Ocean and Sea Ice SAF

Algorithm Theoretical Basis Document for OSI SAF Medium Resolution Sea Ice Drift Product

OSI-407

Version 1.1 - May 2011

Gorm Dybkjaer

Document change record:

Document Version	Date	Author	Description
v0.1	May 28th 2009	GD	Submitted to review
v1.0	July 8th 2009	GD	Review for PCR
v1.1	May 4th	GD	Edition for preop

Table of contents:

Glossary

AAPP - **A**TOVS and **A**VHRR **P**re-processing **P**ackage

AHA - A file format for gridded satellite data, designed at Swedish Met. and Hydro.Inst.

Argos - worldwide location and data collection system

ATBD - Algorithm Theoretical Basis Document (This document)

AVHRR – Advanced Very High Resolution Radiometer

CDOP – Continuous Development and Operations Phase

DAMAP – A common DMI/Met.no software package for processing satellite data

DMI – Danish Meteorological Institute

EPS - EUMETSAT Polar System. The European comp. of a joint Europ./US polar satellite system.

EUMETCast - EUMETSAT's Broadcast System for Environmental Data

EUMETSAT - European Organisation for the Exploitation of Meteorological Satellites

GTS - Global Telecommunication System

ICEDRIFT-GRID - The fixed 20km grid in which the final ice drift product is delivered.

INPUT-GRID – The fixed 1km grid in polar steroid projection containing the input data, either the *IR*- or the *VIS*, from the AVHRR instrument.

IR - Infra Red

KAI – A EUMETSAT tool for processing EPS PFS format products

MCC – Maximum Cross Correlation

met.no – Norwegian Meteorological Institute

Metop – EUMETSAT OPerational METeorological polar orbiting satellite

NETCDF – A file format (network Common Data Form)

NH - Northern Hemisphere

NOAA - National Oceanic and Atmospheric Administration

NWP-SAF – The Numerical Weather Prediction SAF

OSISAF – Ocean and Sea Ice Satellite Application Facilities

PROJ4 – A cartographic projection library

PUM – Product User Manual

RMS – Root Mean Square

SAR – Synthetic aperture radar

SSM/I - Special Sensor Microwave/Imager

VIS – visible

1. Introduction

This product of the Ocean and Sea Ice - Satellite Application Facility (OSI-SAF) estimates sea ice displacement during a period between two satellite data swath that are separated by approximately 24 hours. The information available in the data set is geographical positions of sea ice, at the beginning and at the end of the 24h period. As the data contains no information of the path of the sea ice, the product can NOT be interpreted as a mean sea ice speed estimate for this period.

The product is a gridded subset of the Northern Hemisphere (NH) covering the full OSISAF-NH area (see figure 1). Due to the nature of the input data, each ice drift data set only contains valid ice drift data for parts of the output grid, and dummy values are filled in where no data are calculated.

The justification or strength of this product lies in the relative high temporal and spatial resolution, as well as in the relative short ice drift period, in contrast to fully gridded data sets with longer ice drift periods and coarser spatial resolution. The resolution of this product makes it suited for calibration, and validation purposes as well as for data assimilation.

The product is therefore aiming at modellers dealing with integrated sea-ice-atmosphere models and users dealing with merging of ice drift data sets, for both calibration validation purposed..

1. Scope of this document

This Algorithm Theoretical Basis Document is describing the computational steps implemented for this Medium Resolution Sea Ice Drift processing software, which aspires to run as part of the EUMETSAT OSISAF programme. The document introduces and, to some extent, give justification for the scientific assumptions and choices made, that has led to present near-real-time sea ice motion processing in the EUMETSAT OSI SAF.

User related aspects of the product (like file format and output specifications) are rather to be found in the Product User's Manual [PUM]. Results from validation against ground truth sea ice drift measurements will be gathered in an associated validation report which is in progress [VAL].

General information on the EUMETSAT OSI SAF is available from the OSI-SAF official web site (www.osi-saf.org).

