LIFE Project Number

ENV/FIN/000133

1st EO- data document (years 2001-2008)

Reporting Date

30/11/2009

Actions

Action 2: Satellite Data Processing by FMI
Action 3: Acquisition and extension of GMES-services GSE Polar View and GSE Land (SYKE)

LIFE+ PROJECT NAME or Acronym

SNOWCARBO

Author

Name Beneficiary	Finnish Meteorological Institute (FMI) Finnish Environment Institute (SYKE)
Contact person	FMI: Mr Juha Lemmetyinen SYKE: Mr Mikko Kervinen
Postal address	FMI: P.O. Box 503, FI-00101 Helsinki, Finland SYKE: P.O. Box 140, FI-00251, Helsinki, Finland
Telephone	FMI: +358-09-1929 4663 SYKE: +358 400 397 668
Fax:	FMI: SYKE:+358 9 5490 2690
E-mail	FMI: Juha.Lemmetyinen@fmi.fi SYKE: Mikko.Kervinen@enviroment.fi
Project Website	http://snowcarbo.fmi.fi

Table of contents

List of abbreviations

AKO	Automaattinen Käytön Ohjaus (Automatic usage control)
AMSR-E	Advanced Microwave Scanning Radiometer
AOI	Area Of Interest
ARC	Arctic Research Centre
AVHRR	Advanced Very High Resolution Radiometer
EASE *Grid*	Equal-Area Scalable Earth *Grid*
ECMWF	European Centre for Medium-Range Weather Forecasts
EO	Earth Observation
FMI	Finnish Meteorological Institute
LAADS	Level 1 and Atmosphere Archive and Distribution System
MERIS	Medium Resolution Imaging Spectrometer
MODIS	Moderate Resolution Imaging Spectroradiometer
NAPS	Nearly Autonomous Processing System
NASA	National Aeronautics and Space Administration
NDSI	Normalized Differential Snow Index
NDVI	Normalized Differential Vegetation Index
NDWI	Normalized Differential Water Index
NOAA	National Oceanic and Atmospheric Administration
NSIDC	National Snow and Ice Data Centre
SCA	Snow Covered Area
SSM/I	Special Sensor Microwave/Imager
SWE	Snow Water Equivalent
SYKE	Suomen Ympäristökeskus, Finnish Environmental Institute
TBC	To Be Confirmed
TOA	Top Of Atmosphere

Summary

In this document is described the current status of the Earth observation data products of Actions 2, 3 and 7.

Action 2 Satellite data processing by FMI

Weekly snow water equivalent grids, dates of snow clearance, snow melt onset and soil freeze grids.

Action 3 Acquisition and extension of GMEs Services GSE Polar View and GSE Land (SYKE)

AVHRR dataset for brightness temperature estimation and the MODIS dataset for estimation of the green vegetation status (NDVI) and fractional snow cover (SCA). These datasets include all of the processing chain steps from Level 1 instrument data to the final EO products.

Action 7 Methodology development and implementation (SYKE)

The raw time series which consist of the daily and weekly composites of the EO products NDVI, SCA and brightness temperature estimates from years 2001-2008.

FMI EO Dataset descriptions (Action 2)

Products Summary

General content

The products described in this document are

- **Weekly Snow Water Equivalent Grids:** Gridded weekly average value (mm) of snow water equivalent in Eurasia, obtained by an assimilation of ground station interpolated background maps and passive microwave satellite observations (NASA AMSR-E and SSM/I). The product includes an error estimate of the SWE value. The final dataset spans over 30 years from 1978 to present. *The current prototype dataset includes years 2000 to 2008.*

- **Date of Snow Melt Onset Grids:** Gridded yearly date of spring onset of snow melt in Eurasia, as derived from passive microwave satellite observations (NASA AMSR-E and SSM/I). Optionally daily flagged grid of snow status. The final dataset spans over 30 years from 1978 to present. *Currently no prototype dataset is available.*

- **Date of Snow Clearance Grids:** Gridded yearly date of spring snow clearance in Eurasia, as derived from passive microwave satellite observations (NASA AMSR-E and SSM/I). Optionally daily flagged grid of snow status. The final dataset spans over 30 years from 1978 to present. *The final dataset has been completed.*

- **Date of Soil Freezing Grids:** Gridded yearly date of autumn soil, freezing in Eurasia, as derived from active (scatterometer) microwave satellite observations (NASA QuikScat and ESA ASCAT). The final dataset spans over 10 years from 1999 to present. *Currently no prototype dataset is available.*

Data sources

The data for **Weekly Snow Water Equivalent Grids, Date of Snow Melt Onset Grids** and **Date of Snow Clearance Grids** are based on AMSR-E and SSM/I brightness temperature observations [SD-1], [SD-2], acquired through the National Snow and Ice Data Centre. The products are free of charge to the user. The DMSP SSM/I Pathfinder daily EASE-Grid brightness temperatures span the years 1978-2002. AMSR-E/Aqua daily EASE-Grid brightness temperatures are available from 2003.

The Date of Soil Freezing Grids: final source data is to be confirmed. Possible data sources are NASA QuikScat/SeaWinds or Metop ASCAT scatterometer σ^0 (backscatter) observations. The whole dataset of QuikScat observations is available through the NASA Scatterometer Climate Record Pathfinder project.

Raw Data Access and storage

Summary of the current status of AMSR-E and SSM/I brightness temperature observations is given in Table 1. Data is organized in a yearly/monthly directory structure on duplicated external hard drives. Archive dataset is available for project partners upon request.

