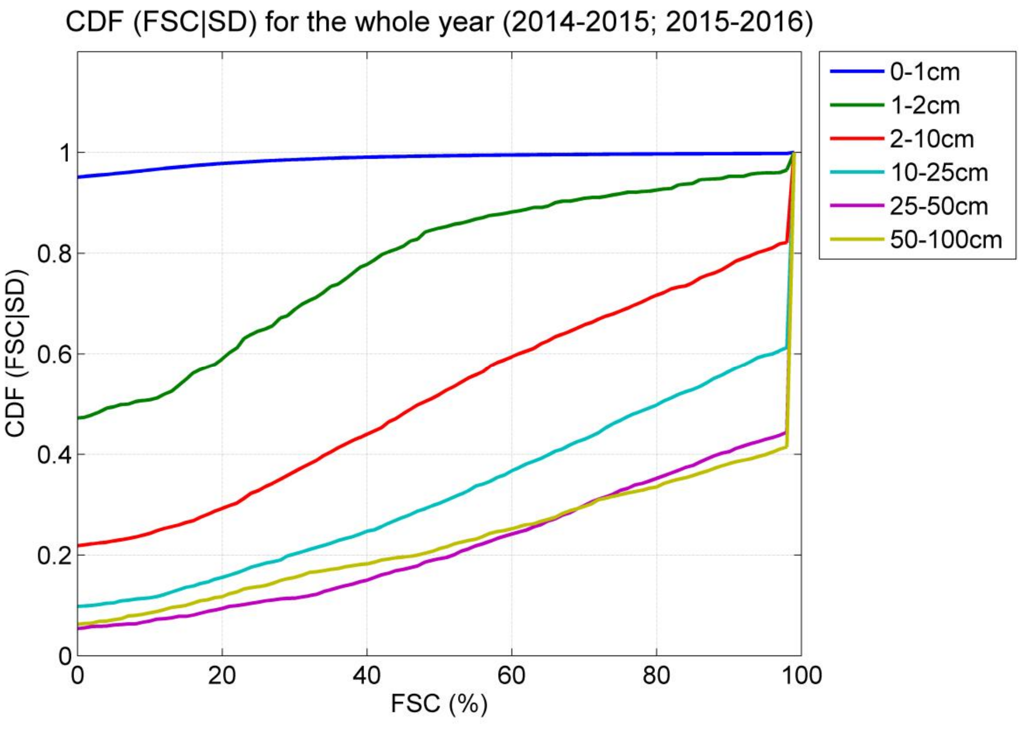


Figure 12.1. Cumulative Distribution function on FSC for different Snow Depths.

Table 12.1 Contingency matrices for the whole year, 2014-2015; 2015-2016. Number of cases for in-situ classes and snow-product classes is presented.

FSC-threshold 25%, SD-threshold; 0 cm			FSC-threshold 25%, SD-threshold; 2 cm		
	Prod. snow	Prod. no-snow		Prod. snow	Prod. no-snow
ref.snow	17590	4937	ref.snow	17101	4071
ref.nosnow	3028	162524	ref.nosnow	3517	163390
FSC-threshold 15%, SD-threshold; 0 cm			FSC-threshold 15%, SD-threshold; 2 cm		
	Prod. snow	Prod. no-snow		Prod. snow	Prod. no-snow
ref.snow	18635	3892	ref.snow	18008	3164
ref.nosnow	4780	160772	ref.nosnow	5407	161500
FSC-threshold 5%, SD-threshold; 0 cm					
	Prod. snow	Prod. no-snow			
ref.snow	19285	3242			
ref.nosnow	7241	158311			

From the contingency tables, we provide the binary metrics described in Table 11.1. For easy reading, the metrics are presented as bars.

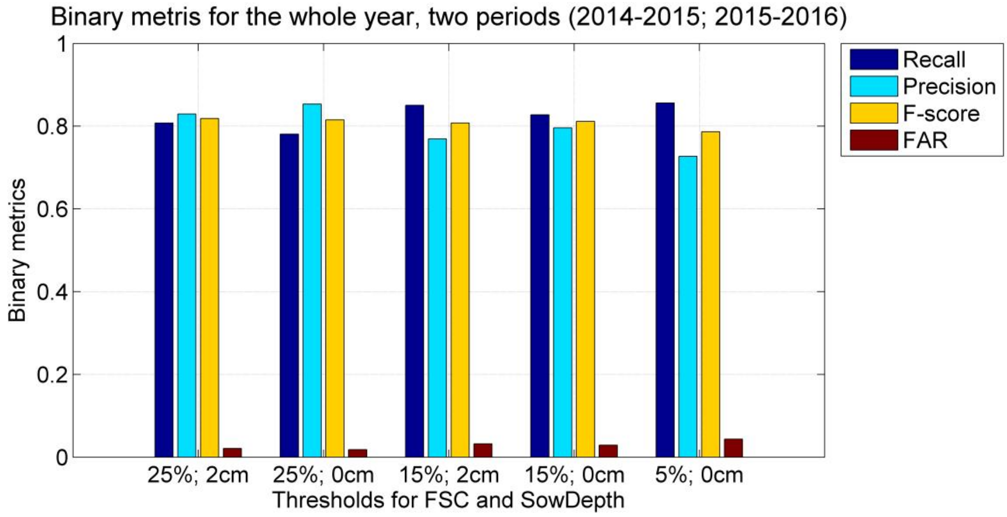


Figure 12.2. Binary metrics for the whole year, separated by the thresholds for classifying FSC and SnowDepth observations to binary 'snow/no-snow' information

The binary metrics in Fig 12.2 indidates that with all thresholds, *F-score* is at least 0.78. This particular value is associated to FSC threshold of 5% and SnowDepth threshold of 0 cm. In this case the *Recall* (product's capability to identify true snow cases) is the highest (~0.87) which is expected as already low FSC is classified as 'snow'. At the same time, however, some false snow identifications (false alarms) occur, which results in an decreased F-score.

Generally, the classification based on FSC thresholds of 25% and SnowDepth threshold of 2cm seems to provide the best validation results in terms of the applied binary metrics. In this case the *F-score*, which is the metrics most usable in the assessment (if only one metrics has to be chosen) of the products' performance is as high as 0.82. *FAR is very low, < 0.05 for all the thresholds.* We will discuss these issues a bit more in section 12.2 below Fig 12.4. (results for snow season).