After introducing some product specific notation in the remaining of current chapter, grid information, input data and processing steps are described in chapter 2. The algorithm behind the motion tracking algorithm is described chapter 3. The Validation strategy is explained in chapter 4. Finally, in chapter 5, the assumptions and limitations are discussed.

2. Common notations

A few product specific notations will be used throughout this document and to ease further reading the most central ones are explained here.

The main input data source for this ice drift detection procedure is thermal infrared data (***IR***) from the Advanced High Resolution Radiometer (AVHRR) on board the Metop platform. During the Arctic summer also visible data (***VIS***) from the same instrument is used as input data.

The applied ice drift detection technique is based on feature recognition from one satellite swath data set to another. Here the data set recorded at time *T* is called the ***reference*** data, and the another data set used for feature comparison, recorded at time *T* + 24h, is called the ***compare*** data.

Two different data grids will be mentioned throughout this document. One grid is the ***input-grid***, which is a fixed 1km grid in polar steroid projection (table 1), containing either the *IR*- or the *VIS* data, i.e. brightness-temperature or albedo data.

The other grid is the product output grid, the ***icedrift-grid***, for which the ice drift data are calculated. The *icedrift-grid* is a 20 km resolution grid (table 2), in the same projection as the *input-grid*.

2. Data and data handling

1. Input data

The input data used for this application is retrieved from the Advanced Very High Resolution Radiometer (AVHRR) instrument on board the polar orbiting Metop satellite. AVHRR instruments have been operating from polar orbiting satellites since the late 1970'ies on board NOAA satellites, and newer versions of the instruments carry 6 spectral bands, 3 in the visible spectrum and 3 in the near Infra Red spectrum. Present application mainly operates with IR data, but during the Artic summer also a visible band is deployed. The bands applied as visible and infra red data are channels 2 and 4, with central wave length of 0.86 and 10.8 microns, respectively.

The spatial resolution of the original input data is approximately 1.1 km at nadir and the scan width is 2045 pixels, providing a swath width of approximately 2800km. The accuracy of the geographical rectification is assumed to be sub pixel [EUMETSAT2007], with a maximum swath angle of ±56 degrees. This result in geographical accuracies between approximately less than 1 km at nadir and less than 2 km, at the edge of the image (see chapter 5). The Metop satellite is sun synchronous, meaning that 2 swath separated by 24h more or less cover the same area (figure 1 - right).

To avoid running the ice drift procedure for areas with no sea ice a sea ice mask is applied in the processing chain. The sea ice mask is the OSISAF ice type product [Andersen2006] valid for the day of the *reference* data set.

2. Grid info

The two grids used in this procedure cover the geographical area OSISAF NH [Andersen2007], illustrated in figure 1. The cartographic projection tool, PROJ.4 [PROJ4], is used to transform the grid positions in the NH subset into geographical coordinates and vice versa. Exact grid projection characteristics and grid size of the *input-grid* and *icedrift-grid* are given in tables 1 and 2, respectively.

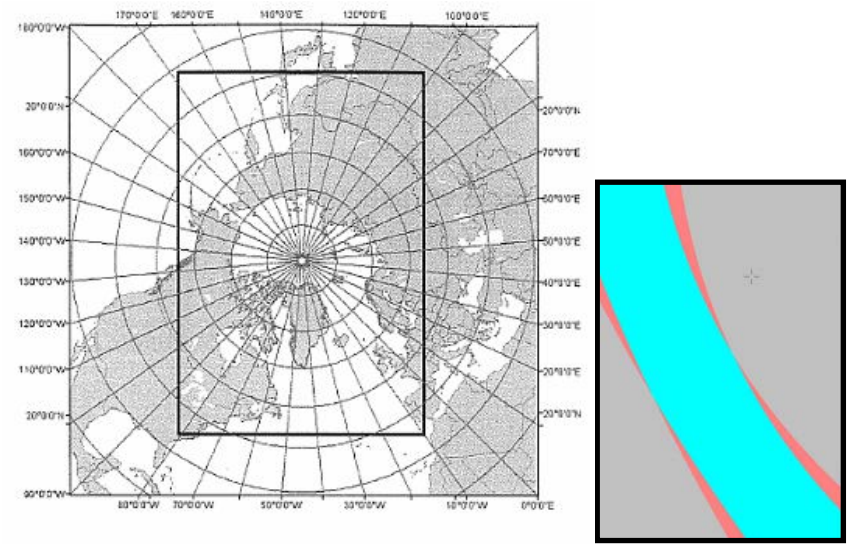

Figure 1 Left image: The OSISAF NH-subset outlined with bold rectangle. Right image: The blue area illustrates the overlap between two input data sets separated by 24 hours for the NH area. Specifications of the grids are given in tables 1 and 2.