Table 1 Summary of the collected AMSR-E/SSM/I raw data archive

	2001	2002	2003	2004	2005	2006	2007	2008
Format	Raw binary	Raw binary	Raw binary	Raw binary	Raw binary	Raw binary	Raw binary / hdfeos	Raw binary / hdfeos
N:o scenes	7300	11980	16598	16810	16610	16608	13084	9516
Total size (Gb)	10.9	12.6	14.2	14.3	14.2	17.1	14.6	9.2

Product nomenclature

Filename convention

Product	***Naming convention***
Weekly Snow Water Equivalent Grids	SNOWCARBO_SWE_<year>_<week>_r<resolution>.<extension> <year> gives the year of the product. <week> gives the julian week of the product (1-52). <resolution> gives the grid resolution (0.05 OR 0.25). <extension> is the proper extension for the file depending on file format.
Date of Snow Melt Onset Grids	SNOWCARBO_SMO_<year>_r<resolution>.<extension> <year> gives the year of the product. <resolution> gives the grid resolution (0.05 OR 0.25). <extension> is the proper extension for the file depending on file format.
Date of Snow Clearance Grids	SNOWCARBO_SCL_<year>_r<resolution>.<extension> <year> gives the year of the product. <resolution> gives the grid resolution (0.05 OR 0.25). <extension> is the proper extension for the file depending on file format.
Date of Soil Freezing Grids (TBC)	SNOWCARBO_SFR_<year>_r<resolution>.<extension> <year> gives the year of the product. <resolution> gives the grid resolution (0.05 OR 0.25). <extension> is the proper extension for the file depending on file format.

Product metadata content

Product	***Metadata content***
Weekly Snow Water Equivalent Grids	Data content, field 1: 'Snow Water Equivalent (mm)' Data content, field 2: 'Variance of SWE estimate (mm)' Data date Processing date Coordinate system Latitude range Longitude range Spatial Resolution Processing software name Processing software version Processing organisation Weather station data date Auxiliary data, land mask name Auxiliary data, land mask version Auxiliary data, mountain mask name Auxiliary data, mountain mask version Auxiliary data, forest mask name Auxiliary data, forest mask version

Date of Snow Melt Onset Grids	Data content, field 1: 'Snow Melt Onset Date (julian day)' Data date Processing date Coordinate system Latitude range Longitude range Spatial Resolution Processing software name Processing software version Processing organisation Auxiliary data, land mask name Auxiliary data, land mask version Auxiliary data, mountain mask name Auxiliary data, mountain mask version Auxiliary data, forest mask name Auxiliary data, forest mask version
Date of Snow Clearance Grids	Data content, field 1: 'Snow Clearance Date (julian day)' Data date Processing date Coordinate system Latitude range Longitude range Spatial Resolution Processing software name Processing software version Processing organisation Auxiliary data, land mask name Auxiliary data, land mask version Auxiliary data, mountain mask name Auxiliary data, mountain mask version Auxiliary data, forest mask name Auxiliary data, forest mask version
Date of Soil Freezing Grids (TBC)	Data content, field 1: 'Soil Freezing Date (julian day)' Data date Processing date Coordinate system Latitude range Longitude range Spatial Resolution Processing software name Processing software version Processing organisation Auxiliary data, land mask name Auxiliary data, land mask version Auxiliary data, mountain mask name Auxiliary data, mountain mask version Auxiliary data, forest mask name Auxiliary data, forest mask version

Product description

Weekly Snow Water Equivalent Grids

The Weekly Snow Water Equivalent (SWE) Grids consist of gridded values of SWE and SWE variance over Eurasia. The product is based on applying passive microwave observations and ECMWF weather station observations in an assimilation scheme to produce global maps of SWE estimates (in EASE-Grid format) over the northern hemisphere, 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. As a novel approach, a priori information of snow depth is used to calibrate the model where data is available.

The basis of the processing system is presented in a study by Pulliainen (2006). As applied for GlobSnow, estimates of SD (snow depth) based on emission model inversion of two frequencies, 18.7 and 36.5 GHz, are first calibrated over EASE grid cells with weather station data of SD available. Snow grain size is used in the model as a scalable model input parameter. These values of grain size are used to construct a Kriging interpolated background map of the effective grain size, including an estimate of the effective grain size error. The map is then used as an input in model inversion over the span of available radiometer observations, providing an estimate of SD. In the inversion process, the effective grain size in each grid cell is weighed with its respective error estimate. A snow density value is applied to each grid cell to connect depth to SWE. Areas of wet snow are masked according to observed brightness temperature values using an empirical equation, as model inversion of SD/SWE over areas of wet snow is not feasible due to the saturated brightness temperature response.

The weather station observations of SD are further interpolated to provide a crude estimate of the SD (or SWE) background. The SWE estimate map and SD map from weather station observations are combined using a Bayesian spatial assimilation approach to provide the final product.

The snow emission model applied is the semi-empirical HUT snow emission model (Pulliainen et al., 1999). The model calculates the brightness temperature from a single homogenous snowpack covering frozen ground in the frequency range of 11 to 94 GHz. Input parameters of the model include snowpack depth, density, effective grain size, snow volumetric moisture and temperature. Separate modules account for ground emission and the effect of vegetation and atmosphere. The model has been validated against tower-based and airborne reference radiometer observations (see e.g. Pulliainen et al., 1999, Lemmetyinen et al., 2009). Studies comparing the model to other equivalent radiative transfer models of snowpack emission show good performance for the HUT model against point-scale observations (see e.g. Tedesco et al., 2006). The model can also be applied for direct inversion of snow properties from space-borne passive microwave data (Pulliainen et al., 2001), with better agreement to validation measurements when compared to traditional empirical algorithms.

For the SnowCarbo SWE Grids, the model is applied in two stages; (1) to match model estimates from satellite observations to a priori information of snow depth, using grain size as a fitting parameter and (2) to provide model inversion estimates of SWE over the whole area of available observations.

An example of the product in EASE grid projection is presented in Figure 1.