12.2. Performance over snow season (October-June)

First we present the Cumulative Distribution Function (see section 12.1). In principal the curves are very similar compare to the ones for the whole years and thus indicate the same success of the VIIRS-based snow products. The only (by glance) difference shows up in the curve for SD = 0-1 cm (blue curve), which is natural as the snow fall over bare ground occur mostly in snow season; therefore it is shown up more clearly in this figure (see Section 12.1) for further explanation.

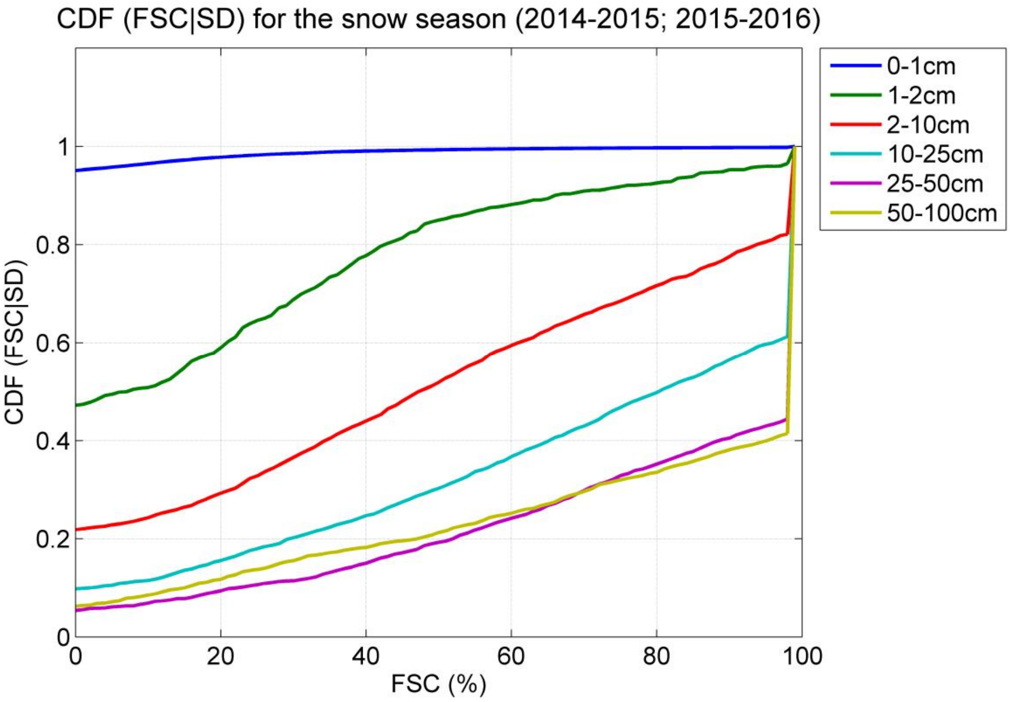


Figure 12.3. Cumulative Distribution function on FSC for different Snow Depths for snow season (October-June).

Table 12.2 presents the obtained contingency matrices for the snow season. Again from these, binary metric for the final evaluation of the product's performance is calculated. The metrics are presented in Fig.

Table 12.2 Contingency matrices for the whole year, 2014-2015; 2015-2016. Number of cases for in-situ classes and snow-product classes is presented.

FSC-threshold 25%, SD-threshold; 0 cm			FSC-threshold 25%, SD-threshold; 2 cm		
	Prod. snow	Prod. no-snow		Prod. snow	Prod. no-snow
ref.snow	17558	4790	ref.snow	17069	3954
ref.nosnow	2749	109776	ref.nosnow	3238	110612
FSC-threshold 15%, SD-threshold; 0 cm			**FSC-threshold 15%, SD-threshold; 2 cm**		
	Prod. snow	Prod. no-snow		Prod. snow	Prod. no-snow
ref.snow	18601	3747	ref.snow	17974	3049
ref.nosnow	4327	108198	ref.nosnow	4954	108896

FSC-threshold 5%, SD-threshold; 0 cm			
	Prod. snow	Prod. no-snow	
ref.snow	19249	3099	
ref.nosnow	6515	106010	

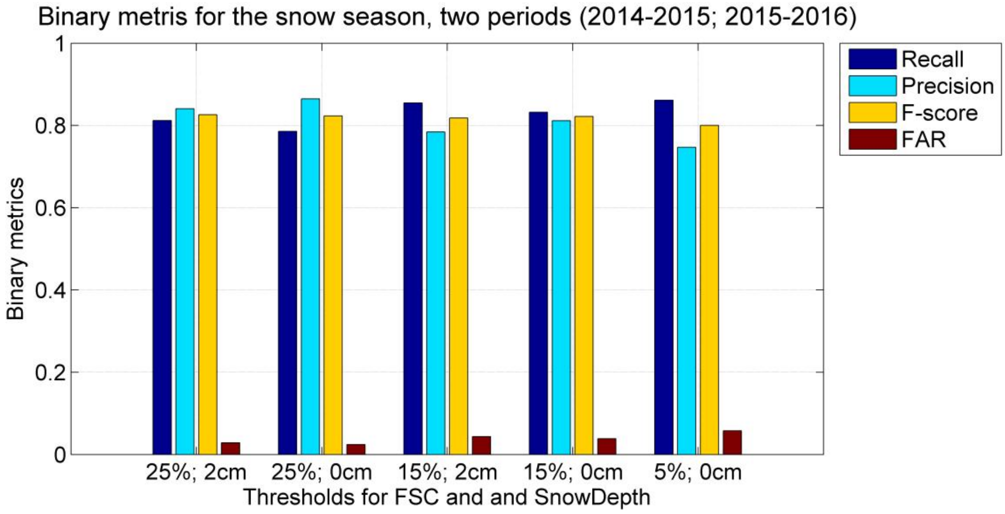


Figure 12.4. Binary metrics for snow-season, separated by the thresholds for classifying FSC and SnowDepth observations to binary 'snow/no-snow' information.