Table 1 Geographical definition of the *input-grid*.

Projection	Polar stereographic projection with true scale at 70°N	
Resolution	1 km	
Size	7600	11200
Central Meridian	45°W	
Corner points UL (dec.degr.)	32.655N	169.160E
Corner points UL (m)	U = -3800000	V = 5600000
Earth axis	a=6378273	b=6356889.44891
PROJ4-string	+proj=stere +a=6378273 +b=6356889.44891 +lat_0=90 +lat_ts=70 +lon_0=-45	

Table 2 Geographical definition of the *icedrift-grid*.

Projection	Polar stereographic projection with true scale at 70°N	
Resolution	20 km	
Size	379	559
Central Meridian	45°W	
Corner points UL (dec.degr.)	32.854N	169.114E
Corner points UL (m)	U = -3780000	V = 5580000
Earth axis	a=6378273	b=6356889.44891
PROJ4-string	+proj=stere +a=6378273 +b=6356889.44891 +lat_0=90 +lat_ts=70 +lon_0=-45	

In the product user manual [PUM] the relations between grid coordinates and geographical coordinates are described.

3. Processing chain and Timeliness

Various processing steps are involved in the ice drift production, from pre-processing procedures to final validation procedures and dissemination. Below these steps are ordered chronologically and processing times are estimated. Processing time of the actual ice drift estimation procedure can be shortened significantly if it is decided to parallelize the process to be able to run on more computing threads.

2h and 15min. EUMETCast: Timeliness of global Metop data level 1b via EUMETCast. The data comes in 3-minute segments in satellite swath projection in native EUMETSAT file format EPS.

0h and 20min. DMI: Concatenation of EPS-segments using the KAI programme and waiting for delayed segments.

0h and 10min. DMI: Running conversion program (convert_avh1b.exe) for converting the concatenated EPS swath data into AAPP level 1b format in swath projection. Where AAPP is **A**TOVS and **A**VHRR **P**re-processing **P**ackage, a part of NWP-SAF.

0h and 10min. DMI: Running a file-format conversion programme to transform the non-rectified AAPP swath data into to the aha-satellite data format (a DAMAP-tool). A format designed to hold rectified and gridded satellite data. In this conversion the data are placed into the fixed grid constituting the NH-subset of OSISAF (see figure 1).

0h and 10min. DMI: Converting the aha-file into netCDF, the input format for the ice drift detection routine.

2h and 30 min. OSISAF: Running the full ice drift estimation programme including filtering. This production time is depending on the sea ice extent and the given time estimate is valid for normal winter coverage and for a non-parallelized cpu usage.

0h and 20 min. OSISAF: verification and dissemination.

The ice drift product will thus be ready and distributed less than 6 hours after the instrument time of the last recorded segment of the *compare* swath.

3. Algorithm Description

Various setup's of Maximum Cross Correlation (MCC) techniques are acknowledged and applied in many feature tracking programmes, not the least in programmes that keep track of drifting sea ice [Haarpaintner2006][Maslanik1998][Ezraty2006]. The technique is relatively strait forward to apply to gridded satellite data, and it is a relative robust method. Moreover, the method is based on sensible assumptions for ice drift tracking (See chapter 5). There is no obviously better solution than the MCC technique for ice drift detection, hence the MCC technique is chosen for this application.

1. Algorithm principle

The applied MCC algorithm is a relative simple pattern tracking technique that performs a section-wise matching of geographical distributed data recorded at time T (*reference* data, figure 2) with data recorded 24h later, at time T + 24h (*compare* data, figure 2). The best match, as measured by the highest correlation, between *reference* data and a sub-image/section of the *compare* data determines the ice drift for a given grid point.