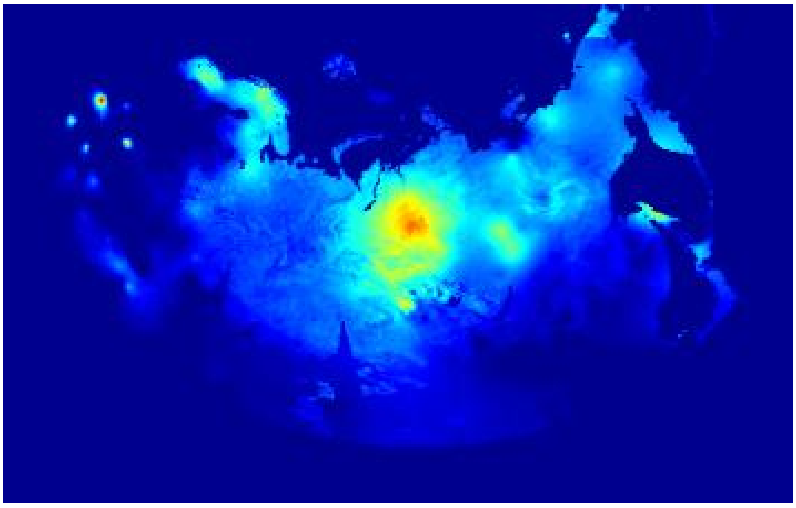

Figure 1 An example of the SWE product (in 0.25 degree EASE-grid) presenting Eurasia hemisphere. The lower limit for the latitude is 35° and the upper limit 85°.

The product source data are EASE gridded AMSR-E and SSM/I brightness temperatures [SD-1], [SD-2]. Although the EASE-grid can represent data almost to the equator the product is limited between latitudes 35° and 85° for physical reasons.

The data type of the product data fields are signed 16-bit integer int16. Positive values and zero are reserved for SWE and negative values for flags. The physical values of SWE are in millimetres. One decimal place is taken into account and the data field value is obtained as follows: SWE_{data}=round(10*SWE_{phys}). For example, if the estimated SWE is 90.234 mm and the corresponding data field value is 902. Negative value -1 means no data, value -2 means water body and value -3 means mountainous area. Summary is presented in Table 2.

Table 2 Legend of the SWE product

Data value	Interpretation
>= 0	$SWE_{phys}=SWE_{data}/10$
-1	No data
-2	Water
-3	Mountains

Date of Snow Clearance Grids

The date of snow melt onset and date of snow clearance are based on a time series analysis of daily SSM/I or AMSR-E brightness temperature observations. The analysis is based on an empirical interpretation of the difference of two passive microwave (radiometer) channels, which allow identification of the snow state (wet/dry snow) and existence of snow (existing snow/snow free ground). The methodology and empirical algorithms forming the basis of the snow state and existence detection are largely reported in literature. The novelty in the SnowCarbo product is the applied time series analysis for snow state and snow existence threshold determination, with increased accuracy compared to direct application of the algorithms. The method is presented by Takala *et al.* (2009).

As a summary, the date of snow clearance is determined as follows:

(1) A pixel wise the time series of the brightness temperature difference of 37 and 19 GHz vertically polarized channels is calculated
(2) The temporal average of the time series with a averaging window of 8 days is calculated
(3) the maximum and minimum values of the averaged time series are identified
(4) A threshold value p for snow clearance (snow free ground) = 90% is identified
(5) Should the averaged time series have value larger than p*(Dmax-Dmin)+Dmin, the vector value is set to 1 (otherwise 0)
(6)Finally, the the last transit from 0->1 is selected and identified as snow clearance date

The Snow Clearance product is formed as a gridded (0.25 degree EASE grid) value corresponding to the date (counted as days from Jan1) in the respective year, when the snow clearance threshold (90%) is passed.

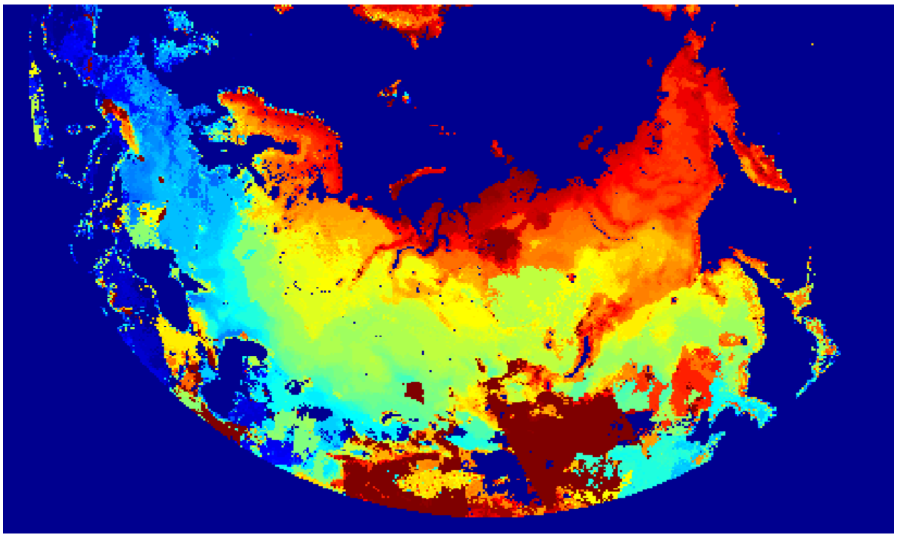

Figure 2 An example of the Date of Snow Clearance product (in 0.25 degree EASE-grid) for Eurasia in 1997. Color codes represent the date of snow clearance from Jan 1.

Table 3 Legend of the Snow Clearance product

Data value	Interpretation
>= 0	Date of Snow Clearance (Julian days from Jan1)
-1	No data
-2	Water
-3	Mountains

Date of Snow Melt Onset Grids

The theoretical and methodological basis of the Snow Melt Onset grids is founded on the same time series analysis and the same raw data as the Date of Snow Clearance. Essentially, the threshold and vector values determined from the timeseries are different. See previous section for description of method.