The binary metrics show a bit better performance for snow season which was expected as the summer period often confuses the snow/cloud discrimination. The FAR is a bit higher than in the results for the whole year (see Fig 12.2) , which is due to the very large number of true 'no-snow' observations at summertime – which are mostly correctly classified in the VIIRS and thus keep the FAR low compared to results in Fig 12.4 where these cases are not present.

An interesting results for this analysis is the improvement of *Precision* for all threshold-combinations. This indicates that during (potential) snow seasons the VIIRS-provided 'snow' really is snow clearly more often than in the cases when also summertime products are included in the analyses. In general the threshold FSC=25% and SnowDepth=2cm gives the best results in terms of binary metrics, hence similar results is gained as with the analysis for the whole year. *The gained F-score is as high as 0.82, and Precision even a bit higher. FAR* is 0.02. This is not surprising as the 15% FSC is sometimes considered the lowest value with an adequate accuracy; estimated values <15% are considered possibly just observational noise (caused by the sensor characteristics, atmosphere etc; Rittger et al. 2013). And, although FSC-threshold of 15% gives the highest *Recall*, the *FAR* for both SnowDepth thresholds (0.04) decreases the *F-score*. The threshold 2cm for in-situ SnowDepth probably works best as it is associated to a proper snow pack ,while for instance if SD>0.05cm is observed at the weather stations, the amount of snow in the area of the VIIRS product pixel may be so low that it is simply impossible to detect if by the means of remote sensing. Therefore 2cm threshold is more realistic and feasible.

13. Evidence to Support Performance Indicator

There are many previous projects and studies where the similar kind (or close) of validation technics has been used. One of the most recent is the European Space agency's SnowPEx-project (ending in 2016), where VIIRS-based product were not involved but the nearly the same validation technique as applied here were applied to several international snow products by different providers. The approach used here in Sen3app was also applied in the EU pre-Copernicus project Cryoland and in ESA-funded project SnowPEx.

14. References

Metsämäki, S., Anttila, S., Huttunen, M., Vepsäläinen, J. (2005). A feasible method for fractional snow cover mapping in boreal zone based on a reflectance model. Remote Sensing of Environment, Vol. 95 (1):77-95.

Metsämäki, S., Mattila, O.-P., Pulliainen, J., Niemi, K., Luojus, K., Böttcher, K. (2012). An optical reflectance model-based method for fractional snow cover mapping applicable to continental scale. Remote Sensing of Environment, 123, 508-521.

Painter, T. H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R. E., & Dozier, J. (2009). Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. *Remote Sensing of Environment*, *113*(4), 868–879. doi:10.1016/j.rse.2009.01.001

Rittger, K., Painter, T. H., & Dozier, J. (2013). Assessment of methods for mapping snow cover from MODIS. *Advances in Water Resources*, *51*, 367–380. doi:10.1016/j.advwatres.2012.03.002.

Performance assessment of the Fractional Snow Cover Products for the Pan-European area, and description of the evaluation methodology

Issue 1.0

SEN3APP

Processing Lines And Operational Services Combining Sentinel And In-Situ Data For Terrestrial Cryosphere And Boreal Forest Zone

FP7 Grant agreement No 607052

Document History

Date	Update	Contributor	Organization
22/09/2016	Initial document	Gabriele Schwaizer	ENVEO

Table of Contents

Introduction

Daily fractional snow cover over the Pan-European area is mapped fully operational by means of medium resolution optical satellite data from Terra/MODIS, as Sentinel-3 data were not available in near-real time during the SEN3APP demonstration phase.

The snow maps are of high interest for meteorological or hydrological services, but also for water suppliers or hydropower companies and geotechnical engineering companies. Users asked for daily accuracy information per pixel for this snow product for assessing the provided snow information correctly for their application. Thus, the daily Root Mean Square Error is provided per pixel as uncertainty information for this product. The purpose of this document is to assess the general performance of the product compared with snow maps generated from higher resolution optical satellite data and with snow information from in-situ data. Internationally accepted standards and protocols defined in the ESA project SnowPEx are used to evaluate the Pan-European FSC product with available reference data.

1. Document Identifier

SEN3APP_ENVEO_FSCPanEU_VR_V1.0

2. Title

Performance assessment of the Fractional Snow Cover Products for the Pan-European area, and description of the evaluation methodology

3. Authority and Contact Information

ENVEO IT GmbH

Contact information:

Gabriele Schwaizer, gabriele.schwaizer@enveo.at, +43-(0)512-507 48302

4. Abstract

The evaluation results for the fractional snow cover products for the Pan-European area are presented. The performance of the products is assessed by comparison with snow maps from selected Landsat scenes and selected in-situ data.

5. Keywords

#fractional snow cover, #Pan-European area, #optical satellite, #MODIS

6. Key terminology

Fractional snow cover	The area of snow in percentage (0 % - 100 %) referred to the area of one pixel (100 %)
Pan-European area	Defined area for the FSC product extending from 72°N/11°W to 35°N/50°E
RMSE	Root Mean Square Error

7. Background, Context and Scope

The Fractional Snow Cover (FSC) service for the Pan-European area from existing optical satellite data has been developed during the EU FP7 project CryoLand (No. 262925), and continued running fully operationally during the EU FP7 project SEN3APP (No. 607052). Key users were interviewed in the beginning of the project period to identify if the already existing service still fulfils their requirements or if any changes are needed.

Preliminary validation activities were already performed during the EU FP7 project CryoLand. The existing reference data base of 44 very high resolution optical satellite data from SPOT-5, WorldView-1/-2, etc., and 59 selected Landsat scenes, acquired between 2000 and 2014, has been further extended by 72 additional Landsat scenes acquired between 2000 and 2015, spatially distributed all over the Pan-European area. As also the reference snow maps from high resolution satellite data have some uncertainty, which has not yet been sufficiently investigated, we generate snow maps from Landsat data by applying 3 different snow detection algorithms. The Pan-European FSC products are then intercompared with all these reference snow maps, to assess the quality of the products. The intercomparison protocols and methods developed and established internationally during the ESA project SnowPEx (lead by ENVEO) are used to update the evaluation results for the snow extent products.

The QA4EO framework was selected as an example of a template for reporting the performance of the Pan-European Fractional Snow Cover products and as an information package about the dataset for the user to easily assess the suitability of the data for the purpose.