For each point in the *input-grid* separated by 20 km, an ice drift vector is attempted retrieved by the iterative best matching routine sketched in figure 2, provided the icedrift-grid point is classified as sea ice according to the applied ice mask (see section 2.1). A matrix around each *icedrift-grid* point is correlated, to any corresponding matrix in the *reference* data that is inside the maximum allowed distance from origin in the *compare* data set, i.e. inside the red circle in figure 2. The "maximum allowed distance from an *icedrift-grid* point" is determined from a maximum allowed ice drift 'speed' multiplied with the time between the *reference* and the *compare* data sets.

The ice drift associated with a given *icedrift-grid* point from time T to T+24h is hence the geographical shift between the *icedrift-grid* point in the *compare* data set and the centre of the best matching matrix in the *reference* data set.

Despite the fact that the input data are highly sensitive to clouds, the production does not use any cloud screening procedures, instead a post MCC filtering routine is applied to remove erroneous data, i.e. ice drift vectors that are not coherent with its neighbourhood will be removed. (See section 3.2).

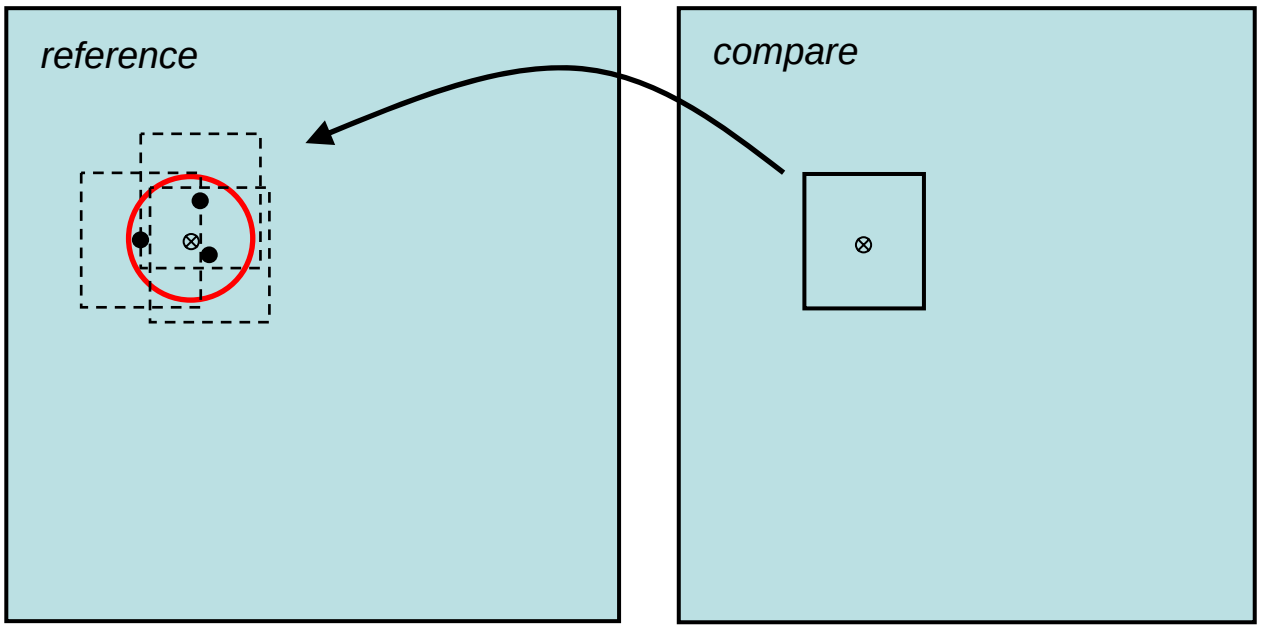

Figure 2 Sketch of the feature tracking procedure. Bold square in *compare* data illustrates the correlation matrix around the *icedrift-grid* point of interest (small circle with cross in *reference* and *compare*). Red circle in *reference* data correspond to the maximum allowed drift distance between the *reference* and *compare* data sets. The three punctured squares, with associate centres (black dots), illustrates 3 possible best matches (or maximum correlation matrices) to the compare matrix.