The Snow Melt Onset product is formed as a gridded (0.25 degree EASE grid) value corresponding to the date (counted as days from Jan1) in the respective year, when the snow melt onset threshold is passed.

Table 4 Legend of the Snow Melt Onset product

Data value	Interpretation
>= 0	Date of Snow Melt onset (Julian days from Jan1)
-1	No data
-2	Water
-3	Mountains

Date of Soil Freezing Grids

The soil freezing grids will be based on available historical and current scatterometer observations. The default instruments will be the NASA QuikSCAT/SeaWinds scatterometer (Ku-band) and Metop ASCAT instrument (C-band).

The final methodology is still under investigation. Possible methods include a time series analysis analogous to the snow melt products, and physical model inversion. A potential physical model for soil freezing detection from active microwave data is presented by Pulliainen *et al.* (1998).

The final format of date of soil freezing grids is still to be determined. It will likely be similar to the snow melt products.

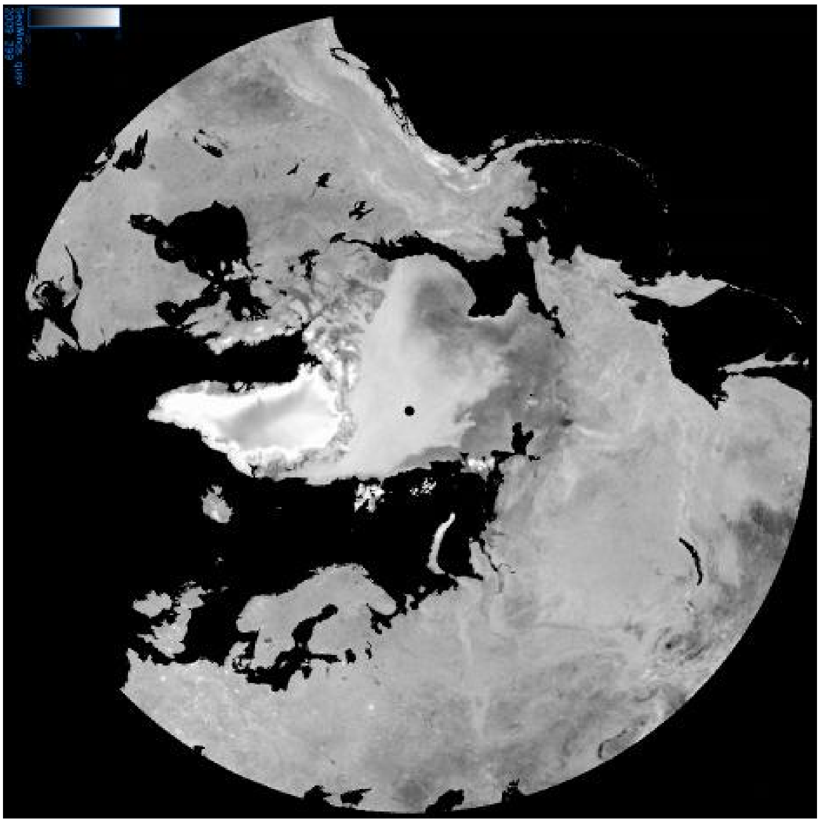

Figure 3 QuikScat backscatter image of Northern Hemisphere.

Data product status

This section summarizes the existing data products and degree of completeness for products still under development.

Table 5 Status of product availability

Product	Raw data available (years)	Product data available	Notes
Weekly Snow Water Equivalent Grids	1978 - present	Prototype dataset: 2000-2008	30-year dataset under production
Date of Snow Melt Onset Grids	1978 - present	-	30-year dataset under production
Date of Snow Clearance Grids	1978 - present	1978 - present	complete
Date of Soil Freezing Grids	1999 - present	-	Methodology under development

Reference Documents

[RD-01] Pulliainen, J., Grandell, J., and Hallikainen, M., 1999. HUT snow emission model and its applicability to snow water equivalent retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 37:1378-1390.

[RD-02] Pulliainen, J., M. T. Hallikainen, 2001. Retrieval of regional snow water equivalent from space-borne passive microwave observations. *Remote Sens. Environ.,*75: 76–85.

[RD-03] 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 Sens. Environ.,* 101: 257-269.

[RD-04] Pulliainen, J.T., Manninen, T., and Hallikainen, M.T., 1998. Application of ERS-1 wind scatterometer data to soil frost and soil moisture monitoring in boreal forest zone. *IEEE Transactions on Geoscience and Remote Sensing*, 36(3): 849 - 863

[RD-05] Tedesco, M., and Kim, E., 2006. Intercomparison of electromagnetic models for passive microwave remote sensing of snow. *IEEE Transactions on Geoscience and Remote Sensing*, 44(10): 2654-2666.

[RD-06] Lemmetyinen J., Derksen C., Pulliainen J., Strapp W., Toose P., Walker A., Tauriainen S., Pihlflyckt J., Kärnä J.-P. and Hallikainen M., 2009. A Comparison of Airborne Microwave Brightness Temperatures and Snowpack Properties across the Boreal Forests of Finland and Western Canada. *IEEE Trans. Geosci. Remote Sensing,* 47: 965-978.

[RD-07] Takala, M., Pulliainen, J., Metsamaki, S.J., and Koskinen, J.T., 2009. Detection of Snowmelt Using Spaceborne Microwave Radiometer Data in Eurasia from 1979 to 2007. *IEEE Transactions on Geoscience and Remote Sensing*, 47(9): 2996 - 3007.

