EUMETSAT Satellite Application Facility on
Support to Operational Hydrology and Water Management

Product Validation Report (PVR) for product H11 - SN-OBS-2

Snow Status (Dry/wet) by MW radiometry

Reference Number:	SAF/HSAF/CDOP2/PVR-11/2.0
Issue/Revision Index:	2.0
Last Change:	24 October 2013
About this document	This Document has been prepared by the Product Validation Cluster Leader, with the support of the Project Management Team and of the Validation and Development Teams of the Precipitation Cluster

DOCUMENT CHANGE RECORD		
Issue / Revision	Date	Description
1.0	16/05/2011	Baseline version prepared for ORR1 Part 2.
2.0	24/10/2013	Baseline version prepared for ORR1 part4

Index

List of figures

Acronyms

AMSU	Advanced Microwave Sounding Unit (on NOAA and MetOp)
AMSU-A	Advanced Microwave Sounding Unit - A (on NOAA and MetOp)
AMSU-B	Advanced Microwave Sounding Unit - B (on NOAA up to 17)
ATDD	Algorithms Theoretical Definition Document
AU	Anadolu University (in Turkey)
BfG	Bundesanstalt für Gewässerkunde (in Germany)
CAF	Central Application Facility (of EUMETSAT)
CDOP	Continuous Development-Operations Phase
CESBIO	Centre d'Etudes Spatiales de la BIOsphere (of CNRS, in France)
CM-SAF	SAF on Climate Monitoring
CNMCA	Centro Nazionale di Meteorologia e Climatologia Aeronautica (in Italy)
CNR	Consiglio Nazionale delle Ricerche (of Italy)
CNRS	Centre Nationale de la Recherche Scientifique (of France)
DMSP	Defense Meteorological Satellite Program
DPC	Dipartimento Protezione Civile (of Italy)
EARS	EUMETSAT Advanced Retransmission Service
ECMWF	European Centre for Medium-range Weather Forecasts
EDC	EUMETSAT Data Centre, previously known as U-MARF
EUM	Short for EUMETSAT
EUMETCast	EUMETSAT’s Broadcast System for Environmental Data
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FMI	Finnish Meteorological Institute
FTP	File Transfer Protocol
GEO	Geostationary Earth Orbit
GRAS-SAF	SAF on GRAS Meteorology
HDF	Hierarchical Data Format
HRV	High Resolution Visible (one SEVIRI channel)
H-SAF	SAF on Support to Operational Hydrology and Water Management
IDL©	Interactive Data Language

IFOV	Instantaneous Field Of View
IMWM	Institute of Meteorology and Water Management (in Poland)
IPF	Institut für Photogrammetrie und Fernerkundung (of TU-Wien, in Austria)
IPWG	International Precipitation Working Group
IR	Infra Red
IRM	Institut Royal Météorologique (of Belgium) (alternative of RMI)
ISAC	Istituto di Scienze dell'Atmosfera e del Clima (of CNR, Italy)
ITU	İstanbul Technical University (in Turkey)
LATMOS	Laboratoire Atmosphères, Milieux, Observations Spatiales (of CNRS, in France)
LEO	Low Earth Orbit
LSA-SAF	SAF on Land Surface Analysis
Météo France	National Meteorological Service of France
METU	Middle East Technical University (in Turkey)
MHS	Microwave Humidity Sounder (on NOAA 18 and 19, and on MetOp)
MSG	Meteosat Second Generation (Meteosat 8, 9, 10, 11)
MVIRI	Meteosat Visible and Infra Red Imager (on Meteosat up to 7)
MW	Micro Wave
NESDIS	National Environmental Satellite, Data and Information Services
NMA	National Meteorological Administration (of Romania)
NOAA	National Oceanic and Atmospheric Administration (Agency and satellite)
NWC-SAF	SAF in support to Nowcasting & Very Short Range Forecasting
NWP	Numerical Weather Prediction
NWP-SAF	SAF on Numerical Weather Prediction
O3M-SAF	SAF on Ozone and Atmospheric Chemistry Monitoring
OMSZ	Hungarian Meteorological Service
ORR	Operations Readiness Review
OSI-SAF	SAF on Ocean and Sea Ice
PDF	Probability Density Function
PEHRPP	Pilot Evaluation of High Resolution Precipitation Products
Pixel	Picture element
PMW	Passive Micro-Wave
PP	Project Plan
PR	Precipitation Radar (on TRMM)
PRD	Product Requirements Document
PUM	Product User Manual
PVR	Product Validation Report
RMI	Royal Meteorological Institute (of Belgium) (alternative of IRM)
RR	Rain Rate
RU	Rapid Update
SAF	Satellite Application Facility
SEVIRI	Spinning Enhanced Visible and Infra-Red Imager (on Meteosat from 8 onwards)
SHMÚ	Slovak Hydro-Meteorological Institute
SSM/I	Special Sensor Microwave / Imager (on DMSP up to F-15)
SSMIS	Special Sensor Microwave Imager/Sounder (on DMSP starting with S-16)
SYKE	Suomen ympäristökeskus (Finnish Environment Institute)
T_{BB}	Equivalent Blackbody Temperature (used for IR)
TKK	Teknillinen korkeakoulu (Helsinki University of Technology)
TMI	TRMM Microwave Imager (on TRMM)
TRMM	Tropical Rainfall Measuring Mission UKMO
TSMS	Turkish State Meteorological Service
TU-Wien	Technische Universität Wien (in Austria)
U-MARF	Unified Meteorological Archive and Retrieval Facility
UniFe	University of Ferrara (in Italy)
UTC	Universal Coordinated Time
VIS	Visible
ZAMG	Zentralanstalt für Meteorologie und Geodynamik (of Austria)

1 Validation strategy, methods and tools

1.1 Validation team and work plan

In order to evaluate the satellite snow product accuracy, a Validation Group has been established since the beginning of the Validation Phase, and enlarged in the CDOP-1 phase. The Snow Product H11 has been verified by experts from National Meteorological and Hydrological Institutes of Belgium, Germany, Finland, Poland, and Turkey .

Validation team for snow products			
Silvia Puca (Leader)	Dipartimento Protezione Civile (DPC)	Italy	silvia.puca@protezionecivile.it
Mario Barbani	Dipartimento Protezione Civile (DPC)	Italy	mario.barbani@protezionecivile.it
Emmanuel Roulin	Institut Royal Météorologique (IRM)	Belgium	Emmanuel.Roulin@oma.be
Pierre Baguis	Institut Royal Météorologique (IRM)	Belgium	Pierre.baguis@oma.be
Peter Krahe	Bundesanstalt für Gewässerkunde (BfG)	Germany	krahe@bafg.de
Panu Lahtinen	Finnish Meteorological Institute (FMI)	Finland	panu.lahtinen@fmi.fi
Bozena Lapeta	Institute of Meteorology and Water Management (IMWM)	Poland	Bozena.Lapeta@imgw.pl
Monika Pajek	Institute of Meteorology and Water Management (IMWM)	Poland	Monika.Pajek@imgw.pl
Pawel Przeniczny	Institute of Meteorology and Water Management (IMWM)	Poland	pawel.przeniczny@imgw.pl
Aydin Erturk	Turkish State Meteorological Service (TSMS)	Turkey	agerturk@dmi.gov.tr
Ibrahim Sonmez	Turkish State Meteorological Service (TSMS)	Turkey	isonmez@dmi.gov.tr
Serdar Surer	Middle East Technical University (METU)	Turkey	serdarsurer@gmail.com
Zuhal Akyurek	Middle East Technical University (METU)	Turkey	zakyurek@metu.edu.tr

Table 1 List of the people involved in the validation of H-SAF snow products

The Snow products validation programme was started soon after the H-SAF Requirements Review (26-27 April 2006). The first activity was to lay down the Validation plan, which was finalised as first draft early as 30 September 2006. After the first Workshop, other ones followed, at least one per year to exchange experiences, problem solutions and discuss possible improvement of the validation methodologies. The results of the Snow Validation Programme are reported in this Product Validation Report (PVR).

1.2 Validation objects and problems

The products validation activity has to serve multiple purposes:

- to provide input to the product developers for improving calibration for better quality of baseline products, and for guidance in the development of more advanced products;
- to characterise the product error structure in order to enable the Hydrological validation programme to appropriately use the data;
- to provide information on product error to accompany the product distribution in an open environment, after the initial phase of distribution limited to the so-called "beta users".

Validation of snow observation from space is a hard work, especially because ground systems are essentially based on in-field measurements, very sparse. Comparison with results of numerical models obviously suffer of the limited skill of NWP in predicting snow parameters (a very downstream product that passes through quantitative snow forecast, that certainly is not the most accurate product of NWP). The validation results are sensitive to the climatic situation and the status of soil.

During the Development phase some main issues have been pointed out. First of all, the need to have a wide validation with the involvement of several countries. Secondly, the opportunity to define a common validation methodology and to develop the software for all steps of this validation procedure; software available to all the members of the SPVG.

1.3 Validation methodology

Since the beginning of the project, the importance of defining a common validation procedure was clear, in order to make the results obtained by several institutes comparable and to better understand their meanings. This methodology has been identified for all the steps in CDOP-1 inside the validation group, in collaboration with the product developers, and with the support of ground data experts. The common validation methodology is based on ground data comparison to produce **large statistic** (multi-categorical) and case studies**.**

The main steps of the validation procedure are:

1. check for consistency of both observation and satellite data series,
2. comparison between the snow observation and the satellite data,
3. mountain mask application,
4. large statistic analysis: multi-categorical scores evaluation,
5. case study analysis.

1.4 Data consistency check

The first data check of the validation procedure consist in verifying the availability of both observation data and satellite product, for all days of the reference season (1.10.2012 – 31.05.2013). Secondly, a 'sanity check' is applied both on the satellite product downloaded from the ftp server and the observational data – a quick look on the filename format and file modification dates can prevent making validation on wrong (e.g. old version of the product) or incomplete datasets (e.g. missing observation data).
A study on the definition of more sophisticated check to evaluate the reliability of the snow ground measurements will be one of the next steps of the validation programme.

1.5 Mountain mask description

The mountain mask (Annex2) developed by METU for internal use of the HSAF Project depends on two main features of topography. One is the slope and the other one is the elevation of the terrain. For the mountain mask development, the GTOPO Digital Elevation Model (DEM) which has 1 km of spatial distribution has been used. The GTOPO DEM covering the H-SAF domain, which is between latitudes from 25 to 70, and longitudes from -20 to 45, has been used as base layer. A vector layer that is formed of parallel polylines with 0.05*0.05 degree intervals was developed in order to make pixel-wise calculations in each individual cell of this mesh. Using this mesh onto the GTOPO DEM, the slope and elevation values of each cell have been calculated. After obtaining this cell-wise DEM values, an algorithm has been run over this layer in order to create a binary true or false map with the rule given below:
("Mean Elevation of Cell" >=1000m) **OR** ("Std_Dev of Slope of Cell" >=2% **AND** " Mean Elevation of Cell " >=700m) **OR** ("Range of Elevation" >=800m **AND** " Mean Elevation of Cell " >=500m).

