DTU Library

Experiences With an Optimal Estimation Algorithm for Surface and Atmospheric Parameter Retrieval From Passive Microwave Data in the Arctic

Scarlat, Raul Cristian; Heygster, Georg; Pedersen, Leif Toudal

Experiences with an optimal estimation algorithm for surface and atmospheric parameter retrieval from passive microwave data in the Arctic

Raul Scarlat, Georg Heygster, Leif Toudal Pedersen

Abstract—We present experiences in using an integrated retrieval method for atmospheric and surface parameters in the Arctic using passive microwave data from the AMSR-E radiometer. The core of the method is a forward model which can ingest bulk data for seven geophysical parameters to reproduce the brightness temperatures observed by a passive microwave radiometer. The retrieval method inverts the forward model and produces ensembles of the seven parameters, wind speed, integrated water vapor, liquid water path, sea and ice temperature, sea ice concentration and multi-year ice fraction. The method is constrained using numerical weather prediction (NWP) data in order to retrieve a set of geophysical parameters that best fit the measurements. A sensitivity study demonstrates the method is robust and that the solution it provides is not dependent on initialization conditions. The retrieval parameters have been compared with the Arctic Systems Reanalysis model data as well as columnar water vapour retrieved from satellite microwave sounders and the Remote Sensing Systems AMSR-E ocean retrieval product in order to determine the feasibility of using the same setup over pure surface with 100% and 0% sea ice cover respectively. Sea ice concentration retrieval shows good skill for pure surface cases. Ice types retrieval is in good agreement with scatterometer backscatter data. Deficiencies have been identified in using the forward model over sea ice for retrieving atmospheric parameters, that are connected to the treatment of surface emissivity and surface temperature. The retrieval agrees well with legacy atmospheric retrieval products in open ocean areas.

Index Terms—Remote sensing, Sea ice, Arctic regions, Atmospheric measurements.

I. INTRODUCTION

In passive remote sensing from satellites, the radiometer signal has contributions from surface and atmosphere. Single parameter retrieval algorithms consider one of the two contributions as the signal of interest, and the other as noise to be ignored or compensated for, e.g. based on other channels of the radiometer or through a priori knowledge. The basic idea of integrated retrieval is to find a set of geophysical parameters which, if applied to a forward model, simultaneously yields in good approximation the observed brightness temperatures of all radiometer channels used. Over open ocean, such integrated retrieval exists and has been applied for more than a decade

[1]. It is based on a forward model which predicts the observed brightness temperatures from the geophysical state of surface and atmosphere. However, over sea ice, an integrated retrieval is much more difficult. The challenges are (i) the high surface emissivity which makes the atmospheric component in the signal much smaller, and (ii) the high variability of the sea ice emissivity in space and time. Although sea ice forward models exist, their use in integrated retrieval has been limited until now because the number of required geophysical parameters is high and their values are generally unknown. For these reasons until now only little effort has been undertaken for integrated retrieval of surface and atmospheric parameters over sea ice [2], [3]. In order to take advantage of the multispectral capabilities of imaging radiometers, in [4] an integrated retrieval method is proposed that can retrieve seven geophysical parameters, wind speed (W), liquid water path (LWP) , total columnar water vapor (TWV) , sea surface temperature (SST), ice temperature (IT), sea ice concentration (SIC) and multi-year ice fraction ($MYIF$). Optimal estimation (OE) techniques are used to invert the forward model and extract the ensemble of seven parameters that optimally match the observed brightness temperatures. A priori information from climatological and meteorological sources is used to constrain the method to the natural variability of each parameter.

In the current paper we demonstrate the feasibility and current limits of extending the coverage of this integrated retrieval using optimal estimation methods from ocean to sea ice. Moreover, we test how influential the a priori constraints are and how robust the method is to different initialization setups. Will the optimal estimation method (referred thereafter as OEM) be biased towards the NWP a priori data or can it move away from those in order to match the measurements? The NWP data is a powerful information source but it does not provide a complete true state of the climate system which is why we want to use satellite observations in a consistent method that can complete the picture.

The OEM can use different channels from a microwave radiometer like the AMSR-E and AMSR2 instruments as an input source. In order to avoid the discrepancies that naturally occur between the different channels as each channel footprint covers a somewhat different region of the Earth surface, we use the AMSR-E Level 2A product which offers resampled brightness temperatures in which all channels represent the same footprint at the same resolution. This product is obtained by applying the Backus-Gilbert method to Level 1B data in their native resolution [5]. The source for a priori information

is the ERA-Interim reanalysis [6], collocated in time and space to the satellite footprint.