Source data references

[SD-1] Knowles, K. W., M. H. Savoie, R. L. Armstrong, and M. J. Brodzik. 2009. *AMSR-E/Aqua daily EASE-Grid brightness temperatures*, 2003-2009. Boulder, Colorado USA: National Snow and Ice Data Center. Digital media.

[SD-2] Armstrong, R. L., K. W. Knowles, M. J. Brodzik and M. A. Hardman. 2009, *DMSP SSM/I Pathfinder daily EASE-Grid brightness temperatures*, 1978-2002. Boulder, Colorado USA: National Snow and Ice Data Center. Digital media.

SYKE EO Dataset descriptions (Action 3)

Introduction

For the purposes of the SnowCarbo project, SYKE produces an extended and harmonized dataset of GMES Service elements GSE Land and GSE PolarView for the generation of time series of intra-annual green vegetation status and snow cover. These products are derived form the observations of the MODIS Terra instrument. The spatial and temporal coverage of the GSE products is extended to the specific needs of this project. In addition, surface brightness temperatures are retrieved using NOAA AVHRR data.

The datasets described in this document are

- AVHRR Data
  - Raw AVHRR dataset
  - Calibrated AVHRRR Dataset
- MODIS Data
  - RAW MODIS dataset
  - Calibrated MODIS dataset
  - EO products derived from the calibrated MODIS dataset

Area of Interest (AOI)

All data is processed to the area of "Baltic EU countries" with latitude ranging from 45 to 71 degrees and longitude from 7 to 45 degrees.

Data grid

All EO instrument data is processed in their nominal resolutions and the final products are aggregated to coarser resolution as required. For MODIS and AVHRR data grids of different resolutions (0.0025, 0.005 and 0.01 degrees) are aligned in a way that cells of different resolutions overlap each other exactly and for all cell sizes the upper-left corner of the top-left pixel is exactly at 71,7 (lat, lon). Thus, each 0.01 cell covers the exact area of 4x4 group of 0.0025 cells (Figure 4)

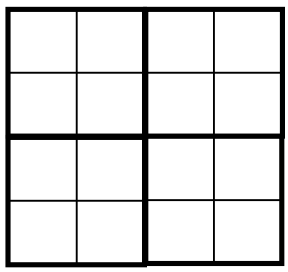

Figure 4 Cell alignment of different resolutions. One 0.01 degree cell consists of four 0.005 degree cells

Temporal coverage

Daily products are generated from February to October for years 2001 to 2010. As only optical instruments are used, fully cloudy scenes are excluded form the dataset.

AVHRR Data

Summary

NOAA AVHRR data is used to retrieve the top of atmosphere brightness temperatures. Processing of the AVHRR dataset includes data retrieval, calibration and rectification. In addition to the calibrated TOA brightness temperatures, cloud masks for the selected image are generated.

Processing steps

The processing steps of AVHRR data are described in Table 6.

Table 6 Processing steps of the AVHRR data

Processing Step	Description	Software
Raw data archiving	AVHRR data files are transferred to the dedicated archive tree with a graphical user interface. Additionally so-called lck-files that are used by the NAPS-processing software are generated for each individual scene. Also software for selecting the images based on measurement time is designed.	Matlab, Python
Data unpacking	Data format change for processing purposes	NAPS, hrpt.exe
Calibration	Calibration of the raw instrument data to the TOA reflectance and TOA brightness temperature.	NAPS, avhrr_cal.exe, avhrr_cor.exe
Rectification	Projection of the calibrated dataset to geodetic projection (WGS84) and subsetting according to the project AOI. Output format is Erdas Imagine (.img)	NAPS, avhrr_geo.exe, grid_to_image.exe
Computation of the cloud masks	Cloud masks are computed using threshold values for brightness temperatures.	Matlab
Preprocessed data archiving	Preprocessed data is transferred to archive tree with search and metadata capabilities	Matlab, Python
Brightness temperature retrieval	On-line extraction of the TOA brightness temperature values from the calibrated .img files	Matlab

Raw Data

Description

Local data archive is generated by combining the SYKE's operational NOAA AVHR archive and historical AVHRR data obtained from the off-line archives of the FMI. AVHRR raw data include files in *.dun* and *.hmf* formats. Instrument track parameters are given in the telex-files (.tlx) in raw ASCIII format.

Data sources

For operative purposes SYKE receives NOAA AVHRR data daily trough ftp from the Sodankylä receiving station. Off-line data is stored in a CD/DVD archive that is copied to RAID disk arrangement for processing in SYKE. Raw data file nomenclature is described in Table 7. For project data processing purposes, the whole CD and on-line dataset available in FMI had to be restored in external hard drives and reorganized. A metadata/search system according to SYKE processing system data format standards was created for it.

Table 7 AVHRR raw data nomenclature.

Name [YYYY][JUL|MMDD][HH][MM]_NOAA[II]_[FMT].[FMT][.gz]		
Field	**Signification**	**Value**
[YYYY]	Year in 4 digits	2000 – 2010
[JUL| MMDD]	Julian date in 3 digits or month and day values in 4 digits	001 – 356 02-12, 01-31
[HH][MM]	Hours and minutes in 2 digits	01-23, 01-59
[II]	Instrument number	14-19
[FMT]	Data format	dun|hmf
[.gz] | [.gz2]	Extension of the compressed files	(optional)

Data storage and access

Data is organized in a yearly/monthly directory structure on duplicated external hard drives. Archive dataset is available for project partners upon request. Summary of the current status of the NOAA AVHRR archive is given in Table 8.