After applying this rule, a binary mountain mask, shown in Fig.1, is obtained and used in the HSAF snow products generation. It can be found on the ftp site of the Turkish State Meteorological Service (212.175.180.253) under /OUT/h1X/mountainmask directory in .hdf file format (Username: *snowtur ,* Password: *rs37kar).*

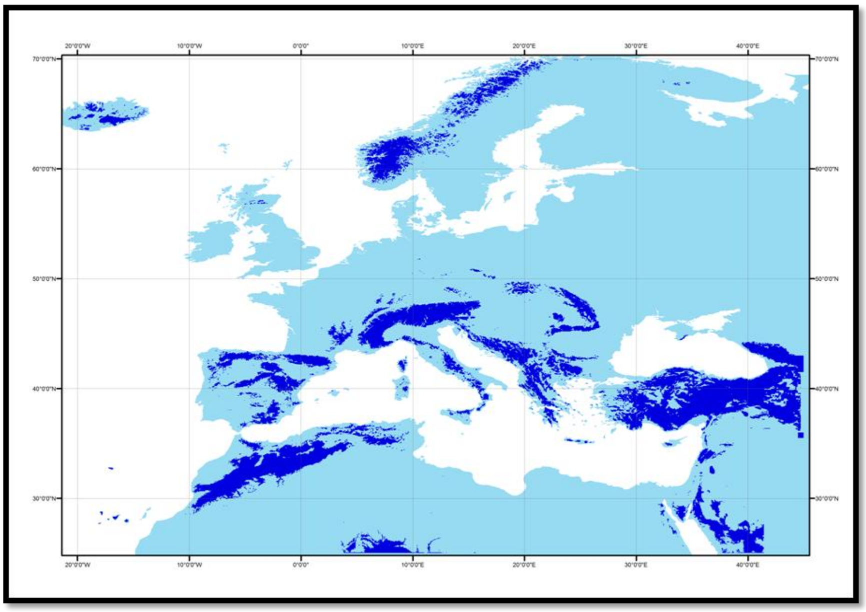

Figure 1 Mm: Mountain mask used in the HSAF Project

1.6 Comparison between the observation data and the product

The validation conditions for existence of dry snow are:

- Automatic snow depth (SD) measurement states that there is two centimetres or more of snow during the day;
- The average 2 meter air temperature between the 18 hour period between 00-18 UTC has been equal or lower than -1.5 degrees centigrade;

- The 2 meter air temperature is not above 0 degrees centigrade for more than one hour (one hourly measurement, four 15-minutely measurements, etc.) during this same period.

Or the same in more formal form

1. SD ≥ 2 cm
2. $T_{ave,air}$(00…18 UTC) ≤ -1.5 °C
3. $t_{T>0}$ ≤ 1 h.

If rule 1) is not true, there is no snow on the ground. If 1) is true, but either one of the rules 2) or 3) are false, the snow is wet. If all the rules are true, snow is considered dry.
There are some shortcomings in these validation rules. Firstly, they are indirect. This approach is based on air temperature data, which does not directly describe the status of the snow pack, especially on wetlands. Secondly, the calibration of the thresholds was made based on limited data on a single location in Finnish Lapland. The selected thresholds might not be representative in other places. Even at the same location, the overall representativeness of these rules is not certain.

The comparison between the station data and the H11 products are made daily on pixel level. For a single day, consider one pixel with several weather stations. All different stations increment corresponding snow status counters for that individual product pixel. There are counter arrays for

- the number of no snow cases;
- the number of dry snow cases;
- the number of wet snow cases.

That is, there are four arrays for every day: the three listed above, and of course the product data itself. Each *pixel* forms a *single validation data point*. To avoid the bias that would be caused by mixed-pixel cases, only the unambiguous pixels (the pixels where all the weather stations agree on the status) are included when calculating the statistical metrics. The mixed-pixel cases can be inspected as auxiliary information, although they cannot be included in the statistics defined in the Product Requirements.

1.7 Validation program package

The Finnish team has developed a program package following the methodology described in this chapter for the validation of H11. This software is available in:

ftp.fmi.fi
/HSAF/products/utilities/h11_validator/src_h11_validator.zip
user: hydrosaf
passwd: 23ywPdrM

In the same directory it is possible to find some examples. A manual related to this software is under development.

1.8 Large statistic: multi-categorical score evaluation

The large statistic validation analysis is based on the evaluation of multi-categorical statistical scores on one *snow season* of data. The *snow season* ranges between October and May. The validation of future versions of H11 will be performed on this reference period in order to assess the product improvements. The scores are monthly and seasonally evaluated.

The main steps to evaluate the statistical scores are:

- all the institute download the H11 and ground data for the reference data;
- all the institutes apply the preliminary data check;
- all the institutes apply the mountain mask to classify flat and mountain areas;
- all the institutes compare the H11 with the snow measurements (pixel to point);
- all the institutes evaluate the monthly and seasonally contingency tables and statistical scores (below reported) for the reference period on flat area, mountain area, and for merged product: flat and mountain area;
- all the institutes send the contingency tables and statistical scores to the validation cluster leader;
- the validation cluster leader verifies the consistency of the results and evaluate the monthly and yearly common statistical results.

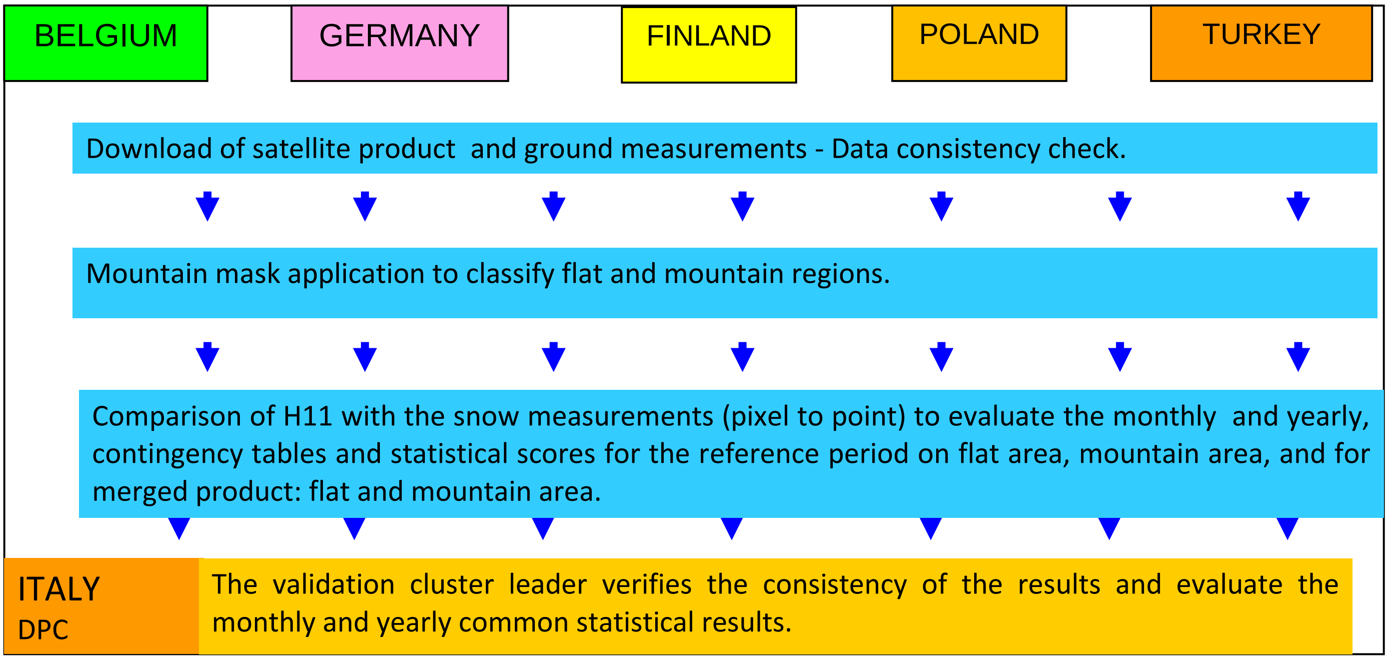

Figure 2 Main steps of the validation procedure in the SPVG

Product has value "Dark" if the sun elevation is too low, namely lower than 10 degrees.

		Station data		
		Dry	Wet	No snow
Product	Dry	Hit	False alarm	False alarm
	Wet	False alarm (check QC)	Hit	False alarm
	Dark	Miss	NA	NA

		(check QC)		
	Cloud	Miss (check QC)	NA	NA
	No snow	Miss	Miss	Correct negative

Table 2 Contingency table

Four fields are marked as not applicable (NA). In these cases it is not possible to get information on the surface status (radiometer instrument detects only dry snow, optical instruments need sunlight and clear skies). Three fields are marked with "check QC". In these cases also the Quality Control files need to be inspected to see if radiometer data has been available. If radiometer data are missing (data gaps between different satellite overpasses), these categories are discarded from the statistics, as the dry snow cannot be detected. From the classification results, different statistical scores can be calculated. To simplify the formulae, let's use the following notation

- A = number of hits
- B = number of false alarms
- C = number of misses
- D = number of correct negatives

- Probability of detection

POD = A/(A+C) Range: 0 to 1. Perfect score: 1

- False alarm ratio

FAR = B/(A+B) Range: 0 to 1. Perfect score: 0

- Probability of false detection Range: 0 to 1. Perfect score: 0

POFD = B/(B+D)

-Accuracy

ACC = (A+D)/(A+B+C+D) Range: 0 to 1. Perfect score: 1

- Critical success index

CSI = A/(A+B+C) Range: 0 to 1. Perfect score: 1

- Heidke skill score

HSS = 2(AD-BC) / [(A+C)(C+D) + (A+B)(B+D)]

Range: $-\infty$ to 1. 0 indicates no skill. Perfect score: 1.

These statistics need to be calculated on as large area as possible to find the overall performance of the product.

The results are reported separately for each country, and also a combined statistics are presented. To better see when there are most errors, the results are displayed also on a monthly basis.

The above-mentioned statistics is calculated both for flat/forested and for mountainous areas for each month of the reference season as well as for the whole season. The statistical scores are evaluated also for the merged product: flat and mountain areas.