2. Filtering ice-drift vectors

Most ice drift estimation routines are associated with filtering routines to remove erroneous ice drift vectors. In this setup no cloud screening procedure is implemented, despite the fact that the input data are very sensitive to atmospheric properties. This consequently produce more erroneous ice drift vectors than routines based on micro wave data, that are much less sensitive to atmospheric opacity than IR and VIS data.

The reason for not applying cloud screening here is that cloud screening in the Arctic is rather dubious, due to comparable properties of cloud and snow/ice surfaces in the VIS and IR spectrum. Therefore, it is decided to ignore the presence of clouds and alternatively to run a comprehensive filter routine for erroneous ice drift vectors after the MCC routine. Whenever an effective cloud screening procedure is available for real time use, this will off cause be implemented in this ice drift procedure. That will save time in the MCC procedure.

Obvious error vectors are recognised by having an abnormal absolute drift compared to neighbouring ice drift vectors or a bearing that is uncorrelated to its neighbours. Figure 3 is shown an example of an ice drift product before and after applying the filter.

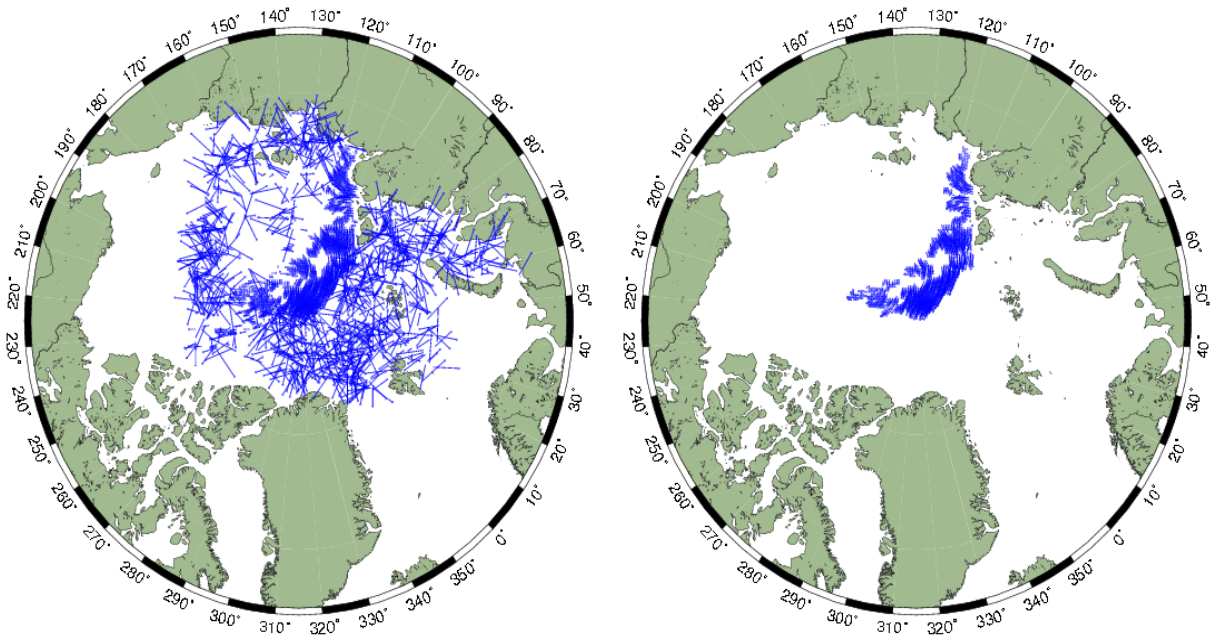

Figure 3 Example of ice drift estimation before applying filter (left) and after filtering of 'obvious' erroneous ice drift vectors (right). The lengths of the vectors are comparable, but scaled for presentation purposes.