In order to test the behaviour of the retrieval, two data sets of AMSR-E level 2A brightness temperatures are used. The first data set comprises two subsets which contains cases of validated 100% (SIC1) sea ice concentration and the other contains cases of open water near, but at a safe distance from the ice edge (SIC0) [7]. The dataset originates from the sea ice project of the Climate Change Initiative and is referred to as the Round Robin Data Package (RRDP). The RRDP provides a source for validating the sea ice concentration retrieval [7], [8] and for testing atmospheric parameter retrieval over pure surface types. For each pixel in these datasets, atmospheric parameter data from the ASR model [9] have been collocated. In order to observe the retrieval behaviour over a wider set of Arctic scenes, a larger comparison dataset was also prepared. This second dataset includes one full days data every ten days of the year 2006, for a total of 18 winter and 18 summer season days, and it covers the entire Arctic region north of $60°$ N. The 2006 dataset starts on January 11th 2006 and ends on January 6th 2007. As water vapour retrieval over the Central Arctic is one of the goals of the integrated retrieval also use as reference a dataset from a microwave sounder based method that is especially accurate for the low atmospheric water vapour values encountered in the Arctic [10].

II. DATA

As a source for the radiometer measurements we used the AMSR-E Level 2A dataset [11]. This data product offered by the NASA National Snow and Ice Data Center (NSIDC) contains several spatially consistent subsets of brightness temperature observations resampled to the footprint sizes of the 6.9, 10.7, 18.7, 23.8, 36.5, and 89 GHz channels. This is achieved by bringing the Level 1A antenna temperatures to the common spatial resolution using a set of weighted coefficients. Every Level 2A observation in a scan line is calculated using the coefficients that correspond to the relative weights of all neighbouring Level 1A observations. These coefficients are unique for every scan position within one scan line but they do not vary between different scan lines. The weighting coefficients for the Level 1A observations are produced using the Backus-Gilbert method. The different subsets are produced by resampling the higher resolution channels to match the larger footprint size of the lower resolution channels. The available spatial resolutions are at 56 km for all channels, 38 km for all channels frequencies above 6.9 GHz, 21 km for frequencies above 10.7 GHz, 12 km for frequencies above 23.8 GHz and 5.4 km for the 89 GHz channels only. In order to include the 6.9 GHz channels in the OEM retrieval, we use the AMSR-E Level 2A set with all channels resampled at 56 km spatial resolution.

The RRDP and the 2006 datasets will be used to test the retrieval in a scenario where the surface conditions are known in order to judge the impact of sea ice presence on the atmospheric retrieval. The RRDP dataset also includes ERA-Interim values for air and skin temperature, wind speed, liquid water path and total water vapour collocated with each data point. Also included in the RRDP are collocated ASCAT (Advanced Scatterometer) backscatter values which is useful in classifying different sea ice types according to surface and volume scattering properties. Because the SIC1 dataset is based on identifying convergence zones from sea ice tracking data, this dataset is only valid for winter months when leads that do not close during ice movement will freeze over. We have only used the data points between October and May of each year between 2007 and 2011.

The Remote Sensing Systems (RSS) AMSR-E ocean products use AMSR-E brightness temperatures and an algorithm adapted from [12] to retrieve wind speed, columnar water vapour, liquid water path and sea surface temperature over ice free ocean regions. This data set covers the entire nine year lifespan of AMSR-E and has been used for creating derived products [13] and validated against ship based observations [14]. Because it uses the same instrument this product provides a sanity check for the quality of the OEM retrieval over open ocean with less concern about collocation induced discrepancies.

In order to test the retrieval we needed a reference dataset covering as many of the parameters as possible. In order to avoid the difficulties from comparing our results with different single parameter retrieval products we selected the Arctic Systems Reanalysis model [9] as a suitable source of data as it provides a high resolution ensemble of consistent atmospheric and surface temperature parameters. The data assimilation systems have been optimized for the Arctic. The resolution of the product dataset is 30 km and fields of atmospheric water vapour, liquid water path, wind speed and skin surface temperature have been collocated with the RRDP data points.

As an additional comparison tool for atmospheric water vapour retrieval we selected the AMSU-B total water vapour product [10]. This product is specially designed for the Arctic region and it uses the high frequency sounding channels of the AMSU-B instrument to detect the low water vapour values typical for the atmosphere over sea ice in the Central Arctic. One issue connected with this product is that retrieval uncertainty increases with increasing water vapour values from about 1 mm below 2 mm to around 3 mm above 14 mm.