The results are presented in the form of contingency tables and statistical scores.

2 Ground data used for validation activities

2.1 Introduction

In the following sections, the snow ground data networks used in the SPVG to validate H11 is described. The data of countries (Belgium, Finland, Germany, Poland, and Turkey) with different climatology and orography have been used in the H11 validation activity. The ground data inventory shows that synoptic, automatic, handmade snow measurements are used in the SPVG. It is well known that ground data are influenced by several error sources that should be carefully handled and characterized before using it as a reference for ground validation of any satellite-based snow products.

The classification results are first added together, i.e. hits, false alarms, misses and correct negatives of each country are combined, and the total POD/FAR/CSI are calculated from these number.
The filtering flat/mountain is done by using the mountain mask of the HSAF product.

2.2 Description of ground data used

The parameter used for validation is snow depth and 2-meter air temperature. The time period from **01.10.2012** to **31.05.2013** has been considered as reference period for the validation of H11. The satellite product has been validated with ground data both for flat and mountain regions in Belgium, Finland, Germany, Poland and Turkey.
The ground data used in the SPVG does not come from the same type of measurement stations. In the following table data coming from different instrument sources are shown. A description of the snow ground measurement national networks used for the validation of H11 is reported in the following sub-sections.

Country	Type	Flat	Mountain	Total
Belgium	Synoptic	16	0	16
Finland	Automatic	204	0	204
Germany	Synoptic	61	6	67
	automatic	941	270	1211
Poland	Synoptic	58	3	61
Turkey	Climatic	61	71	132
	Synoptic	97	35	132
	Awos	-	6	6

Table 3 Number of ground stations used for H11 validation

2.3 Ground data in Belgium (IRM)

Ground data consist of the information contained in the reports of 9 synoptic stations. These stations are a sub-set of the stations used for the validation of the snow detection i.e. stations in Southern Belgium for which the hourly temperature data needed for the common validation procedure are available. The comparison is made between the synoptic station and the pixel where the station is situated.

2.4 Ground data in Finland

Data used for validation of H11 (Snow status) in Finland is described in this section.

The parameters used for validation are snow depth and 2-meter air temperature. The data are measured at automatic synoptic (WMO) weather stations operated by FMI around Finland. The stations are equipped with Campbell Scientific SR50 Sonic Ranging Sensors. Accuracy of this sensor is 1 cm or 0.4 % (whichever is greater). Air temperatures are measured using PT100 resistive thermometers housed in WMO standard enclosures that protect the sensor from direct Sun light and rain, but allows air to flow through.

Snow depth values are from 06 UTC (official WMO snow depth report). The temperatures are collected for every hour. Alongside with these data, the station coordinates are collected from FMI database. These data are assembled to two monthly CSV files with a line for every individual report, one file for snow depths and one for temperatures.

For the validation period (Oct 1, 2009 – September 30, 2010), there were on average 65 separate weather stations reporting the snow depth (the limiting factor) in Finland. During the snow-free season, most of these stations stop reporting SD, but there were still at least 15 daily reports during the summer.

For snow season, on average 85 stations provided snow depth data. Map in the figure below shows the distribution of the stations across Finland.

For description of validation methodology, please see document "Validation methodology for H11 – snow status".

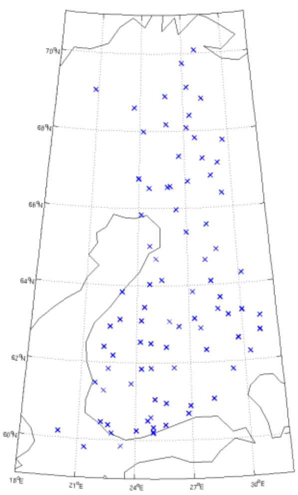

Figure 3 FMI operated weather stations used in validation activities ofH11. Due to coarse coastline data, few stations seem to be located on the Baltic sea, but they are actually on islands

2.5 Ground data in Germany

The station measurements contain daily information on snow depth and snow water equivalent (only fewer stations) from more than 2,000 stations in Germany. They are part of the station network operated by the Deutsche Wetterdienst DWD. The spatial distribution of the stations is illustrated in Figure 2. In all areas the density of stations is very high. Especially in the higher areas of the midmountains and the South of Germany, that are most affected by snow, a huge number of stations is available as ground truth.

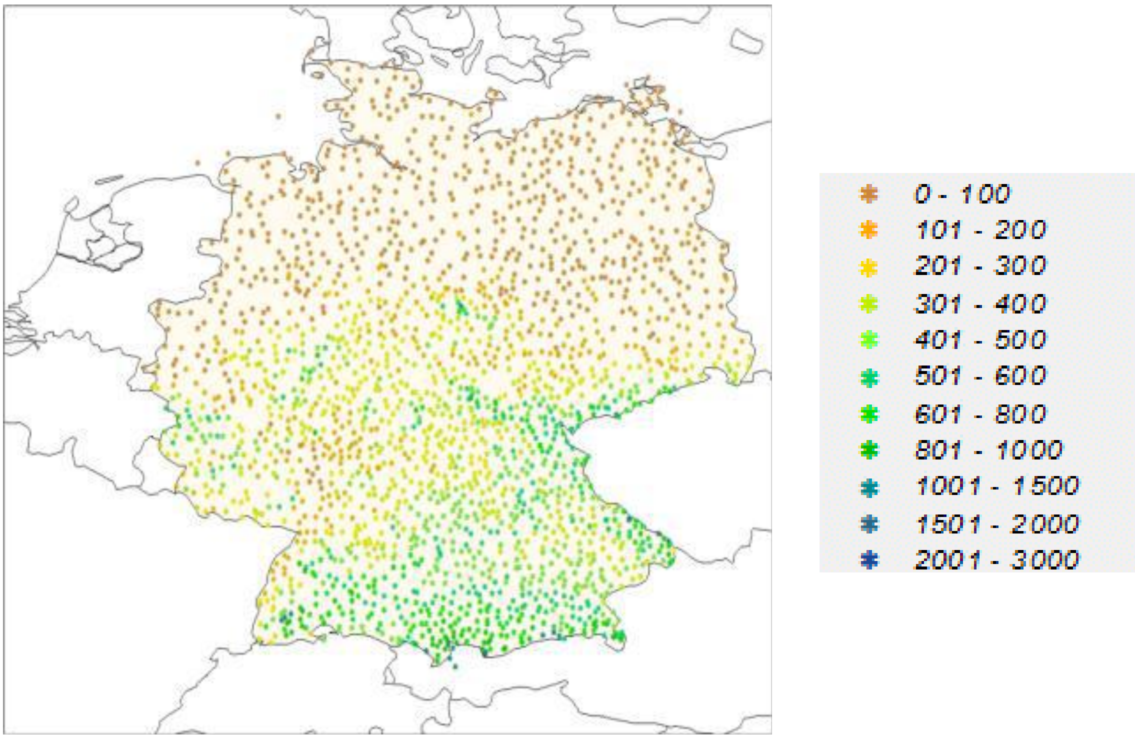

Figure 4 Distribution of the DWD stations measuring snow in the year 2012-2013 (2067 stations)
Stations are coloured with respect to their elevation in meter a.s.l (legend on the right

The in-situ data includes a sub-set of weather stations with 1278 stations where snow as well as temperature measurements are available. To avoid uncertainties due to interpolation of temperature from neighbouring stations only the stations with both measurements were taken into account. In the mean 125 stations have been reported daily a snow depth >= 2cm in the period =ct. 2012 to May 2013.

2.6 Ground data in Poland

In Poland, for validation of H11 product, measurements of snow depth parameter and temperature from ground stations (SYNOP network) were used.
No Quality Control has been performed on this data, however, measurements on SYNOP stations are made by qualified staff and good quality of this dataset must be emphasized. Measurements of snow

depth are made on a daily basis at 0600UTC. Measurements of temperature are performed at 1 hour regime.
Data was retrieved from local SH (System of Hydrology) database. For the summer season we expect the dataset is not complete, especially when the mountainous areas are concerned.
Maximum number of stations yielding both snow depth and hourly temperature measurements is 53. Three stations belong to the mountainous area.

code	name	height [a.s.l.]	lon	lat
100	KOŁOBRZEG	3	15.579722	54.1825
105	KOSZALIN	32	16.1555555	54.2044444
115	USTKA	6	16.8622222	54.5883333
120	ŁEBA	2	17.5347222	54.7536111
125	LĘBORK	38	17.723611	54.553055
135	HEL	1	18.8116666	54.6036111
140	GDAŃSK - PORT PÓŁNOCNY	2	18.6977777	54.3997222
160	ELBLĄG	40	19.4313888	54.1619444
185	KĘTRZYN	108	21.369444	54.068333
195	SUWAŁKI	184	22.9488888	54.1308333
200	ŚWINOUJŚCIE	6	14.2411111	53.9227777
205	SZCZECIN-DąBIE	1	14.6227777	53.3952777
210	RESKO	52	15.414722	53.771666
215	SZCZECINEK	137	16.746666	53.714722
230	PIŁA	72	16.7483333	53.1311111
235	CHOJNICE	164	17.5325	53.7152778
250	TORUŃ	69	18.595	53.0419444
270	MŁAWA	147	20.3611111	53.1041666
272	OLSZTYN	133	20.4227777	53.7711111
280	MIKOŁAJKI	127	21.58333	53.78333
285	OSTROŁĘKA	94	21.535277	53.0675
295	BIAŁYSTOK	148	23.1722222	53.1083333
300	GORZÓW WIELKOPOLSKI	72	15.2772222	52.7411111
310	SŁUBICE	21	14.5938888	52.3486111
330	POZNAŃ	83	16.8358333	52.4172222
345	KOŁO	115	18.6613888	52.2002777
360	PŁOCK	106	19.7258333	52.5883333
375	WARSZAWA OKĘCIE	106	20.9611111	52.1627777
385	SIEDLCE	152	22.2447222	52.1811111
399	TERESPOL	133	23.6213888	52.0783333
400	ZIELONA GÓRA	192	15.5244444	51.93
415	LEGNICA	122	16.2075	51.1925
418	LESZNO STRZYŻEWICE	91	16.5375	51.8277777
424	WROCŁAW	120	16.9011111	51.1055555
435	KALISZ	138	18.0819444	51.7819444
455	WIELUŃ	200	18.5577777	51.2111111
465	ŁÓDŹ	187	19.4022222	51.7272222
469	SULEJÓW	188	19.8663888	51.35
488	KOZIENICE	123	21.5436111	51.5647222
495	LUBLIN-RADAWIEC	239	22.3936111	51.2166666
497	WŁODAWA	177	23.5308333	51.5533333

code	name	height [a.s.l.]	lon	lat
500	JELENIA GÓRA	342	15.7886111	50.9002777
510	ŚNIEŻKA	1603	15.7413888	50.7369444
520	KŁODZKO	356	16.6141666	50.4369444
530	OPOLE	165	17.97	50.6277777
540	RACIBÓRZ	205	18.1916666	50.0616666
550	CZĘSTOCHOWA	293	19.0925	50.8125
560	KATOWICE MUCHOWIEC	284	19.0322222	50.2413888
566	KRAKÓW-BALICE	237	19.8019444	50.0719444
570	KIELCE	260	20.6922222	50.8105555
575	TARNÓW	209	20.9844444	50.0188888
580	RZESZÓW-JASIONKA	195	22.0291666	50.1152777
585	SANDOMIERZ	217	21.7158333	50.6966666
595	ZAMOŚĆ	212	23.206388	50.697777
600	BIELSKO-BIAŁA	398	19.0011111	49.8080555
625	ZAKOPANE	855	19.9602777	49.2938888
650	KASPROWY WIERCH	1991	19.9819444	49.2325
660	NOWY SąCZ	292	20.6891666	49.6272222
670	KROSNO	326	21.7691666	49.7066666
690	LESKO	420	22.3416666	49.4663888
695	PRZEMYŚL	279	22.771944	49.804166