8. Product performance and uncertainty

The mean unbiased RMSE value derived from the intercomparisons with the selected Landsat scenes is about 16 %, and mean Bias values range between about -3.5 % and + 1 %, dependent on the algorithm applied on the Landsat scenes for generating the reference snow maps. But, depending on the surface characteristics covered by particular Landsat scenes, the unbiased RMSE and Bias values can be much higher or lower. The derived unbiased RMSE values for the total area covered by selected Landsat scenes range between about 0.5 % and about 38.5 %, and the Bias values range between minimum about -33.5 % and maximum about +18.5 %.

Table 8.1: Results of the evaluation of the Pan-European FSC product with snow maps from Landsat data generated by different algorithms, for different snow and surface classes, as well as for the total area of all Landsat scenes. The number (#) of samples is the number of intercomparison pairs including the particular snow or surface class.

ENVEO Pan-European Fractional Snow Cover						
	Dozier		*Klein*		*Salomonson*	
Measure	*Unbiased RMSE*	*Bias*	*Unbiased RMSE*	*Bias*	*Unbiased RMSE*	*Bias*
Mean	15,85	-0,37	16,16	-3,44	15,68	0,63
Minimum	0,45	-17,1	1,7	-33,28	2,38	-19,61
Maximum	30,37	12,48	38,34	8,14	30,93	18,33

The main validation results derived for the intercomparison of the Pan-European FSC products with snow maps generated from Landsat data by applying different algorithms are summarized in Table 8.1.

9. Inputs

Input	Description	Link
ENVEO Pan-European Fractional Snow Cover Products	Daily fractional snow cover maps for the Pan-European area	http://neso1.cryoland.enveo.at/cryoclient/
Snow maps from Landsat data	Snow maps generated by applying different snow classification algorithms on selected Landsat scenes (selected within ESA project SnowPEx), spatially distributed all over Europe	http://snowpex.enveo.at/LS_data_processing.html
Corine Land Cover 2012 (CLC 2012)	Surface classification by the European Environment Agency (EEA), V16	http://land.copernicus.eu/pan-european/corine-land-cover/clc-2012/view
GlobCover 2009	ESA surface classification (water, forest), for regions not covered by the CLC2012	http://due.esrin.esa.int/page_globcover.php
EU DEM	Digital Elevation Model provided by the European Environment Agency, used for topographic correction, and generation of mountain mask	http://www.eea.europa.eu/data-and-maps/data/eu-dem
Scripts for snow map comparison and statistical analysis	Software developed at ENVEO for a pixel-by-pixel intercomparison of the Pan-European FSC product and the snow maps from Landsat data, and for the generation of statistical analyses	

10. Standards and Traceability

Standard/ Documentation	Description	Link
Topographic correction	Correction of illumination effects and atmospheric propagation due to topography	Ekstrand (1996)
NDSI	Normalized Difference Snow Index (used for snow pre-classification of the Pan-European FSC product)	Hall, Riggs, Salomonson, DiGiromamo, & Bayr, 2002

Standard/ Documentation	Description	Link
SCAmod	Algorithm used for the generation of the Pan-European FSC product	Metsämäki et al., 2012
Dozier	Snow detection algorithm applied on Landsat scenes	Dozier & Painter, 2004
Klein	Snow detection algorithm applied on Landsat scenes	Klein, Hall, & Riggs, 1998
Salomonson	Snow detection algorithm applied on Landsat scenes	Salomonson & Appel, 2004, 2006

11. Methodology, Processing

1) Select and download snow covered Landsat scenes at nearly clear sky conditions to be used for generating reference snow maps, and associated auxiliary data (DEM, surface classification, water mask, etc.)
2) Pre-process all needed reference data sets, including radiometric calibration of Landsat data, reprojection and resampling of auxiliary data as needed, topographic correction (Ekstrand, 1996) of Landsat top of atmosphere reflectance, generation of reference snow maps by applying different snow detection methods (Dozier, Klein, Salomonson)
3) Resample, reproject and aggregate high resolution reference snow maps to fractional snow map at the grid size of the Pan-European FSC product
4) Run pixel-by-pixel intercomparison between reference snow map and the Pan-European FSC product and calculate statistics

12. Evaluation of Performance

A validation data base of 59 Landsat data, 44 very high resolution optical satellite data and in-situ data was used for the product evaluation during the EU FP7 project CryoLand.

The following statistical measures are used to describe the product performance:

- the Bias between two products, the number of used pixels, specified by N_{ui}, are used as calculation basis:

$$BIAS = \frac{1}{N_{ui}} \sum_{j=0}^{y} \sum_{i=0}^{x} (FSC_{EXT}(i,j) - FSC_{REF}(i,j))$$

- the root-mean-square error, RMSE, between two products, using all pixels suitable for inter-comparison (N_{ui}):

$$RMSE = \sqrt{\frac{1}{N_{ui}} \sum_{j=0}^{y} \sum_{i=0}^{x} (FSC_{EXT}(i,j) - FSC_{REF}(i,j))^2}$$

- the unbiased RMSE using the same input dataset (N_{ui}) as for the RMSE:

$$unbiasedRMSE = \sqrt{\frac{1}{N_{ui}} \sum_{j=0}^{y} \sum_{i=0}^{x} \left(\left(FSC_{EXT}(i,j) - \overline{FSC_{EXT}}\right) - \left(FSC_{REF}(i,j) - \overline{FSC_{REF}}\right)\right)^2}$$

- the correlation coefficient between two products (EXT = SCF Extent in Product 1, REF = Product 2 or Reference snow map, e.g. from Landsat) using only the valid pixels for the inter-comparison (N_{ui}):

$$CorrCoef = \frac{\sum_{j=0}^{y} \sum_{i=0}^{x} \left(FSC_{EXT}(i,j) - \overline{FSC_{EXT}}\right)\left(FSC_{REF}(i,j) - \overline{FSC_{REF}}\right)}{\sqrt{\sum_{j=0}^{y} \sum_{i=0}^{x} \left(FSC_{EXT}(i,j) - \overline{FSC_{EXT}}\right)^2 \sum_{j=0}^{y} \sum_{i=0}^{x} \left(FSC_{REF}(i,j) - \overline{FSC_{REF}}\right)^2}}$$

$\overline{FSC}$ is the average fractional snow cover value.