Any given ice drift vector (*this-vector*) has to pass a number of filters before final release. The full filtering routine works in 5 steps in the following order:

1. Minimum Cross Correlation threshold
    - If the Maximum Cross Correlation between the *reference* and the *compare* matrices is less than 0.6, *this-vector* is dismissed.
2. Displacement length - neighbourhood homogeneity.
    - If the vector length difference between *this-vector* and the mean of all neighbouring pixels is larger than a given threshold, *this-vector* is dismissed.
3. Minimum number of neighbours.
    - If *this-vector* has less than 4 neighbouring drift vectors, *this-vector* is dismissed.
4. Direction - neighbourhood homogeneity
    - If the bearing of *this-vector* diverges more than a given threshold from the mean bearing of the neighbouring ice drift vectors - *this-vector* is dismissed.
5. Re-running filter number 3
    - After applying filter 1-4 to the non-filtered ice drift estimates, filter number 3 is re-applied on the remaining data.

The effect of the applied filter can be seen in figure 3, showing the un-filtered ice drift estimates and the final product. ‘Neighbourhood’ in filtering context is a 7 by 7 grid point matrix, around *this-vector*.

3. Algorithm characteristics

The characteristic numbers for this ice drift estimation setup are:

- The correlation matrix is 41*41 pixels in the *input-grid*, i.e. 41*41km
- The *icedrift-grid* is 20 by 20 km
- The maximum allowed ice drift speed over 24h is 0.3 m/s, i.e. fixing the maximum allowed 24h ice drift to 25.92 km.
- ‘Neighbourhood’ is a 5x5 *ice drift-grid* matrix around the *icedrift-grid* point of interest.

4. Validation strategy

A product specific validation report is prepared [VAL] and also an ice drift product comparison report is finalized through a visiting scientist program [Hwang and Lavergne]. Here comprehensive validation work is describet. The validation report contains statistics for approximately one year of ice drift data. The report contains RMS errors, absolute error and correlation statistics, like shown in table 3. Here are correlations between the satellite ice drift and buoy drift for two directional components, U and V. Also the RMS error, mean absolute error and bias are calculated.

Beside the general validation report, monthly validation is worked out with statistics comparable to the general validation report. This monthly report will be based on ‘on the fly’ validation statistics generated for each ice drift data set.

Table 3 Validation example for the ice drift product based on all available Argos buoy data in the GTS network. Valid for January 2009 (see also figure 4).

Directional validation	
Correlation delta U	0.937015
Correlation delta V	0.903716
Number of samples	1483
Absolut displacement validation	
RMS Error	1.37 km
MEAN absolut Error	0.91 km
Bias (buoy-sat)	-0.03 km

In figure 4 the data basis for the validation values in table 1 are plotted as a scatter plot.

All validation is based on buoy data retrieved from the GTS network at DMI. Each buoy data are paired to all ice drift estimates within a vicinity of 50 km. Though it is well known that the Argos positioning system can be associated with errors, these data are never the less a comprehensive data set available in near real time, which is needed for ‘on the fly’ validation.