Table 4 SYNOP stations in Poland (metadata)

2.7 Ground data in Turkey

Turkey has not taken part in product generation and validation of the snow status product (H11). However, the daily snow depth dataset as well as the hourly temperature observations from synoptic stations are provided to Finnish colleagues to be used in the product generation and/or validation of the H11 product.

The snow depth and hourly temperature dataset covers the time period from 01.10.2009 to 31.09.2010 which has been considered as a common period for the validation of H11 product. The daily snow depth observations are reported by each synoptic station in centimeters at 06:00 UTC while temperature observations are available on top of the every UTC hours.

The total number of the synoptic station considered for providing snow depth and hourly temperature data over Turkey is 132. The spatial distribution of the stations including the elevation information is provided in the following figure.

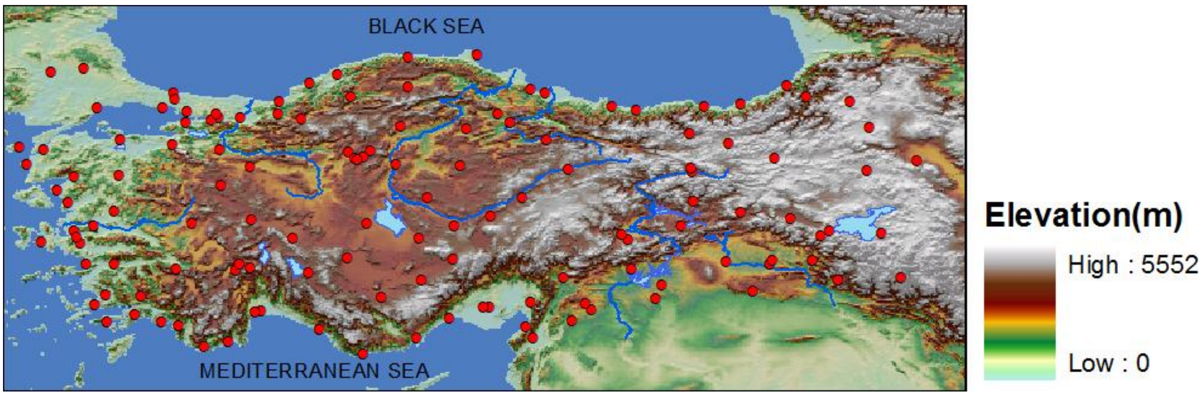

Figure 5 Spatial distribution of synoptic stations used for providing snow depth and hourly temperature dataset for validation of H11 product

2.8 Some conclusions

After this inventory some conclusions can be drawn. First, the H11 has been validated in five countries: Belgium, Finland, Poland. Slovakia, Turkey. These countries have different climatology and orography. From the Table 3 it is evident that the snow data used in this validation activities derived by different types of measurement stations.

In the future an analysis of the ground data error will be conducted in order to associate a ground data error to the validation results and to select the ground data really representative of the ground snow.

3 Validation results: case study analysis

3.1 Introduction

As reported in the Section 3 the common validation methodology is composed of large statistic (and case study analysis. Both components (large statistic and case study analysis) are considered complementary in assessing the accuracy of the implemented algorithms. Large statistics helps in identifying existence of pathological behaviour; selected case studies are useful in identifying the roots of such behaviour, when present.

This Chapter collects the case study analysis performed by all the Institutes involved in the SPVG on H11 for the reference period 01.10.2012 to 31.05.2013. The Chapter is structured by Country / Team, one section each. The analysis has been conducted to provide information to the User of the product on the variability of the performances with climatological and morphological conditions.

Each section presents the case studies analysed giving the following information:

- description of case study,
- ground data and tools used,
- comparison,
- some conclusions.

In the future the SPVG will test the possibility to present case study analysis also in the test sites indicated by the hydrological validation team, in order to provide a complete product accuracy and hydrological validation analysis to the users.

3.2 Two Case study analysis in Belgium (IRM)

Two cases of snow cover are presented. The first one concerns the abundant snowfalls that were observed during March 2013. More precisely, in Figures 1 and 2 are presented the original satellite image, the result of the processing of satellite data for Belgium and the station status for March 13, 2013.

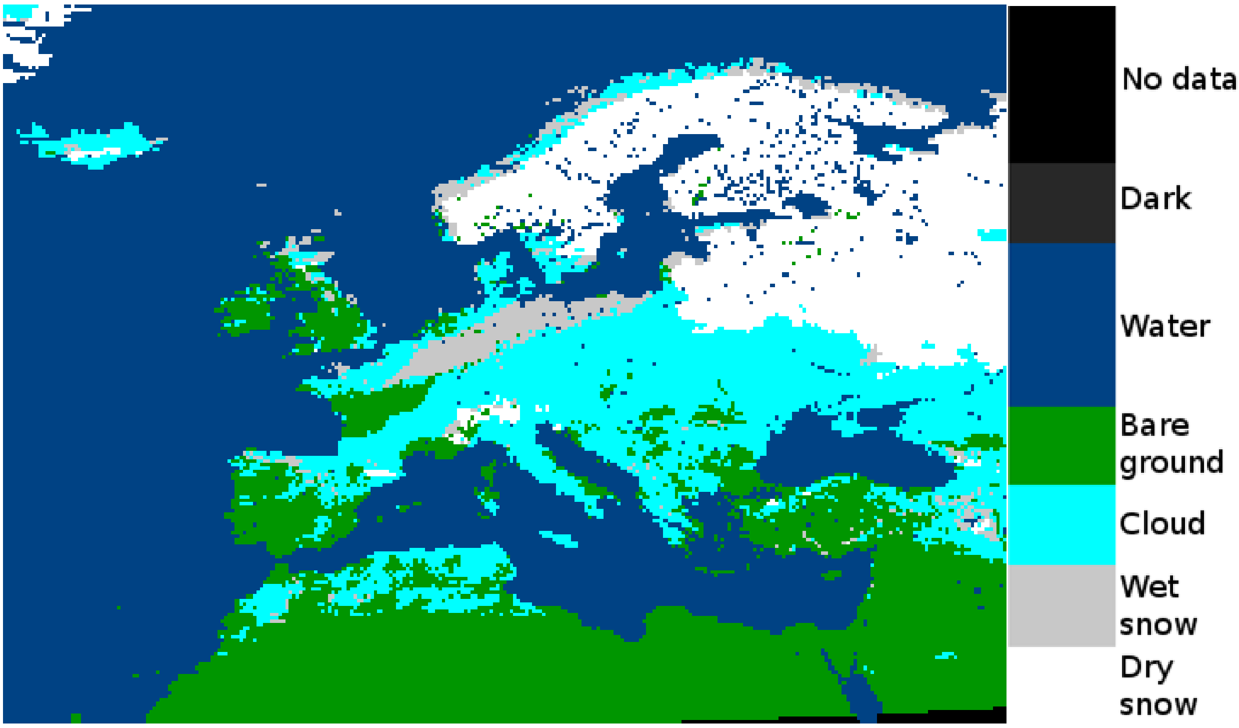

Figure 6 Satellite image of snow status over the H-SAF domain (March 13, 2013).

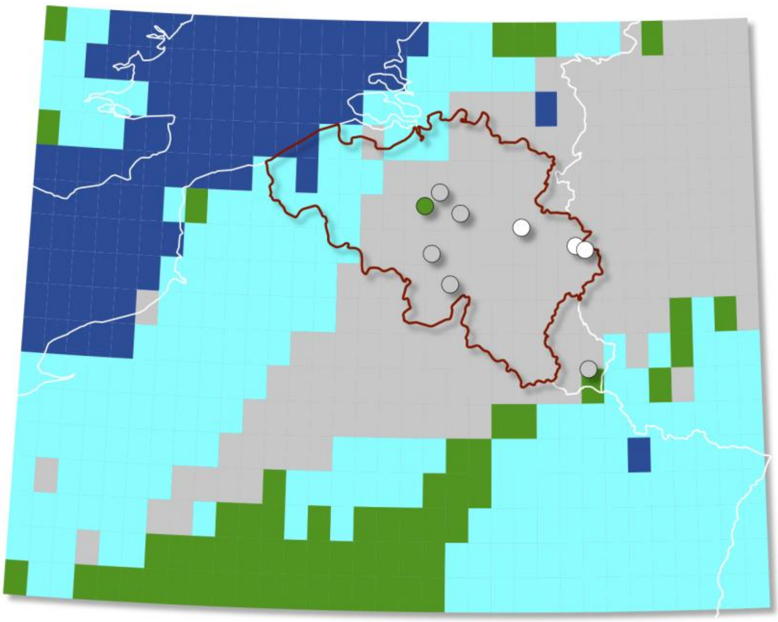

Figure 7 Satellite image for snow status over Belgium, including data from 9 stations (March 13, 2013). The colour code for stations is the same as for the satellite).

In this case, only 1 out of 9 stations reports no snow while the satellite image shows snow over all the stations. The satellite detection shows wet snow but we see three stations in the Ardennes area reporting dry snow.

The second case refers to the snow cover of January 16, 2013. In Figures 3 and 4 one can see the situation at the scales of the H-SAF domain and of Belgium. Presence of snow regardless of its status (dry or wet) is rather well detected, with 5 or 6 stations reporting snow located on or near satellite pixels detecting also snow. However the snow status is detected correctly for the data of only 1 out of 9 stations.