The results of the Pan-European FSC evaluation performed during the EU FP7 project CryoLand are summarized in Table 12.1. These intercomparisons of total 59 Landsat scenes and 44 VHR scenes resulted in a mean unbiased RMSE value of about 16 % and a mean positive Bias ranging between about -2 % and +3 %, and the mean correlation coefficients were 0.74 and 0.75.

Table 12.1: Results of the Pan-European FSC product evaluation with snow maps from high resolution optical satellite data performed during the EU FP7 project CryoLand.

	Landsat scenes			*VHR scenes*
Algorithm / *Measure*	*Dozier*	*Klein*	*Salomonson*	*Manual mapping*
Number of scenes	59	59	59	44
unbiased RMSE	15,93	15,86	15,76	15,36
Bias	-0,70	2,78	-1,89	3,02
Correlation Coefficient	0,75	0,74	0,75	0,74

56 of the Landsat scenes used already during the CryoLand validation activities, and additional 72 newly selected Landsat scenes were now used to update the evaluation of the Pan-European Fractional Snow Cover product for the service operation during SEN3APP.

The intercomparison of the Pan-European FSC product with the snow maps generated by applying different algorithms, Dozier, Klein and Salomonson, on the 128 Landsat scenes results in a mean unbiased RMSE value of about 16 % for each Landsat snow map classification, but the mean Bias values range between about -3.5 % for the Klein algorithm applied on Landsat and about +0.6% for the Salomonson algorithms applied on Landsat. The detailed unbiased RMSE and Bias values derived for each of the intercomparison cases are shown per month in Figure 12.1.

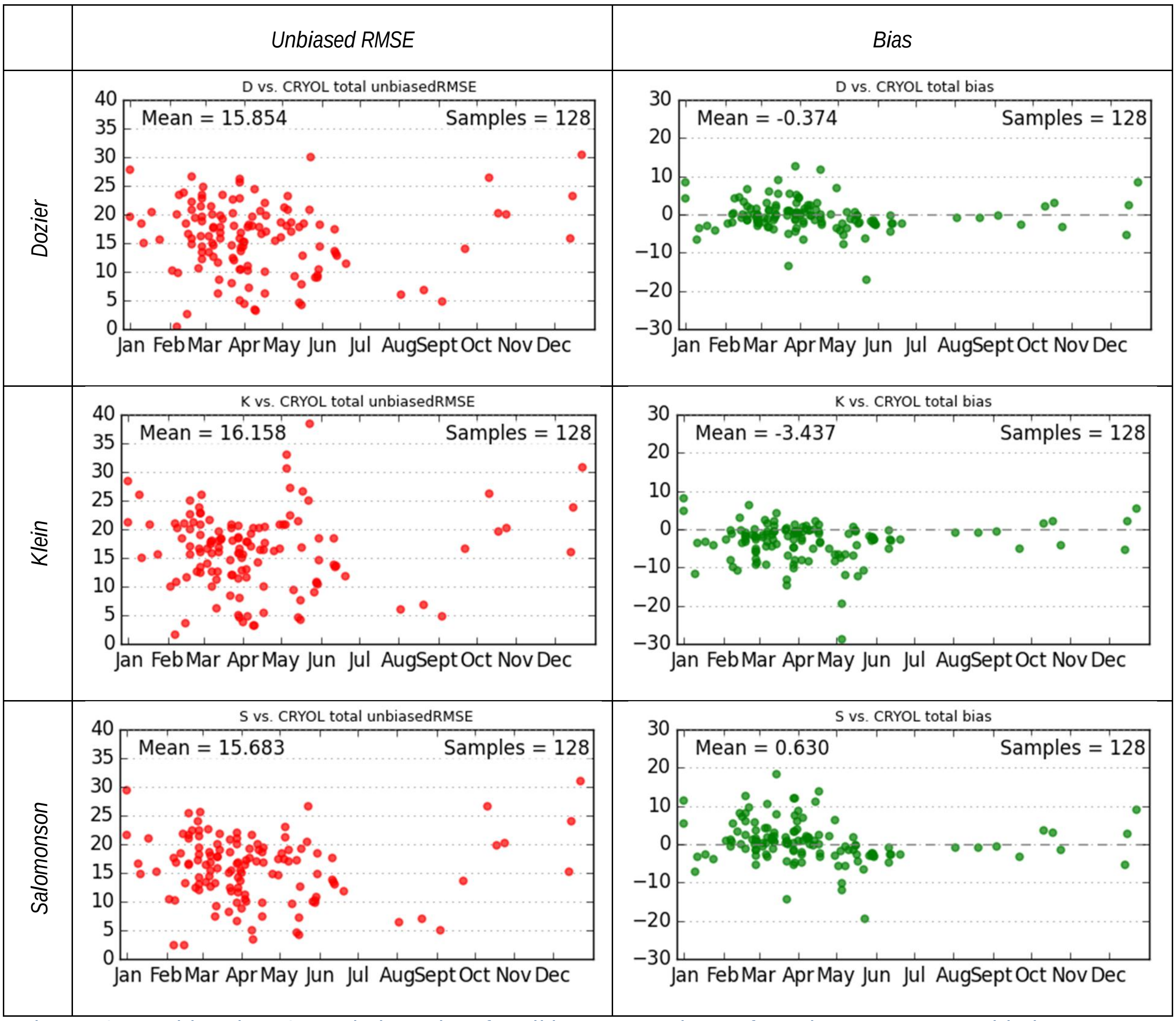


Figure 12.1: Unbiased RMSE and Bias values for all intercomparisons of Landsat snow maps with the Pan-European FSC product developed during the EU FP7 project CryoLand, and run as operational service during SEN3APP.

The mean correlation coefficients between all the different Landsat snow maps and the Pan-European FSC product are all very similar, ranging between 0.73 and 0.75, but showing each also a wide variability from very low to very high values (Figure 12.2).

These results of the extended reference data base are also in line with the evaluation results derived during the project CryoLand, when also snow maps analysed from very high resolution optical satellite data were intercompared with the product (cf. Table 12.1).