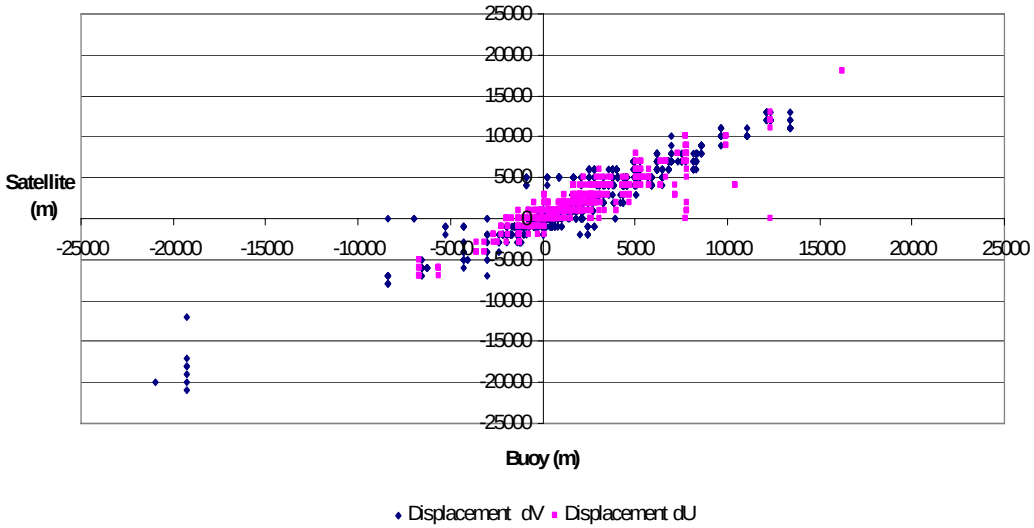

Figure 4 A scatter plot showing buoy drift and satellite estimated ice drift for U- and V-directions, where the U and V directions are oriented right-left and up-down, respectively, in the NH subset in figure 1 and specified in tables 1 and 2. The data are the basis for the error statistics in table 3.

5. Limitations and assumptions

There exist no single data sources that can fulfil all needs for ice drift data sets. Large scale and low resolution data, from passive microwave instruments like SSMI and AMSR-E, can provide full Arctic coverage on a daily basis. However, the spatial resolution is coarse and using such data ice drift estimation for less than 48h only makes limited sense (see Laverne2009). Such data are perfectly suited as input for large scale modelling, general sea ice circulation studies and climate studies.

On the other hand, SAR sensors, can provide very high resolution ice drift information unaffected by atmospheric properties, but these data have only limited spatial coverage, resulting in only partial Arctic coverage on daily basis.

In between these high and low resolution microwave data sources are the AVHRR VIS and IR data. They provide very wide swath data, at medium spatial resolution and high repetition rate, suitable for 24h ice drift estimation. Short ice drift period and high spatial accuracy and resolution make the data suited for calibration and validation purposes and to some extent also for data assimilation.

1. Limitations

Due to the sensitivity of VIS and IR data to atmospheric water, the AVHRR data can not provide fully gridded ice drift data on a daily basis. For a given area of interest both *reference* and *compare* data must have clear sky conditions in order to calculate ice drift. This limits the use of AVHRR data for surface analysis.

During the arctic summer this limitation is pronounced, as clouds often cover large parts of the arctic region. Moreover, in periods with surface melting (summer), the contrasts in the IR data are drastically reduced and consequently making surface feature analysis difficult. During that period, this ice drift procedure uses VIS data, as this naturally coincide with periods with midnight sun. Despite the substitution of IR data with VIS data in the summer period, the data frequency drops to about 12 % of data frequency around January, where the cloud cover is at a minimum. Ice drift data frequencies for both IR and VIS data are plotted in figure 5, for a 9 month period.

In an earlier setup of this ice drift procedure the possibility of deriving 12h ice drift was tested. This test failed, due to large validation errors, with absolute mean drift errors of approximately 5 km between satellite drift and buoy drift. The reason for large error statistics for the 12h ice drift setup is that an ascending swath is compared to a descending swath, resulting in overlapping of mainly swath edge areas. Here the geo-rectification errors are much larger than corresponding errors for the swath centre, and much higher than sub-pixel accuracy (see section 2.1). In general applies: “approaching shorter drift period the error to signal ratio will approach infinite”.

In contrast, the swath overlap between two swath data sets separated by 24 hours will overlap almost perfectly and provide the highest number of well rectified pixels for the ice drift detection procedure. Therefore this procedure is set up to generate 24h ice drift product.

2. Assumptions

The basic principle behind this feature tracking routine is the assumption of conservation in the features being tracked, i.e. the shape of the features must appear relative equal in both *reference* and *compare* data. This must comply to a degree where the correlation between the *compare* correlation matrix correlates to the *reference* correlation matrix with a correlation value, r, greater than 0.6.

It is also assumed that features of interest are moving with a minimum of rotation.

It is further assumed that the net 24h displacement does not exceed 0.3 m/s and finally the feature being tracked must have 2-dimentional characteristics. I.e. if the feature is a strait lead exceeding the correlation matrix, the MCC routine will calculate a 'ridge' of almost equally high correlated match ups between the two correlation matrices and hence make the best match between *reference* and *compare* data dubious.