It is noteworthy that the coastal pixels from the satellite are not the same across the two cases considered. Moreover, there are two land pixels in the image from March 13 for which water is encoded in the corresponding GRIB file. These pixels are located in places where no known large water areas exist. Quality control check did not show any anomaly that could account for such differences.

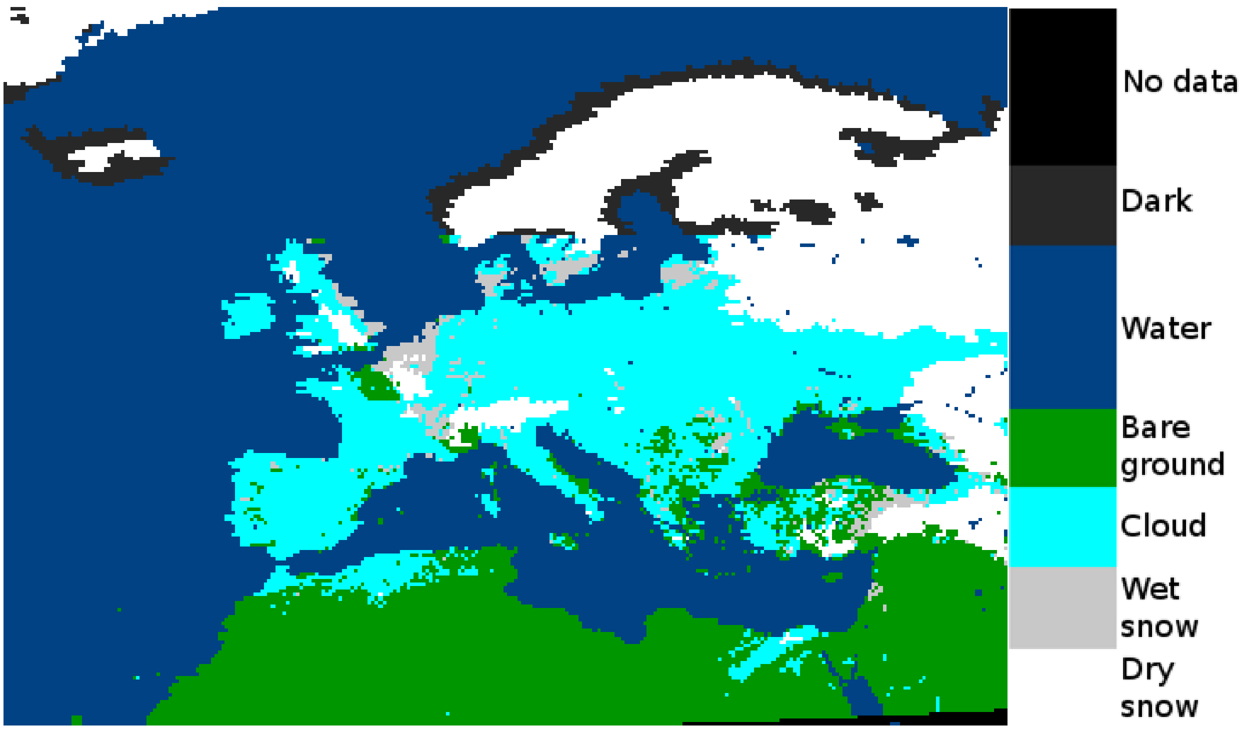

Figure 8 Satellite image of snow status over the H-SAF domain (January 16, 2013)

3.3 Two Case study analysis in Poland (IMWM)

Two cases of snow cover are presented.

The first one concerns 21 December 2013. The weather in the northern and north-western parts of Poland was sunny and frosty. In the south-western part clouds were dominating. The temperature range was between -15 and -5 degrees Celsius. Also couple of days before the day the temperatures were mostly below 0 deg. C.

H-SAF product performance is quite good. Wet snow detected in areas in vicinity of the Baltic sea was mostly confirmed by station measurements. Dry snow in large areas in the eastern part also fits station data. For mountainous areas the conditions were rather cloudy but some stations detected snow which is not visible on the product, thus negatively affecting probability of detection.

Figures 1 and 2 present the H11 product and NOAA channels 1,2,3 composition product respectively.

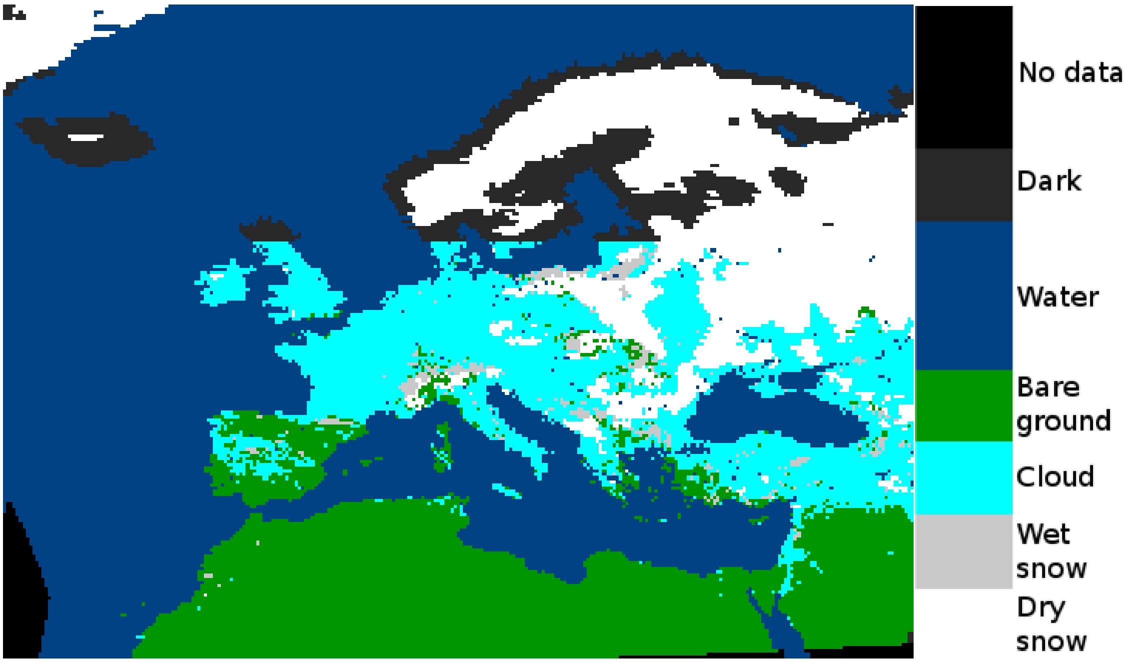

Figure 9 Satellite image of snow status over the whole H-SAF domain (21 December 2012)

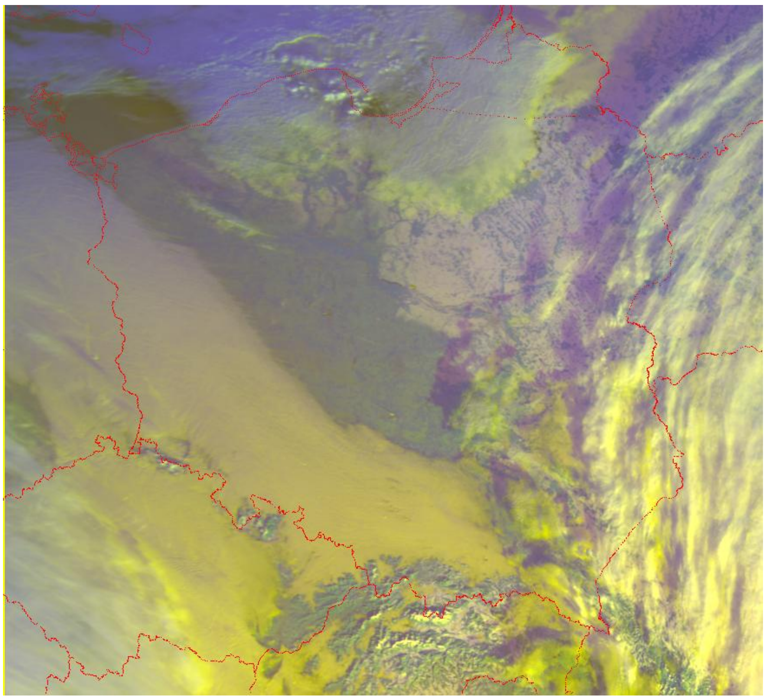

Figure 10 Satellite image with NOAA 1,2,3 channels colour composition. Clouds visible over south-western territory of Poland. Snow below thin clouds in north and north-eastern parts of the country (21 December 2012).)

The second case refers to the snow cover of 3 April 2013. Masses of cold and dry arctic air influenced the weather in Poland. It was rather cold for this time of the year, with temperatures ranging from -2 to +4 deg. C during the day. In the southern and south-eastern part of the country clouds were dominating and snow precipitation was measured.

Presence of wet snow by the seaside is rather well detected. Large areas covered by snow in the north were mainly confirmed as dry by station data. However for some points the snow considered as dry by satellite product was detected as wet by the station. Green ground patches detected in the north are confirmed by SYNOP measurements.