III. METHOD

A. Optimal estimation method (maximum a posteriori solution)

Satellite radiometers measure the radiation naturally emitted by the atmosphere and the Earth surface below.

Following the radiative transfer theory we can represent the vector T_A of brightness temperatures measured at the top of the atmosphere by a microwave radiometer as a function of a number of geophysical parameters.

$$T_A = F(p) \qquad (1)$$

where p is the state vector containing both surface parameters, atmospheric profiles and sea ice parameters. F is the forward

operator (see Section III-D) that maps the state vector parameters to the observed brightness temperatures.

The estimation method used here follows [15] and is called the maximum a posteriori method. Through an iterative process the prior state vector p_n is nudged to p_{n+1} in the direction of minimizing the difference between the satellite measured brightness temperatures and the brightness temperatures simulated by forward model while not deviating too much from the background state.

The optimal estimation method uses a priori information on the state of the atmosphere and ocean as an input for a forward model in order to obtain simulated brightness temperatures. The method iterates towards the optimal set of geophysical parameters which when used as input for the forward model best matches the observed brightness temperatures. A cost function that balances the penalties of departing from the background state vector with those for deviating from the observed brightness temperatures is evaluated at each iteration step. A constraint that reduces the number of possible state vector ensembles is the background covariance matrix, while the microwave measurement covariance matrix restricts the observation space.

B. The inversion method

The solution is found by inverting

$$T_A = F(p) + e \qquad (2)$$

where T_A is the vector of satellite measured brightness temperatures. F is the forward model, p is the state vector comprising in our case the geophysical parameters wind speed W, total water vapour TWV, liquid water path LWP, open water temperature SST, sea ice temperature IT, total ice concentration SIC and multi-year ice fraction $MYIF$. e is the measurement error. The solution is found by using the Jacobian of the forward model

$$M = \frac{\partial F(p)}{\partial p} \qquad (3)$$

and iterating the state vector following the Gauss-Newton method

$$p_{n+1} = p_n + \hat{S}_n^{-1}(M_n^T S_e^{-1}(T_A - F(p_n)) + S_p^{-1}(p_a - p_n)) \quad (4)$$

where $M_n = M(p_n)$ and

$$\hat{S}_n = (S_p^{-1} + M_n^T S_e^{-1} M_n)^{-1} \qquad (5)$$

is the a posteriori covariance matrix of the state vector.

The Gauss-Newton iteration nudges the state vector to follow a quadratic approximation of the cost function. In the case that the forward model is too non-linear and cannot be well represented by a quadratic function, the iteration step towards finding a solution needs to be chosen at each iteration in order to decrease the cost function [15]. This can be achieved by adding an additional parameter to the iteration rule as proposed in the Levenberg-Marquardt method

$$\hat{S}_n = ((1 + \gamma)S_p^{-1} + M_n^T S_e^{-1} M_n)^{-1}. \qquad (6)$$

γ is chosen at each iteration step and the cost function is reevaluated. The initial value used for γ is 10^{-5}. If the cost function decreases the iteration values are accepted. If the cost function does not decrease then γ is increased by a factor of 10 to force the iteration on a steeper decent (and a smaller step size) towards the minimum and the same iteration step is repeated. The current version of the retrieval uses the Levenberg-Marquardt form (6) of the iterative equation.

The convergence conditions suggested by [15] are implemented in order to stop the iteration process and output the optimal solution. The method will check if

$$d^2 = (p_{n+1} - p_n)^T \hat{S}_n^{-1}(p_{n+1} - p_n) < 7 \qquad (7)$$

is true where d^2 represents the degrees of freedom of the system and the number 7 was chosen as a limit because it is the dimension of the state vector. An additional condition for the solution to be accepted is check for the minimum of

$$\Delta = ||T_A - F(p_n)||_2. \qquad (8)$$

If convergence has been reached the step p_n for which the cost function is minimised is accepted as the solution and sent to output.

C. Error treatment and covariance matrices

The background values represent long term means from climatological or other sources which together with the individual parameter variances in the background covariance matrix S_p constrain the retrieved parameters to physically realistic values. The diagonal elements of this covariance matrix should represent the natural variability of each parameter in order to allow for a consistent retrieval, but this information is limited by the quality of the prior information we have about the climate system. In order to speed up the iterative process for finding the optimal solution, the method also uses a first guess point that serves as initialization state. The first guess state can also come from a background state but that is not the only source [15]. According to [2] the retrieval accuracy is influenced by the quality of the first guess data. In order to test the sensitivity of the retrieval method to the background covariance matrix and to the first guess conditions, different implementations of the optimal estimation retrieval are compared. For each of these tests one reference version was kept the same. It uses a diagonal background covariance matrix (see Appendix A, Table XI) obtained from year long ECMWF ERA Interim mean variances for the atmospheric parameters and the surface temperatures. For the sea ice concentration and multi-year fraction variances, a locally processed dataset using the NASA Team algorithm [16] was used for the same time period. The background state vector for all versions tested below represents the yearly mean values for each parameter (Appendix A, Table XIV), and is calculated from the same sources as the reference background covariance matrix. For this reference OEM version, the start guess comes from temporally and spatially collocated ECMWF ERA Interim data.