Table 8 Summary of the collected AVHRR Raw data archive

	2001	2002	2003	2004	2005	2006	2007	2008
Format	dun	dun	dun	hmf	hmf	hmf	hmf	hmf
N:o scenes	3301	3491	4220	4436	4607	5810	5204	4624
Total size (Gb)	110	110	113	196	196	240	271	238

Preprocessed data

Description

AVHRR data is calibrated and rectified using a software suite designed for the preprocessing of AVHRR data (1). Used programs include *hrpt* for data unpacking, *avhrr_cal* and *avhrr_cor* for data calibration, *avhrr_geo* for geocoding,m and *correct_grid* and *grid_to_image* for extraction of the instrument and solar angel data. Processing is done using SYKE's in-house processing software NAPS. Dataset processing is currently being done.

Temperature data retrieval

TOA brightness temperatures can be directly obtained from the calibrated .img files. Software for extraction of the temperature data is written in Matlab. TOA temperature data is directly used to assess the maximum ground temperature potentially available in the night time imagery.

MODIS Data

Summary

Collected and processed dataset can be divided into three main categories:

1. Raw EO data, which currently includes the unprocessed files form years 2001 to 2008. The dataset are pre-selected by the operator by excluding totally cloudy scenes.
2. Preprocessed EO data, including the full set of calibrated bands of the instrument. Calibrated bands are further rectified to their nominal resolutions and cropped to the area of interest of the SnowCarbo project.
3. Product data computed from the preprocessed bands, including green vegetation status, snow covered area information and the winter and summertime cloud masks.
4. Daily and weekly composites of the product data.

Processing steps

Processing of the MODIS dataset is described in Table 9

Table 9 MODIS data processing steps

Processing Step	Description	Software
Raw data archiving	Operator-selected MODIS datasets are inspected and transferred to the dedicated archive tree with a graphical user interface. Additionally so-called lck-files that are used by the NAPS-processing software are generated for each individual scene.	Matlab
Data unpacking	Data format change for processing purposes	NAPS, Envimon
Calibration	Calibration of the raw instrument data to the TOA reflectances. Observations with solar zenith angle greater than 71 degrees and/or instrument viewing angle greater than 60 degrees are marked as no-data.	NAPS, Envimon
Rectification	Projection of the calibrated dataset to geodetic projection (WGS84) and subsetting according to the project AOI. Output format is Erdas Imagine (.img)	NAPS, Envimon
Preprocessed data archiving	Preprocessed data is transferred to archive tree with a graphical user interface.	Matlab
NDVI calculation	NDVI algorithm is applied to QKM data and the results are saved as Matlab mat-files.	Matlab
SCA Calculation	SCA algorithm is applied to HKM data and results are saved as Matlab mat-files. Observations with solar zenith angle greater than 63 degrees are marked as no-data	Matlab
Cloud mask calculation	Dedicated algorithms are used to generate cloud masks for winter- and summertime images. Results are saved as 1-bit tiff files with georeferencing information in accompanying .tfw files	Matlab
Daily compositing of NDVI	Daily scenes are composited by applying the cloud masks to single scenes and in case of multiple observations of single pixel, selecting the maximum value of NDVI.	Matlab
Daily compositing of SCA	Daily scenes are composited by applying the cloud masks and in case of multiple observations of single pixel selecting the mean value of the SCA. If the observed values differ by	Matlab

	more than 25 percentage units, pixels is marked as no-data.	
Weekly compositing of NDVI & SCA	For each calendar week the daily products are composited using the same criteria as for the daily composites	Matlab

Raw Data

Description

Local data archive contains daily MODIS Level 1B data from February to October for years 2001 – 2008. Images are operator-selected by checking the coverage with respect to the project's area of interest and excluding totally cloudy scenes from the dataset. Two different sources of raw images are used, namely NASA's LAADS archive and SYKE's existing internal MODIS data archive that contains Level 1B mosaics of the AOI, processed by FMI (Sodankylä receiving station) . In addition to the difference in the spatial extent of these two data sources, the data format is different. LAADS data is in hdfeos format and FMI data is in hdf4-format. Starting from year 2007, LAADS-data is used only to fill in the dates that are missing from the SYKE's operative archives. Summary of the current status of the archive is given in Table 11.

Data Sources

LAADS Data

MODIS Level 1 B datasets are previewed, selected and ordered using the "LAADS Web" service (http://ladsweb.nascom.nasa.gov/index.html) and orders are further downloaded to SYKE trough http-protocol using wget. Because of the swath size, typically 2-4 scenes/day are required to get full coverage of the AOI.

A single scene consists of 4 separate files (detailed descriptions are available at http://ladsweb.nascom.nasa.gov/filespecs/):
MOD021KM – Level 1B 1km product
MOD02HKM – Level 1B 500m product
MOD02QKM – Level 1B 250m product
MOD03 – Geolocation file

FMI data

FMI Sodankylä receiving station processes MODIS Level 1B swath data according to the specific needs of PolarView-project. Process combines the separate swaths to single file covering the PolarView project's AOI. Process takes into account the overlapping field of views, different calibrations of the separate swaths and the metadata within the combined files. The resulting dataset is automatically transferred to SYKE's ftp-server daily.

Nomenclature

Table 10 Nomenclature of the raw data files.

LAADS	FMI
MOD021KM.A[DATE].005.[PROCID].hdf	1km_calibrated_ter_ILFMI_dd-mmm-YYYY_HH_MM_00.hdf
MOD02HKM.A[DATE].005.[PROCID].hdf	hkm_calibrated_ter_ILFMI_dd-mmm-YYYY_HH_MM_00.hdf
MOD02QKM.A[DATE].005.[PROCID].hdf	qkm_calibrated_ter_ILFMI_dd-mmm-YYYY_HH_MM_00.hdf
MOD03.A[DATE].005.[PROCID].hdf	geolocated_ter_ILFMI_dd-mmm-YYYY_HH_MM_00.hdf
Where: [DATE] = YYYYDDD.HHMM [PROCID] = YYYYDDDHHMMSS [YYYY] = Year in 4 digits [DDD] = Julian date	[dd] = day of month [mmm] = abbreviation of the month (English) [HH],[MM],[SS] = Hours, minutes and seconds

Data Storage and Access

Summary of the current status of the MODIS L1B archive is given in Table 11. Data is organized in a yearly/monthly directory structure on duplicated external hard drives. Archive dataset is available for project partners upon request.