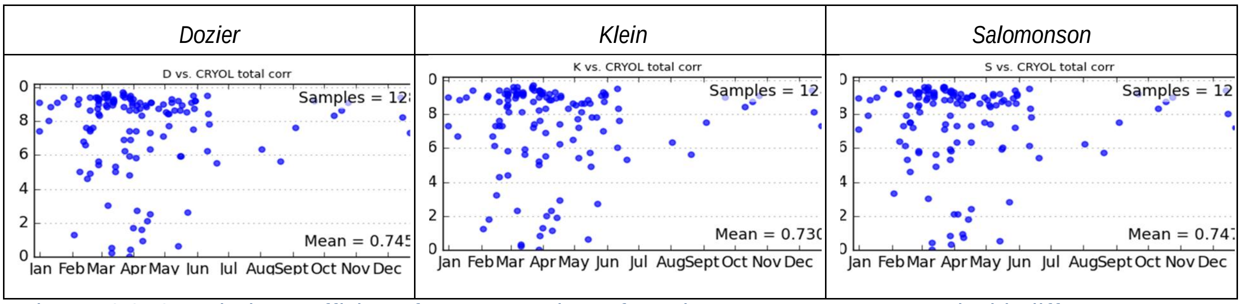


Figure 12.2: Correlation coefficients from comparison of Landsat snow maps generated with different algorithms and the CryoLand FSC product, for the total area covered by each of the Landsat scenes.

For a more detailed assessment of the product performance, the snow areas were classified into particular snow and surface categories. Therefore, the agreement of the Pan-European FSC product with the reference snow maps is investigated for pre-defined fractional snow classes, for areas within and outside forests, as well as for mountainous and plain areas. Also combinations of surface types were analysed, but are not shown here in order to keep the validation report concise and focused. Not all classes were available in all Landsat scenes, for instance mountain areas are only detected in 88 of the 128 selected Landsat scenes.

Table 12.2: Detailed statistical results of the intercomparison of the Pan-European FSC product with the selected Landsat scenes for particular snow and surface classes. Results are shown separately for each snow mapping algorithm applied on all the Landsat scenes.

ENVEO Pan-European Fractional Snow Cover									
	Dozier			Klein			Salomonson		
Class	# samples	Unbiased RMSE	Bias	# samples	Unbiased RMSE	Bias	# samples	Unbiased RMSE	Bias
0 – 25 % FSC	126	14,78	11,11	124	13,53	7,01	122	14,05	6,45
26 – 50 % FSC	126	30,29	9,89	126	29,49	-0,17	127	28,91	5,08
51 – 75 % FSC	127	28,42	0,95	127	29,85	-7,32	127	27,25	2,29
76 – 100% FSC	127	17,31	-9,28	127	18,99	-11,98	127	16,40	-6,34
Forested total	128	17,81	2,07	128	18,14	-3,19	128	17,38	3,86
Unforested total	128	13,85	-2,52	128	14,43	-4,09	128	13,67	-2,08
Mountains total	88	17,50	-0,43	88	18,05	-2,27	88	17,62	0,70
Plains total	128	10,32	-0,44	128	10,55	-3,29	128	10,09	0,25
Total area	128	15,85	-0,37	128	16,16	-3,44	128	15,68	0,63

13. Evidence to Support Performance Indicator

The evaluation of Pan-European or hemispheric snow products is a challenging and time-consuming task. Within the ESA project SnowPEx, the intercomparison and evaluation exercise for satellite-based snow products, protocols and standards for product intercomparisons and evaluations were elaborated in close collaboration with the international snow community. During this project, multiple global, hemispheric and Pan-European snow extent products have been intercompared and evaluated for the first time with the same standards and approaches. The Pan-European Fractional Snow Cover product provided as operational service during SEN3APP was participating in this intercomparison and evaluation exercises. Results of these intercomparison and evaluation activities of satellite-based snow products will be published soon.

14. References

Dozier, J., & Painter, T. H. (2004). Multispectral and Hyperspectral Remote Sensing of Alpine Snow Properties. Annual Review of Earth and Planetary Sciences, 32(1), 465–494. doi:10.1146/annurev.earth.32.101802.120404

Ekstrand, S. (1996). Landsat TM-Based Forest Damage Assessment: Correction for Topographic Effects. Photogrammetric Engineering and Remote Sensing, 62(2), 151–161.

Hall DK, Riggs GA, Salomonson V V., DiGiromamo N, Bayr KJ (2002) MODIS snow-cover products. Remote Sensing of Environment 83:181–194.

Klein, A. G., Hall, D. K., & Riggs, G. A. (1998). Improving snow cover mapping in forests through the use of a canopy reflectance model. Hydrological Processes, 12, 1723–1744.

Metsämäki S, Mattila O-P, Pulliainen J, Niemi K, Luojus K, Böttcher K (2012) An optical reflectance model-based method for fractional snow cover mapping applicable to continental scale. Remote Sensing of Environment 123:508–521

Salomonson, V. ., & Appel, I. (2004). Estimating fractional snow cover from MODIS using the normalized difference snow index. Remote Sensing of Environment, 89(3), 351–360. doi:10.1016/j.rse.2003.10.016

Salomonson, V. V., & Appel, I. (2006). Development of the Aqua MODIS NDSI fractional snow cover algorithm and validation results. IEEE Transactions on Geoscience and Remote Sensing, 44(7), 1747–1756. doi:10.1109/TGRS.2006.876029

Assessment of performance of the High resolution (5km) Pan-European SWE product (augmented using optical FSC data) and description of the evaluation methodology – QA4EO Documentation

Issue 1.0

SEN3APP

Processing Lines And Operational Services Combining Sentinel And In-Situ Data For Terrestrial Cryosphere And Boreal Forest Zone

FP7 Grant agreement No 607052

Document History

Date	Update	Contributor	Organization
29/09/2016	Initial version	K. Luojus	FMI

Table of Contents

Introduction

Daily Snow Water Equivalent products for Pan-European domain have been produced on a semi-operational basis in the SEN3App project. The retrieval is based on combination of satellite-based passive microwave radiometer data, with ground-based synoptic weather station observations. The retrieval applies the

The SWE processing system applies passive microwave observations and weather station observations in an assimilation scheme to produce maps of SWE estimates (in lat/lon coordinate system with an approximate resolution of 5km) for Pan-European domain, covering all land surface areas with the exception of mountainous regions. A semi-empirical snow emission model is used for interpreting the passive microwave (radiometer) observations through model inversion to the corresponding SWE estimates.