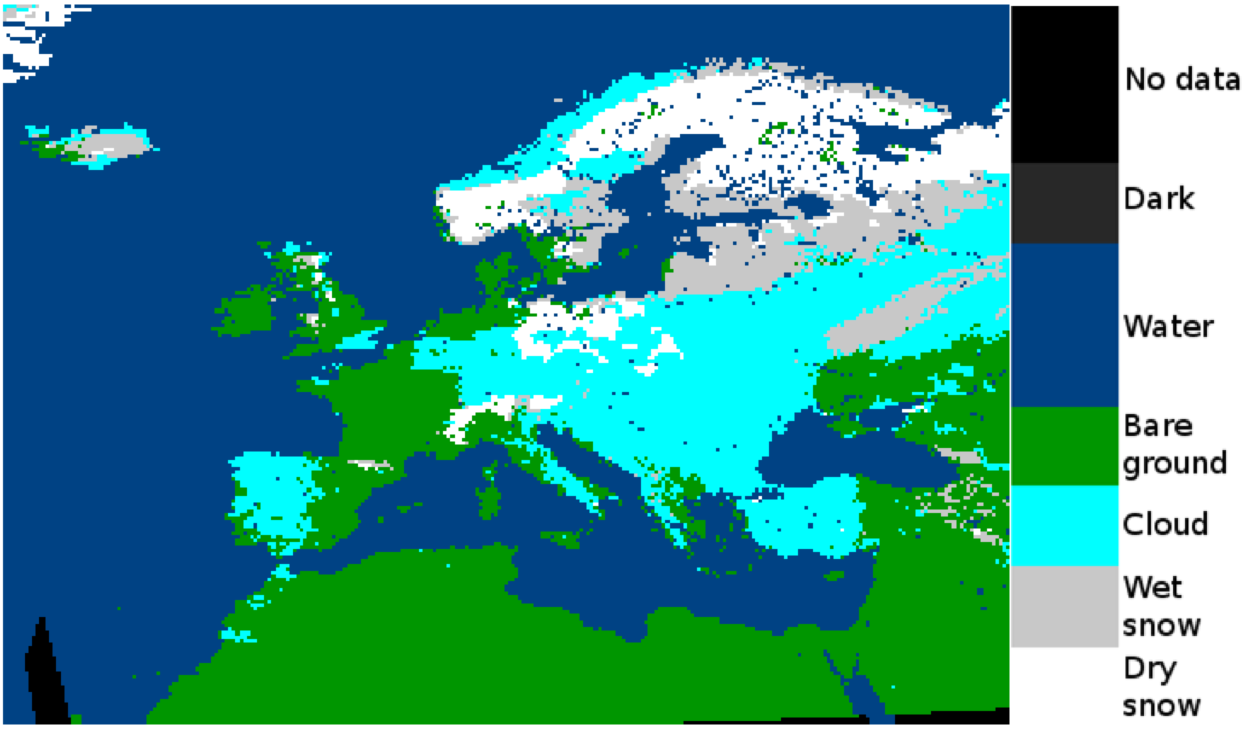

Figure 11 Satellite image of snow status over the H-SAF domain (3 April 2013).)

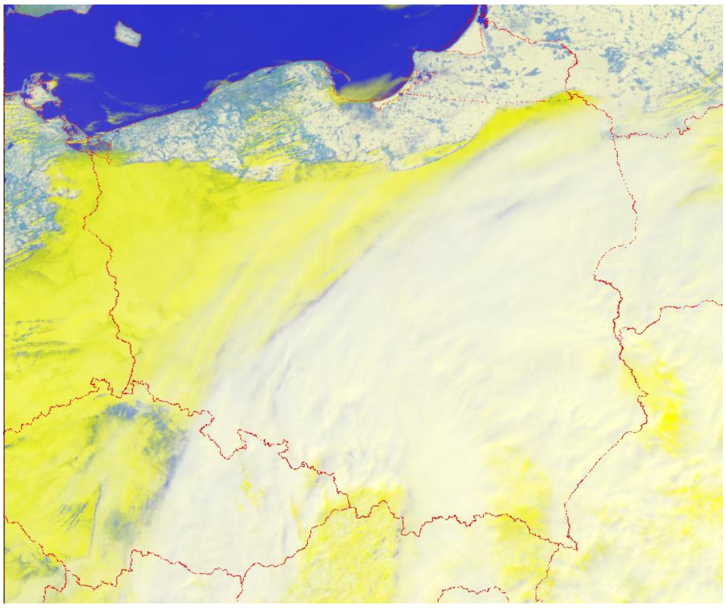

Figure 12 Satellite image with NOAA 1,2,3 channels colour composition. Clouds visible over south-western territory of Poland. Snow below thin clouds in north and north-eastern parts of the country (3 April 2013).)

4 Validation results: long statistic analysis

4.1 Introduction

In this Chapter the validation results of the H11 long statistic analysis for the reference period (1.10.2012 – 31.05.2013) are reported. The validation has been performed on the product release currently in force at the time of writing.
Each Country/Team contributes to long statistic validation by providing the monthly contingency table and the statistical scores. The results are showed for flat and mountainous areas, and for merged product to provide complete error information to the User on the product performances related to the orography.
In order to assess the degree of compliance of H11 with product requirements, all the SPVG members provided the long statistic results following the validation methodology reported in Chapter 3.
For H11, product requirements are recorded in the next table:

Score	threshold	target	optimal
POD	0.60	0.80	0.90
FAR	0.20	0.10	0.05

Table 5 Accuracy requirements for product SN-OBS-1

This implies that the main score to be evaluated are the Probability Of Detection (POD) and the False Alarm Rate (FAR). However, in order to give a more complete idea of the product error structure, several statistical scores have been evaluated, as reported in Section 3.9.

The average performance of H11 for all sites is presented in a compact, synoptic way in Section 6.9. The contents of the monthly statistical scores have been provided by the individual Countries/Teams and verified by the Validation Cluster Leader, step by step, as described in the Section 3. As stressed in Section 4, the average scores reported on Tables 10-11 have been obtained on measurements collected in heterogeneous geographical, orographical and climatological conditions. The variability of the validation results has been 'measured' by the evaluation of the standard deviation of the country statistical scores.

The results obtained in Belgium, Finland, Germany, Poland, and Turkey are here reported in a synoptic way. The country validation results are here reported in order to respond not only to the question whether the product meets the requirements or not, but also if where meets or approaches or fails the requirements.

4.2 Long statistic results in Belgium

Belgium										
	Oct 2012	Nov 2012	Dec 2012	Jan 2013	Feb 2013	Mar 2013	Apr 2013	May 2013	Total	
Flat area										
hits	0	0	0	4	3	7	0	0	14	465
false al.	0	0	13	12	18	16	4	0	63	69
misses	0	0	8	54	35	20	3	0	120	100
corr neg.	130	40	4	0	14	41	126	101	456	263
# obs	130	40	25	70	70	84	133	101	653	681
pod	-	-	0	0,07	0,08	0,26	0	-	0,1	0,04
far	-	-	1	0,75	0,86	0,7	1	-	0,82	0,15
csi	-	-	0	0,06	0,05	0,16	0	-	0,07	0,12
pofd	0	0	0,76	1	0,56	0,28	0,03	0	0,12	0,14
acc	1	1	0,16	0,06	0,24	0,57	0,95	1	0,72	0,08
hss	-!	-	-0,66	-0,39	-0,46	-0,02	-0,03	-	-0,02	0,19

Table 6 Monthly contingency tables and the statistical scores evaluated for flat product in Belgium

The altitude of Belgium is lower than 694 m and this is the reason why we have the statistics only for flat area.
Weather circulation over Belgium is dominated by low pressure systems coming from the Atlantic. Therefore the cloud cover of the country is frequent and high. This is reflected in Table 2, where we can see the number of cases where the stations were under the clouds, sorted by snow status (dry, wet and bare ground). In this case the cloud detection is provided by the satellite.

		Oct 2012	Nov 2012	Dec 2012	Jan 2013	Feb 2013	Mar 2013	Apr 2013	May 2013
Number of stations under clouds reporting:	**Dry snow**	0	0	8	54	35	20	3	0
	Wet snow	0	0	20	17	34	27	0	0
	Bare ground	113	212	234	183	148	141	137	169

Table 2. Station status under clouds (from satellite detection).

The effect of this condition is strong also on the scores, yielding a very low POD and a high FAR over the country. In fact, if snow is reported on the ground and clouds are reported by the satellite product, the comparisons are classified as “misses”, and this is strongly affecting the results if the sky is often cloudy.

4.3 Long statistic results in Finland

Finland									
	Oct 2012	Nov 2012	Dec 2012	Jan 2013	Feb 2013	Mar 2013	Apr 2013	May 2013	Total
Flat area									
hits	22	173	933	1671	1286	1303	335	33	5756
false al.	155	236	21	116	231	331	387	32	1515
misses	106	86	1268	111	58	22	27	21	1699
corr neg.	64	15	0	0	0	0	207	2008	6415
# obs	347	510	2222	1898	1575	1656	956	2094	15385
pod	0,17	0,67	0,42	0,94	0,96	0,98	0,93	0,61	0,77
far	0,88	0,58	0,02	0,06	0,15	0,20	0,54	0,49	0,21
csi	0,09	0,41	0,98	0,94	0,85	0,80	0,36	0,02	0,64
pofd	0,59	0,73	0,02	0,51	0,80	0,94	0,93	0,60	0,19
acc	0,37	0,51	0,99	0,94	0,85	0,80	0,38	0,03	0,79
hss	-0,25	0,15	0,98	0,63	0,29	0,09	-0,33	-0,03	0,58

Table 7 Monthly contingency tables and the statistical scores evaluated for flat and merged product in Finland

In Finland the overall results are really good for winter months with lots of snow, although monthly results have quite many errors (misses and false alarms) during periods with either thin snow cover or patchy snow. Some of the errors might be a result of these conditions. In the first case, microwave data cannot detect the dry snow due to large penetration depth, and thus the ground surface dominates the emission received by the satellite. The later situation is also problematic, as the snow is patchy and the conditions at the weather station (problem also in good snow cases) may be in minority of the large (roughly 25 km by 25 km) output pixel. If these purely logical error sources are "filtered out", it still seems that the algorithm used for radiometer data has problems with snow which is in the border of being wet or dry. (see validation report of H10 to see the accuracy of optical snow recognition product used as a part of this snow status product).

4.4 Long statistic results in Germany

Germany									
	Oct 2012	Nov 2012	Dec 2012	Jan 2013	Feb 2013	Mar 2013	Apr 2013	May 2013	Total
mountainous									
Hits	3	0	56	19	10	14	0	0	102
false alarms	1	0	47	14	50	11	0	0	123
misses	1	0	47	14	50	11	0	0	123
corr negatives	3	0	56	19	10	14	0	0	102
numbers of obs	8	0	206	66	120	50	0	0	450
Pod	0,75	-	0,54	0,58	0,17	0,56	-	-	0,45
Far	0,25	-	0,46	0,42	0,83	0,44	-	-	0,55
Csi	0,60	-	0,37	0,40	0,09	0,39	-	-	0,29
Pofd	0,25	-	0,46	0,42	0,83	0,44	-	-	0,55
Acc	0,75	-	0,54	0,58	0,17	0,56	-	-	0,45
Hss	0,50	-	0,09	0,15	-0,67	0,12	-	-	-0,09
flat area									
Hits	22	13	754	429	601	292	54	5	2170
false alarms	33	6	795	595	579	693	33	2	2736
misses	33	6	795	595	579	693	33	2	2736
corr negatives	22	13	754	429	601	292	54	5	2170
numbers of obs	110	38	3098	2048	2360	1970	174	14	9812
Pod	0,40	0,68	0,49	0,42	0,51	0,30	0,62	0,71	0,44
Far	0,60	0,32	0,51	0,58	0,49	0,70	0,38	0,29	0,56
Csi	0,25	0,52	0,32	0,26	0,34	0,17	0,45	0,56	0,28
Pofd	0,60	0,32	0,51	0,58	0,49	0,70	0,38	0,29	0,56
Acc	0,40	0,68	0,49	0,42	0,51	0,30	0,62	0,71	0,44
Hss	-0,20	0,37	-0,03	-0,16	0,02	-0,41	0,24	0,43	-0,12
mountainous and flat area									
Hits	25	13	810	448	611	306	54	5	2272
false alarms	34	6	842	609	629	704	33	2	2859
misses	34	6	842	609	629	704	33	2	2859
corr negatives	25	13	810	448	611	306	54	5	2272
numbers of obs	118	38	3304	2114	2480	2020	174	14	10262
Pod	0,42	0,68	0,49	0,42	0,49	0,30	0,62	0,71	0,44
Far	0,58	0,32	0,51	0,58	0,51	0,70	0,38	0,29	0,56
Csi	0,27	0,52	0,32	0,27	0,33	0,18	0,45	0,56	0,28
Pofd	0,58	0,32	0,51	0,58	0,51	0,70	0,38	0,29	0,56
Acc	0,42	0,68	0,49	0,42	0,49	0,30	0,62	0,71	0,44
Hss	-0,15	0,37	-0,02	-0,15	-0,01	-0,39	0,24	0,43	-0,11

In Germany the overall results are generally moderate (POD mean 0.44) with exception of the months November/April and May (0.68/0.62/0.71). The false alarm rate reach moderate values too (Far mean 0.56). The lowest values are again the months November/April and May (0.32/0.38/0.29).