Table 11 Summary of the collected MODIS Raw data archive

	2001	2002	2003	2004	2005	2006	2007	2008
Format	hdf4	hdf4	hdf4	hdf4	hdf4	hdf4	hdf4 / hdfeos	hdf4 / hdfeos
N:o scenes	440	359	355	589	428	451	312	387
Total size (Gb)	223	183	180	302	219	230	311	359

Preprocessed data

Description

SYKE's in-house processing software NAPS (Nearly Antonymous Processing System) is used to schedule and execute preprocessing tasks. NAPS controls the execution of dedicated EO-data processing software and takes care of data transfer, logging and distributing the results to archive. Process parameters, scheduling and logs can be controlled by a web-interface on the server running NAPS.

EO-data preprocessing software Envimon (2) is used to unpack, calibrate and rectify MODIS Level 1B data. Resulting data files are listed in Table 17. Data is stored in Erdas imagine .img – format.

Radiometric calibration is done separately for IR/NIR bands and for the thermal bands. No atmospheric calibration is done at this stage, but sun angle correction is taken into account.

Geometric correction and projection is done for the calibrated files. Geodetic projection (datum WGS84) is used for all output files. Also satellite and solar zenith and azimuth angles are extracted and rectified. Nomenclature of the resulting files is described in Table 12 and detailed description of the data files in Table 16 and Table 17.

Nomenclature

Table 12 Nomenclature of the preprocessed files

Name	*MOD_[RES]_[DATE]_[DATATYPE]_rect.img*		
Field	**Subfields**	**Signification**	**Value**
RES		Input data resolution	QKM = 0.00025 deg HKM-=0.005 deg 1KM = 0.01 deg LL = 0.01 deg
DATE		Timestamp of the observation	[YYMMMDD_HHMM]
	YY	Year with 2 digits	01 – 10
	MMM	Month abbreviation in English	jan, feb,.., dec
	DD	Day of month with 2 digits	01 – 31
	HH	Hour with 2 digits	00-23
	MM	Minutes with 2 digits	00-59
DATATYPE		Dataset identifier	See below
		Sun and satellite zenith and azimuth angles	[Solar|Sensor][Zenith|Azimuth]
		Emissive channel data	EV_1KM_Emissive_kelvin
		TOA Calibrated channel data	EV_[RES]_RefSB_toa
_rect.img		Rectified data flag	

Data Storage and Access

Preprocessed files are archived for later use, enabling for example quick re-computation of end products and development of new algorithms. Datasets are stored in a directory tree, organized according to EO instrument, calibration and data type and time of the observation. Preprocessed data are stored on duplicated external hard drives. Dataset is available for project partners upon request.

Table 13 Summary of the preprocessed data archive

	2001	2002	2003	2004	2005	2006	2007	2008	
Number of scenes	437	359	355	587	428	451	322	358	
Size (GB)	517	439	436	686	520	541	511	555	

EO Product Data

Description

For the purposes of carbon balance mapping, dedicated data products are generated from the calibrated MODIS data. The products include NDVI for green vegetation status and fractional snow covered area SCA information.

NDVI (Normalized differential vegetation index) is computed using MODIS QKM bands 1 and 2 with equation.

$$NDVI = \frac{ch_1 - ch_2}{ch_1 + ch_2}$$

Produced NDVI estimate corresponds to the GSELand NDVI product.

SCA Fractional snow cover is estimated using SYKE's operational snow cover algorithm that is described in the documents (3) and (4). Produced estimate corresponds to the product of GSE PolarView project.

Cloud masks are generated using methods developed at SYKE for operational production of SCA and NDVI estimates for GSE Land and GSE PolarView projects.

Additionally for each product a 8bit indexed qeotiff quicklook is generated. This cannot be used to retrieve the actual estimate value but is used for quick visual inspection only.

For more flexible handling of the dataset, both products are initially stored as raw output of the algorithm, meaning that land, water and cloud classification as well as the removal of unlikely values is done in the later parts of the product processing chain. This allows, for example, re-computation of the cloud masks without the need to recalculate the product estimates.

All products are computed for each individual scene and thereafter daily and weekly composites of the products are generated. Details of the composition are described in section SYKE composite products and time-series data.

Nomenclature

Names of the EO product files are derived directly from the preprocessed files that are used in the computation.

Table 14 NDVI product details

NDVI	Name	Value range	Masks	Datatype
Single Scene	Modis_qkm_[YYMMMDD_HHMM]_EV_250_RefSB_toa_rect_ndvi.mat	unlimited	none	float32
Daily	Modis_qkm_[YYMMDD]_EV_250_RefSB_toa_rect_ndvi_max.mat	[0-1]	Clouds	float32
Weekly	Modis_qkm_[YYYY]wk[WW]_EV_250_RefSB_toa_rect_ndvi_max.mat	[0-1]	Clouds	float32

Table 15 SCA product details

SCA	Name	Value Range	Masks	Datatype
Single Scene	Modis_hkm_[YYMMMDD_HHMM]_EV_250_RefSB_toa_rect_ndvi.mat	unlimited	none	float32
Daily	Modis_hkm_[YYMMDD]_EV_250_RefSB_toa_rect_sca_mean.mat	[0-105]	Clouds	float32
Weekly	Modis_hkm_[YYYY]wk[WW]_EV_250_RefSB_toa_rect_sca_mean.mat	[0-105]	Clouds	float32

Data Storage and Access

EO product data is stored on duplicated external hard drives and on the shared network folders in SYKE's internal use. The dataset is available for project partners upon request.