The SWE maps are of high interest for meteorological or hydrological services, for climate investigators and for water suppliers and hydropower companies. Users have asked for a daily product and a general information of the retrieval accuracy is needed.

Internationally accepted standards and protocols defined in the ESA project SnowPEx are used to evaluate the Pan-European FSC product with available reference data. The retrieval accuracy (as defined in ESA SnowPEx) is determined using ground based snow course observations as a reference and determining general statistical measures to describe the accuracy.

1. Document Identifier

SEN3APP_FMI_Super-SWE_VR_V1.0

2. Title

Assessment of performance of the High resolution (5km) Pan-European SWE product (augmented using optical FSC data) and description of the evaluation methodology.

3. Authority

Finnish Meteorological Institute (FMI)

Contact information:

Kari Luojus, kari.luojus@fmi.fi, +358 40 5058417.

4. Abstract

The evaluation results for the 5km Snow Water Equivalent product for the Pan-European domain are presented. The performance of the products is assessed by comparison with distributed snow transect data.

5. Keywords

#snow water equivalent, #SWE, #Pan-European, #passive microwave, #satellite product

6. Key terminology

Snow water equivalent	The amount of water stored in snowpack in mm.
Pan-European area	Defined area for the SWE product extending from 72°N/11°W to 35°N/50°E
RMSE	Root Mean Square Error

7. Background, Context and Scope

The Pan-European snow water equivalent (SWE) product for the SEN3App is derived from a legacy of ESA GlobSnow 25km Northern Hemisphere and EU FP7 CryoLand 10km Pan-European SWE product. In the SEN3App project, the retrieval was enhanced to retrieve SWE on the pan-European domain in 5km spatial resolution. The modifications to the retrieval are described in the SEN3App D3.2. The main driver to move into higher spatial resolution are due to requirements from end users to provide the SWE information in as high a spatial resolution as possible.

Validation activities for the earlier product versions have been carried out in the ESA GlobSnow-1/2 projects, covering the whole Northern Hemisphere. And in the EU FP7 project CryoLand for Pan-European domain. In addition to these projects, a more recent work in the ESA SnowPEx project has resulted in community agreed and adopted intercomparison protocols and methods which were considered in this work as well, among the best practices from ESA GlobSnow and EC CryoLand projects.

The key reference data available for the Pan-European domain (72°N/11°W to 35°N/50°E) , for the time frame of the SEN3App SWE demonstration (winter 2015-2016) were the snow course information from Finnish Environment Institute. Alternative, point-wise measurements would be available from other sources as well, but the protocols and best practices from the ESA SnowPEx project point out that validation of coarse resolution products need to be carried out with distributed snow data, such as snow transect and/or snow course data. Such data for the time-frame of the product were not available from elsewhere but Finnish Environment Institute.

The QA4EO framework was selected as an example of a template for reporting the performance of the Pan-European Fractional Snow Cover products and as an information package about the dataset for the user to easily assess the suitability of the data for the purpose.

8. Outcomes

A detailed inter-comparison of 5km Pan-EU SWE estimates against SYKE snow course observations is provided in Section 12. As a brief summary for the validation period of Oct. 1, 2015 to May, 31 2016 the statistical measures from 760 samples show a bias of -1.7 mm, RMSE 37.7 mm and a correlation coefficient of 0.81. This is in line with the results obtained for the EC CryoLand SWE product validation and the GlobSnow NH SWE product during the same season.

9. Inputs

Input	Description	Link
Satellite data	Passive microwave radiometer data from the DMSP F-17 satellite, SSMIS sensor	
Snow depth data	Snow depth observations from the ECMWF archive; including all available weather station data collected via the WMO GTS network	
Validation data	Snow course data from the Finnish Environment Institute	

10. Standards and Traceability

Standard/ Documentation	Description	Link
High resolution SWE-algorithm	Document describing the algorithm used to create the 5km SWE product	Takala et al. 2016
General description of performance assessment using snow course data	Document describing the method for snow accuracy analyses	Takala et al. 2011
General description of the overall SWE retrieval process	Document describing the general SWE retrieval algorithm in detail	Pulliainen 2006

11. Methodology, Processing

The product accuracy was determined using the following procedure

1) Acquire the SYKE snow course data for the assessment time frame
2) Acquire the 5km Pan-European SWE product for analyses
3) Pre-process the reference data, determine the locations of the reference data in the SWE coordinate system
4) Run a sample-wise inter-comparison between reference SWE values from the snow courses and the Pan-European 5km SWE product and calculate statistics, values from the exact same date and location are inter-compared.

12. Evaluation of Performance

The detailed retrieval accuracy is shown in Figures 1 and 2, with a detailed break-down of the results described below.

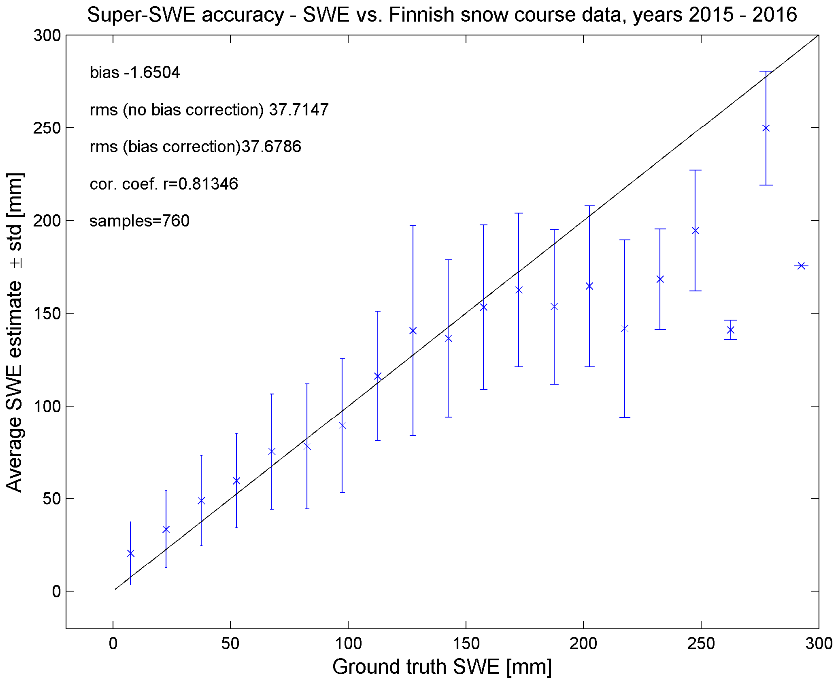


Figure 1. Validation results of estimated SWE against SYKE measured ground truth SWE for winter 2015-2016. Reference dataset consisted of 760 samples during the winter time.