In summary the assumptions can be expressed like this:

a) No change in feature shape between *reference* and *compare* data.
b) no or little rotation.
c) maximum 24h drift of ~25km.
d) 2-dimentional feature inside the correlation matrix.

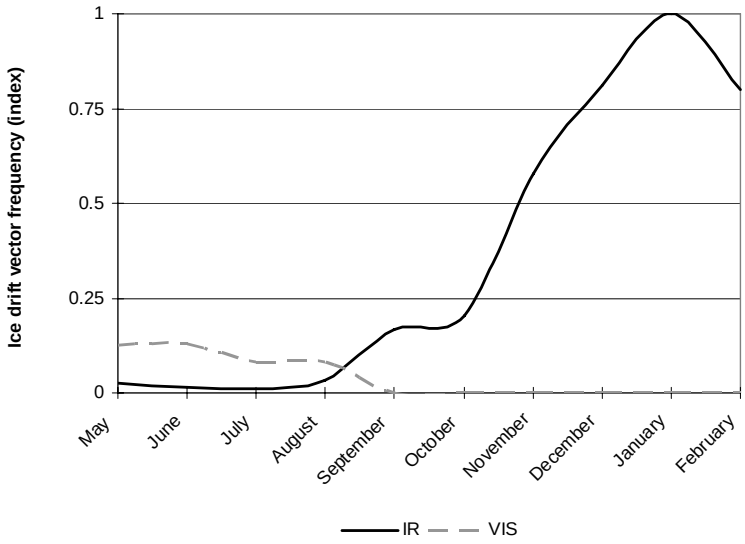

Figure 5 Standardized ice drift vector frequency distribution for IR and VIS data, for an area North of Greenland, during 9 month of 2005-2006. During summer the successful retrieved ice drift vectors from IR data are practically zero in comparison to the number of ice drift vectors during winter. During spring and summer the successfully retrieved ice drift vectors from VIS data are approximately 12 percent of the maximum ice drift vector frequency in January and February.

6. Reference

Andersen S., Breivik L.A., Eastwood S., Godøy Ø., Lind M., Porcires M., Schyberg H. 2007. OSISAF Sea Ice Product Manual v3.5, January.

EUMETSAT. 2007. EPS Operations Services Specification, January 17th.

Ezraty, R., F. Arduin and Jean-Francios Piollé. 2006. Sea Ice Drift in the central Arctic estimated from Seawinds/Quickscat backscatter maps. IFREMER, Users Manual version 2.2.

Haarpaintner, J. 2006. Arctic-wide operational sea ice drift from enhanced-resolutionQuikScat/SeaWinds scatterometry ands its validation, IEEE Trans. Geoscie. Remote Sens.,vol. 44, no.1, pp.102-107.

Hwang, P. and T. Lavergne, 2010. Validation and Comarison of ISI SAF Low and Medium Resolution and IFREMER/Cersat Sea Ice drift products. Reference: CDOP-SG06-VS02. http://osisaf.met.no/docs/OSISAF_IntercomparisonIceDriftProducts_V1p2.pdf

Laverne, Thomas. 2009. Algorithm Theoretical Basis Document for the OSI SAF Low Resolution Sea Ice Drift Product. SAF/OSI/CDOP/met.no/SCI/MA/130.

Maslanik, J., M. Drinkwater, W. Emery, C. Fowler, R. Kwok and A. Liu. 1998. Summary of ice-motion mapping using passive microwave data. National Snow and Ice Data Center (NSIDC) Special Publication 8.

PROJ4. http://trac.osgeo.org/proj/

PUM. 2009. Medium Resolution Sea Ice Drift Product User Manual. SAF/OSI/CDOP/DMI/TEC/MA/137. GBL MR SID – OSI 407.

VAL. 2009. Validation and Monitoring of the OSI SAF Medium Resolution Sea Ice Product. SAF/OSI/CDOP/DMI/T&V/RP/133. OSI – 407.