Furthermore, the validation statistic was carried out for dry and wet snow separately. Out of this statistic it becomes clear that the classification in dry snow conditions at station locations relative low POD of 0.4 are reached while und wet conditions at station location the POD becomes 0.82 in the mean. Unfortunately the FAR is converse of these situations (0.1 dry and 0.75 wet).

Known uncertainties affecting the validation results are that the procedure for calculating areal mean of an H11 raster size out of the station data has discover the problem of a clear distinction between dry and wet status over an area based on station data. In the 266 cases were a value of 1.5 was calculated it has been seen that this could be attributed to the transient state between a day with dry snow condition a day before and a wet snow condition a day after. Therefore, the problem of transient state becomes prominent in the classification of the station data, too.

The uncertainty introduced by the aggregation method of station data as well as for defining the thresholds for defining wet and dry status out of station data need more detailed analysis and has to be quantified.

4.5 Long statistic results in Poland

POLAND									
	Oct 2012	Nov 2012	Dec 2012	Jan 2013	Feb 2013	Mar 2013	Apr 2013	May 2013	Total
Mountainous									
Hits	6	3	13	11	11	9	4	3	60
false al.	1	2	2	3	1	1	0	1	11
Misses	2	0	6	5	6	3	3	0	25
corr neg.	12	8	9	4	4	6	6	8	57
# of obs	21	13	30	23	22	19	13	12	153
Pod	0,75	1,00	0,68	0,69	0,65	0,75	0,57	1,00	0,71
Far	0,14	0,40	0,13	0,21	0,08	0,10	0,00	0,25	0,15
Csi	0,67	0,60	0,62	0,58	0,61	0,69	0,57	0,75	0,63
Pofd	0,08	0,20	0,18	0,43	0,20	0,14	0,00	0,11	0,16
Acc	0,86	0,85	0,73	0,65	0,68	0,79	0,77	0,92	0,76
Hss	0,69	0,65	0,47	0,24	0,33	0,57	0,55	0,80	0,53
flat area									
Hits	32	14	44	30	16	50	38	10	234
false al.	8	3	10	11	6	9	10	0	57
Misses	11	6	15	6	4	17	11	13	83
corr neg.	80	39	40	6	19	15	60	94	353
# of obs	131	62	109	53	45	91	119	117	727
Pod	0,74	0,70	0,75	0,83	0,80	0,75	0,78	0,43	0,74
far	0,20	0,18	0,19	0,27	0,27	0,15	0,21	0,00	0,20
csi	0,63	0,61	0,64	0,64	0,62	0,66	0,64	0,43	0,63
pofd	0,09	0,07	0,20	0,65	0,24	0,38	0,14	0,00	0,14
acc	0,85	0,85	0,77	0,68	0,78	0,71	0,82	0,89	0,81
hss	0,67	0,65	0,54	0,20	0,55	0,34	0,63	0,55	0,60
mountainous and flat area									
hits	38	17	57	41	27	59	42	13	294
false al.	9	5	12	14	7	10	10	1	68
misses	13	6	21	11	10	20	14	13	108
corr neg.	92	47	49	10	23	21	66	102	410
# of obs	152	75	139	76	67	110	132	129	880
pod	0,75	0,74	0,73	0,79	0,73	0,75	0,75	0,50	0,73
far	0,19	0,23	0,17	0,25	0,21	0,14	0,19	0,07	0,19
csi	0,63	0,61	0,63	0,62	0,61	0,66	0,64	0,48	0,63
pofd	0,09	0,10	0,20	0,58	0,23	0,32	0,13	0,01	0,14
acc	0,86	0,85	0,76	0,67	0,75	0,73	0,82	0,89	0,80
hss	0,67	0,65	0,53	0,21	0,49	0,39	0,62	0,59	0,59

Table 8 Monthly contingency tables and the statistical scores evaluated for mountainous, flat and merged product in Poland. ***Validation period: 01/10/2012 – 31/05/2013***

Weather conditions in Poland are mostly dominated by moist polar air masses from North-West bringing clouds and precipitating rain in winter season on one hand and by dry, continental polar masses from North-East on the other hand. Therefore the cloud cover of the country is frequent and high. This limits remarkably the number of validation measurements taken into consideration in this study.

4.6 Long statistic results in Turkey

Turkey									
	Oct 2012	Nov 2012	Dec 2012	Jan 2013	Feb 2013	Mar 2013	Apr 2013	May 2013	Total
mountainous									
hits	0	0	43	221	100	59	0	0	423
false alarms	0	0	93	181	148	82	0	0	504
misses	0	0	43	40	8	15	0	0	106
corr negatives	0	0	21	21	37	10	0	0	89
numbers of obs	0	0	200	463	293	166	0	0	1122
pod	-	-	0,50	0,85	0,93	0,80	-	-	0,80
far	-	-	0,68	0,45	0,60	0,58	-	-	0,54
csi	-	-	0,24	0,50	0,39	0,38	-	-	0,41
pofd	-	-	0,82	0,90	0,80	0,89	-	-	0,85
acc	-	-	0,32	0,52	0,47	0,42	-	-	0,46
hss	-	-	-0,29	-0,05	0,10	-0,09	-	-	-0,05
flat area									
hits	0	0	0	19	1	1	0	0	21
false alarms	0	0	2	7	1	0	0	0	10
misses	0	0	1	7	0	0	0	0	8
corr negatives	0	0	2	10	0	2	0	0	14
numbers of obs	0	0	5	43	2	3	0	0	53
pod	-	-	0,00	0,73	1,00	1,00	-	-	0,72
far	-	-	1,00	0,27	0,50	0,00	-	-	0,32
csi	-	-	0,00	0,58	0,50	1,00	-	-	0,54
pofd	-	-	0,50	0,41	1,00	0,00	-	-	0,42
acc	-	-	0,40	0,67	0,50	1,00	-	-	0,66
hss	-	-	-0,36	0,32	0,00	1,00	-	-	0,31
mountainous and flat area									
hits	0	0	43	240	101	60	0	0	444
false alarms	0	0	95	188	149	82	0	0	514
misses	0	0	44	47	8	15	0	0	114
corr negatives	0	0	23	31	37	12	0	0	103
numbers of obs	0	0	205	506	295	169	0	0	1175
pod	-	-	0,49	0,84	0,93	0,80	-	-	0,80
far	-	-	0,69	0,44	0,60	0,58	-	-	0,54
csi	-	-	0,24	0,51	0,39	0,38	-	-	0,41
pofd	-	-	0,81	0,86	0,80	0,87	-	-	0,83
acc	-	-	0,32	0,54	0,47	0,43	-	-	0,47
hss	-	-	-0,29	-0,02	0,10	-0,07	-	-	-0,04

Table 9 Monthly contingency tables and the statistical scores evaluated for mountainous and merged product in Turkey. ***Validation period: 01/10/2012 – 31/05/2013***

The observations indicate that the validation criteria were satisfied for the months of December, January, February and March. More observations were obtained from mountainous areas, the number of observations satisfying the wetness criteria for flat areas were comparatively less. The POD values are quite high and satisfying the expected target threshold values. The number of false alarms is high for the mountainous areas, which increases the FAR values in the calculated statistics. The high number of false alarms is due to the location of the automatic synoptic weather stations, where they are mostly located in the urbanized areas and the overall condition for the H11 product pixel cannot be represented by the location of the station. Enough number of observations cannot be obtained from the automatic weather stations in 2013 snow season in Turkey, this affected the results in worse direction.

4.7 Synopsis of validation results

In the previous sections the Countries/Teams have concluded with highlighting the main positive aspects of the product and the main failures, according to the experience on their area of investigation.

In this Section, the synthesis of the results is presented. Standard deviation (DEV.ST) has been reported to 'measure' the variability of the statistical scores obtained by several countries.

	Finland	Belgium	Poland	Germany	Turkey	TOTAL
mountainous						
hits	0	0	60	102	423	585
false alarms	0	0	11	123	504	638
misses	0	0	25	123	106	254
corr negatives	0	0	57	102	89	248
numbers of obs	0	0	153	450	1122	1725
pod	-	-	0,71	0,45	0,80	0,70
far	-	-	0,15	0,55	0,54	0,52
csi	-	-	0,63	0,29	0,41	0,40
pofd	-	-	0,16	0,55	0,85	0,72
acc	-	-	0,76	0,45	0,46	0,48
hss	-	-	0,53	-0,09	-0,05	-0,02
flat area						
hits	5756	14	234	2170	21	8195
false alarms	1509	63	57	2736	10	4375
misses	1699	120	83	2736	8	4646
corr negatives	2294	456	353	2170	14	5287
numbers of obs	11258	653	727	9812	53	22503
pod	0,77	0,10	0,74	0,44	0,72	0,64
far	0,21	0,82	0,20	0,56	0,32	0,35
csi	0,64	0,07	0,63	0,28	0,54	0,48
pofd	0,40	0,12	0,14	0,56	0,42	0,45
acc	0,72	0,72	0,81	0,44	0,66	0,60
hss	0,37	-0,02	0,60	-0,12	0,31	0,18
mountainous and flat area						
hits	5756	14	294	2272	444	8780
false alarms	1509	63	68	2859	514	5013
misses	1699	120	108	2859	114	4900
corr negatives	2294	456	410	2272	103	5535
numbers of obs	11258	653	880	10262	1175	24228
pod	0,77	0,10	0,73	0,44	0,80	0,64
far	0,21	0,82	0,19	0,56	0,54	0,36
csi	0,64	0,07	0,63	0,28	0,41	0,47
pofd	0,40	0,12	0,14	0,56	0,83	0,48
acc	0,72	0,72	0,80	0,44	0,47	0,59
hss	0,37	-0,02	0,59	-0,11	-0,04	0,17

Table 10 Statistical scores for H11 for period 1. 10.2012-31.5.2013 all statistic. The results are separate for flat, mountainous and merged product

The scores showed in the previous table lead to the following general considerations:

- The effect of cloud coverage is very strong over Belgium, leading to poor scores, while in Germany there is a strong difference of H11 performance between wet snow and dry snow conditions, to be investigated. The other countries report higher performances;
- Unexpectedly, mountainous areas show better performances. This is mainly due to the large predominance of mountains over Turkey, where all results are quite good (the procedure is carefully checked, thus the good results obtained are liable).
- As expected, the FAR is poorer over mountains than flat areas.