The comparison with the SYKE snow course data show that the RMS-error for the 5km Pan-EU SWE (analyses consisting of 760 samples) was 37.7mm. The bias for the same dataset was -1.7mm. Both values are a significant improvement over the traditional SWE retrieval methods that rely solely on satellite microwave radiometer information, as described in (Takala et al., 2011). The results for this dataset are similar to the evaluations of the ESA GlobSnow product versions those of EC CryoLand product.

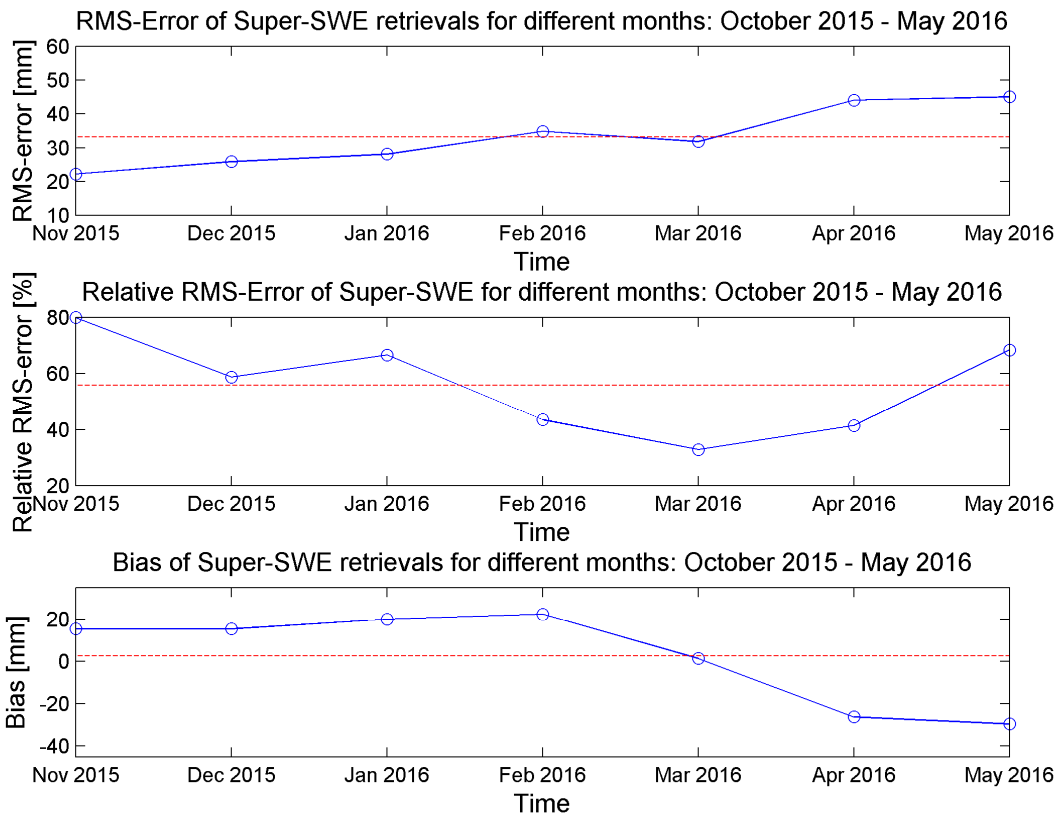


Figure 2. Validation results of estimated SWE against SYKE measured ground truth SWE for winter 2015-2016, presented separately for different months.

The RMSE and bias are presented per month during the winter season Oct 1 2015 to May 31 2016 in Figure 2. The lowest absolute RMSE values are obtained in the beginning of the snow season and the worst in the end of the season. However, the lowest relative RMSE (RMSE/monthly mean SWE) is obtained for March with the highest level of SWE, which is the most important time from the hydrological point of view (middle panel). The bias has positive values and rises steadily until February 2016 and then goes to almost zero for March and to negative for the last two winter months (April and May). Months Dec 2015 to Apr 2016 have more than 50 samples but Oct 2015, Nov 2015 and May 2016 only a few available samples.

13. Evidence to Support Performance Indicator

The retrieval performance for the legacy 25km NH SWE products are similar as acquired here for the 5km Pan-European SWE product (Takala et al. 2011)

The results acquired here are also in line with the preliminary assessment of the high resolution SWE retrieval for winter 2012-2013 presented in (Takala et al. 2016).

These facts indicate that the retrieval for the 5km spatial resolution is feasible and the processing chain for the retrieval is performing as expected.

14. References

Pulliainen, J. 2006. Mapping of snow water equivalent and snow depth in boreal and sub-arctic zones by assimilating space-borne microwave radiometer data and ground-based observations. Remote Sensing of Environment, 101, 257-269.

Takala Takala, M., Luojus, K., Pulliainen, J., Derksen, C., Lemmetyinen, J., Kärnä, J.-P., Koskinen, J. and Bojkov, B.,"Estimating northern hemisphere snow water equivalent for climate research through assimilation of space-borne radiometer data and ground-based measurements", Remote Sensing of Environment, Vol. 115 no. 12, December 2011, pp. 3517-3529, doi: 10.1016/j.rse.2011.08.014.

Takala M., Ikonen, J., Luojus, K., Lemmetyinen, J., Metsämäki, S., Cohen, J., Arslan, A. N. and Pulliainen, J., "New Snow Water Equivalent processing system with improved resolution over Europe and its applications in hydrology", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016, vol.PP, no.99, pp.1-9. doi: 10.1109/JSTARS.2016.2586179.