SYKE composite products and time-series data (Action 7)

Description

Computed NDVI, SCA and brightness temperature products are further composited to daily and weekly products which are the "raw time series"-deliverable of Action 7 and are further analyzed under that action.

For MODIS dataset, there are typically 1-4 separate, partly overlapping observations over the project's AOI. The EO product data is generated for each of these scenes individually and then composited to obtain single estimate for each day.

Resulting composite products have cloudy areas flagged with NaN's and the data range has been reduced to the valid range of the corresponding product.

Nomenclature and data formats of the produced product estimates are described in Table 14 and Table 15.

NDVI Composites

NDVI estimates of single date are composited by first applying the individual cloud masks to the separate estimates and, in case of multiple observations, selecting the maximum of the observed NDVI values for each individual cloud-free pixel.

SCA Composites

SCA estimates of single date are composited by first applying the individual cloud masks to the separate estimates and, in case of multiple observations and when the absolute difference between the observations is less than 25 percentage units, computing the mean of the observations for each individual cloud free pixel. For differences larger than 25 percentage units, pixels are marked as no-data.

Weekly NDVI and SCA composites

For MODIS products, the weekly composites of each product are generated by applying the same method as in case of daily composite.

Brightness temperatures

Since only one scene of AVHRR data is needed to get the full coverage of the projects AOI, daily composites are not required for the brightness temperature product. Weekly composites are generated by first applying the daily cloud masks and, in case of multiple observations, selecting the maximum of the observed night-time brightness temperature values for each individual cloud-free pixel.

Data Storage and Access

Composite data of NDVI and SCA is stored on shared network folders as Matlab mat-files. Dataset is available for project partners upon request.

Reference Documents

1. **Andersson, Kaj.** *NOAA AVHRR Data processing software user's guide (version 3.0).* s.l. : VTT Automation, 2001.
2. —. *Envimon pre-processing software user's guide (Version 2.08)* . s.l. : VTT Information Technology, 2009.
3. *A Comparison of Finnish SCAmod Snow Maps and MODIS Snow Maps in Boreal Forests in Finland and in Manitoba, Canada.* **Anttila, Saku, Metsämäki, Sari and Derksen, C.** Denver, Colorado : s.n., 2006. Proceedings of IEEE 2006 International Geoscience and Remote Sensing Symposium (IGARSS'06). July 31- August 4.
4. *A feasible method for fractional snow cover mapping in boreal zone based on a reflectance model.* **Metsämäki, Sari, et al.** 1, 2005, Remote Sensing of Environment, Vol. 95, pp. 77-95.

Table 16 MODIS bands in preprocessed files

MODIS Band combinations/bandwidths in SYKE's ERDAS Imagine files

[DATATYPE] - tag in filename	250_RefSB_toa		500_RefSB_toa		1km_RefSB_toa		1km_Emissive_kelvin	
Band number in img-file	MODIS n:o	(nm)	MODIS n:o	(nm)	MODIS n:o	(nm)	MODIS N:o	(um)
1	1	620 - 670	3	459 - 479	8	405 - 420	20	3.660 - 3.840
2	2	841 - 876	4	545 - 565	9	438 - 448	21	3.929 - 3.989
3			5	1230 - 1250	10	483 - 493	22	3.929 - 3.989
4			6	1628 - 1652	11	526 - 536	23	4.020 - 4.080
5			7	2105 - 2155	12	546 - 556	24	4.433 - 4.498
6					13lo	662 - 672	25	4.482 - 4.549
7					13hi	n/a	27	6.535 - 6.895
8					14lo	673 - 683	28	7.175 - 7.475
9					14hi	n/a	29	8.400 - 8.700
10					15	673 - 683	30	9.580 - 9.880
11					16	862 - 877	31	10.780 - 11.280
12					17	890 - 920	32	11.770 - 12.270
13					18	931 - 941	33	13.185 - 13.485
14					19	915 - 965	34	13.485 - 13.785
15					26	1360 - 1390	35	13.785 - 14.085
16							36	14.085 - 14.385

Table 17 Preprocessed MODIS data, [DATE] is in format YYMMMDD_HHMM

File name	Resolution (deg)	N:o bands	MODIS Band numbers	Datatype	Description
Modis_qkm_[DATE]_EV_250_RefSB_toa_rect.img	0.00025	2	1 - 2	float32	TOA reflectance with sun angle correction (%)
Modis_hkm_[DATE]_EV_500_RefSB_toa_rect.img	0.005	5	3 - 7	float32	TOA reflectance with sun angle correction (%)
Modis_1km_[DATE]_EV_1KM_RefSB_toa_rect.img	0.01	15	8 – 19, 26	float32	TOA reflectance with sun angle correction (%)
Modis_1KM_[DATE]_EV_1KM_Emissive_kelvin_rect .img	0.01	16	20-36, excl. 26	float32	Thermal data (Kelvin degrees)
Modis_LL_[DATE]_SolarZenith_rect.img	0.01	1	n/a	int16	Solar zenith angle (degrees*100)
Modis_LL_[DATE]_SolarAzimuth_rect.img	0.01	1	n/a	int16	Solar azimuth angle (degrees*100)
Modis_LL_Y[DATE]_SensorZenith_rect.img	0.01	1	n/a	int16	Satellite zenith angle (degrees*100)
Modis_LL_[DATE]_SensorAzimuth_rect.img	0.01	1	n/a	int16	Satellite azimuth angle (degrees*100)
Modis_[DATE]_NAPS.xml	n/a	n/a	n/a	ASCII-text	NAPS-process metadata including processing parameters etc.
Modis_QKM_[DATE]_EV_250_RefSB_toa_rect_logs. zip	n/a	n/a	n/a	ASCII-text	Compressed collection of preprocessing log files