4.8 Product requirements compliance

As shown in Table 11 **the statistical scores obtained by the validation of H11 in SPVG reach the thresholds stated in the Product Requirements for POD, but the threshold is not reached by FAR.** The statistical scores have been evaluated on 22503 samples on flat areas and 1725 samples on mountainous area for the reference period (1.10.2012 – 31.05.2013) across five countries with different orographic and climatological characteristics.

Between target and optimal	Between threshold and target	Threshold exceeded by < 50 %	Threshold exceeded by ≥ 50 %

H-SAF Accuracy requirements for H11					
product requirements				*H11*	
Score	**threshold**	**target**	**optimal**	**total**	**DEV.ST**
POD	0.60	0.80	0.90	0,64	0,16
FAR	0.20	0.10	0.05	0,36	0,19

Table 11 Simplified compliance analysis for product H11

Some general consideration when comparing the results from the validation activity with the stated product requirements from the URD:

a. Product requirements - There are reasons to believe that the current set of H-SAF product requirements are too strict;

b. The results obtained by the current validation procedure represent the convolution of at least three factors: the satellite product accuracy, the accuracy of the ground data used and the limitations of the comparison methodology (e.g., errors of space and time co-location, representativeness changing with scale, etc.). Therefore, the results currently found are by far pessimistic in respect of what is the real product performance.

4.9 Some conclusions

In this Section the validation results of the H11 long statistic analysis have been reported for the reference period (1.10.2012 – 31.05.2013). Each Country/Team has provided the monthly contingency table and the statistical scores for flat and mountainous areas, and for merged product, to the Validation Cluster Leader which verified the validation results step by step, as described in the Section 3.

All the SPVG members have performed the long analysis following the validation methodology reported in the Chapter 3.

The average scores reported on the Table 10 have been obtained on measurements collected in heterogeneous geographical, orographical and climatological conditions. The variability of the validation results has been 'measured' by the evaluation of the standard deviation of the country statistical scores (see Table 10).

The statistical scores evaluated in Belgium, Germany, Finland, Poland, and Turkey have pointed out that in flat areas H11 performances are still not fully satisfying. This can be due also to the cloud cover effect, which can be kept into account when calculating the statistics, and to some discrepancies between results in wet and dry snow conditions. The statistics over both areas is dominated by the results over flat areas, due to the small sample of validated data over mountains.

As shown in Table 11 **the statistical scores obtained by the validation of H11 in SPVG reach the thresholds stated in the Product Requirements for POD, but the threshold is not reached by FAR.** The statistical scores have been evaluated on 22503 samples on flat areas and 1725 samples on mountainous area for the reference period (1.10.2012 – 31.05.2013) across five countries with different orographic and climatological characteristics.

Annex 1: Validation methodology for H11 – Snow status

Panu Lahtinen, panu.lahtinen@fmi.fi
Latest update: January 20, 2011

This document describes the methodology applied when validating H-SAF snow product H11 – Snow status (wet/dry) by MW radiometry.

Validation procedure
To properly validate H11 product, the following steps have to be taken:

1. Observation data containing snow depth and air temperatures need to be collected
2. Satellite products and corresponding quality flag files need to be acquired.
3. Both observation and satellite data series need to be checked for consistency.
4. Comparison between the observation data and the product has to be performed.
5. Results of the comparison need to be presented.

1. Observation data
As there are no direct continuous measurements of the snow wetness available, an alternative approach has been developed. During the SNORTEX campaigns in 2008 – 2010, snow moisture was measured using Toikka snow fork. Also, a qualitative test of snow moisture was made at every snow pit from the surface and from the ground level; if it was possible to make a snow ball, the snow had to contain a significant amount of liquid water.

From these data, the wet snow cases were selected and compared to meteorological data, collected at the FMI Sodankylä weather station. By investigating the 2 meter air temperature history, the following rules were formed for dry snow:

The average air temperature from the previous 18 hours before the measurement has to be equal or lower than -1.5 °C

During this 18 hour period, there should not be more than three measurements above the freezing level, i.e. > 0 °C

As the data from the meteorological station was available every 10 minutes, and the official synoptic stations report hourly, the rules need to be slightly adjusted to be more universal. By setting the nominal measurement time of the product to 1800 UTC, the collecting of the temperature data becomes easier, as data are needed only from one day.

Also, it is necessary to make a check that there really is snow on the ground. The WMO operational measurement uncertainty requirement and instrument performance (WMO, 2008) states that snow depth should be measured with an accuracy of 1 cm for depths less or equal to 20 cm. However, in the same document, it is noted that the most common instrument for automatic snow depth measurements (ultrasonic range sensor) has an uncertainty of 2.5 cm. Based on these values, we select 2 cm of snow as "snow on the ground". This also reduces the snow patchiness compared to lower thresholds.

2. Satellite product

H-SAF H11 products are available at the FMI ftp server, *ftp://ftp.fmi.fi* under the directory */HSAF/products/h11*. The data can downloaded using a ftp client, such as FileZilla.

The product is in GRIB2 format, and consists of two separate files, which both are needed for validation. The files named like "h11_yyyymmdd_day_FMI.grib2" hold the actual product, and "h11_yyyymmdd_QC_day_FMI.grib2" has the associated quality control information. Also quicklook-images are supplied in corresponding monthly directories.

Software to convert the data to ASCII is available at the same server: */HSAF/products/utilities/snobs2_grib2_to_ascii.tar.gz*.

3. Data consistency check

To guarantee high quality of the validation it is advised to check if both the observation data and the satellite product are available for all days of the reference season.

4. Comparison between the observation data and the product

The validation conditions for existence of dry snow are:

5 Automatic snow depth measurement states that there is two centimetres or more of snow during the day
6 The average 2 meter air temperature between the 18 hour period between 00-18 UTC has been equal or lower than -1.5 degrees centigrade
7 The 2 meter air temperature is not above 0 degrees centigrade for more than one hour (one hourly measurement, four 15-minutely measurements, etc.) during this same period.

Or the same in more mathematical form

5 $SD \geq 2$ cm
6 $T_{ave,air}(00...18 \text{ UTC}) \leq -1.5$ °C
7 $t_{T>0} \leq 1$ h.

If rule 1) is not true, there is no snow on the ground. If 1) is true, but either one of the rules 2) or 3) are false, the snow is wet. If all the rules are true, snow is considered dry.
There are some shortcomings in these validation rules. Firstly, it is indirect. This approach is based on air temperature data, which does not directly describe the status of the snow pack, especially on wetlands. Secondly, the calibration of the thresholds was made based on limited data on a single location in Finnish Lapland. The selected thresholds might not be representative in other places. Even at the same location it is not certain what the overall representativeness of these rules is.

The comparison between the station data and the H11 products are made daily on pixel level. For a single day, consider one pixel with several weather stations. All different stations increment corresponding snow status counters for that individual product pixel. There are counter arrays for

- the number of no snow cases
- the number of dry snow cases

- the number of wet snow cases.

That is, there are four arrays for every day: the three listed above and of course the product data itself. Each *pixel* forms a *single validation data point*. To avoid the bias that would be caused by mixed-pixel cases, only the unambiguous pixels (the pixels where all the weather stations agree on the status) are included when calculating the statistical metrics. The mixed-pixel cases can be inspected as auxiliary information, although they cannot be included in the statistics defined in the User Requirements.

To standardise the validation, software implementing the methodology described here can be found from */HSAF/products/utilities/h11_validator*. Also a sample ground data for testing purposes are included. The same folder also holds an up-to-date copy of this document and a spreadsheet for calculating the statistical scores.

5. Results of the comparison

The table below shows the contingency table with details how different combinations of the validation results are merged to "hits", "misses", "false alarms" and "correct negatives". Product has value "Dark" if the sun elevation is too low, namely lower than 10 degrees.

		Station data		
		Dry	Wet	No snow
Product	Dry	Hit	False alarm	False alarm
	Wet	False alarm (check QC)	Hit	False alarm
	Dark	Miss (check QC)	NA	NA
	Cloud	Miss (check QC)	NA	NA
	No snow	Miss	Miss	Correct negative

Four fields are marked as not applicable (NA). In these cases it is not possible to get information on the surface status (radiometer instrument detects only dry snow, optical instruments need sunlight and clear skies). Three fields are marked with "check QC". In these cases also the Quality Control files need to be inspected to see if radiometer data has been available. If radiometer data are missing (data gaps between different satellite overpasses), these categories are discarded from the statistics, as the dry snow cannot be detected. From the classification results, different statistical scores can be calculated. To simplify the formulae, let's use the following notation

- A = number of hits
- B = number of false alarms
- C = number of misses
- D = number of corect negatives

Probability of detection

- POD = A/(A+C)

False alarm ratio

- FAR = B/(A+B)

Probability of false detection

6. POFD = B/(B+D)

Accuracy

- ACC = (A+D)/(A+B+C+D)

Critical success index

- CSI = A/(A+B+C)

Heidke skill score

- HSS = 2(AD-BC) / [(A+C)(C+D) + (A+B)(B+D)]

These statistics need to be calculated on as large area as possible to find the overall performance of the product.

The results are reported separately for each country, and also a combined statistics are presented. To better see when there are most errors, the results are displayed also on a monthly basis. A sample validation report table is shown in Appendix (a separate spreadsheet with the formulae included). This spreadsheet can be used as a template for reporting and calculating the statistics. Some of the statistical scores cannot be calculated, if one of the categories (hits, false alarms, misses, correct negatives) in the dataset is empty.

References

WMO Guide to Meteorological Instruments and Methods of Observation. WMO-No. 8, Seventh edition. 2008.