For testing the background covariance matrix S_p, a reference implementation of the OEM uses data from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim atmospheric reanalysis for calculating the diagonal elements of the covariance matrix. We compare this with two different implementations that use climatology derived values from the Max Planck Institute (MPI) Climate model (F. Bunzel, personal communication). The main difference between the ERA-Interim and the MPI derived parameter variances is that the former included one year of data while the latter only one winter month. From the two MPI derived matrices one is a diagonal covariance matrix of the same shape as the reference, while the other, besides the same diagonal elements, also includes the off-diagonal elements that represent covariances between parameters.

All three versions of the OEM are run for the RRDP SIC1 dataset and the retrieval parameters are compared to collocated ASR fields for *TWV*. The same background values calculated from ECMWF yearly means were used for all versions, the only difference between these OEM implementations being the shape of the background covariance matrix. The results are shown in Table I.

TABLE I
COMPARISON OF OEM RETRIEVAL AND ASR USING DIFFERENT
BACKGROUND COVARIANCE MATRICES.

Parameter	Stat	Ref	Climate-diag	Climate-full
TWV [mm]	Bias	-0.28	-0.09	2.32
	RMSD	2.43	2.08	3.35
	R	0.21	0.22	0.11
LWP [mm]	Bias	0.20	0.17	0.16
	RMSD	0.27	0.19	0.19
	R	0.01	0.04	0.07
IST [K]	Bias	15.21	16.60	17.05
	RMSD	16.01	17.24	17.73
	R	0.69	0.71	0.69
	Convergent pts.	11681	8180	10377
	Avg. iterations	13.9	14.37	15.25

The largest difference between versions can be seen in the number of convergent cases and the average number of iterations required to reach convergence. Ideally the off diagonal elements of the background covariance matrix should introduce more information about the state of the climate system before the measurement are taken. However, it is difficult to evaluate whether the covariances used are representative for the yearly time scales used in the retrieval testing. The results of this comparison seem to indicate that these covariances impose an additional constraint on the method which in certain cases cannot be satisfied within 50 iterations resulting in non-convergent results. The comparison with ASR shows small differences between versions for retrieval accuracy, therefore we chose the best case scenario of using ERA-Interim data as a source in order to achieve the highest number of valid retrieval pixels.

For each AMSR-E Level 2A footprint, spatially and temporally collocated ECMWF values are used as the first guess in the iterative process. This represents a reference version for testing the sensitivity to initialization conditions. For this test, a different implementation of the retrieval was used, that has just one fixed state vector as first guess for all pixels. The

results of comparing these two OEM versions with ASR data are shown in Table II.

TABLE II
COMPARISON OF OEM RETRIEVAL AND ASR USING DIFFERENT FIRST
GUESS CONDITIONS.

Parameter	Stat	Ref	Static first guess
TWV [mm]	Bias	-0.28	-0.21
	RMSD	2.43	2.33
	R	0.21	0.18
LWP [mm]	Bias	0.20	0.21
	RMSD	0.27	0.27
	R	0.01	0.02
IST [K]	Bias	15.21	15.87
	RMSD	16.01	16.58
	R	0.69	0.69
	Convergent pts.	11681	11893
	Avg. iterations	13.90	13.09

The differences between the two cases are small. The individually adapted start points reduce slightly the number of required iterations, but do not change the comparison results, indicating the independence of the iteration procedure to the start guess values. We would expect normally that a fist guess that is closer to the solution will result in a faster convergence and more convergence cases than if one fixed value is used as the starting point. However, this result seems to indicate that the collocated ECMWF is in some few cases farther away from the solution than one fixed mean set of parameters. It should be noted however that the difference in number of convergent points between the two cases is small (1.8% difference).

The satellite measurements are connected with a measurement error due to instrumental noise. In addition to this noise, the simulations themselves contain modelling errors. These error sources need to be taken into account when constructing the constraints under which the OEM has to match the measurements. Following the testing done for the background covariance matrix and the start guess position, we compared four different brightness temperature covariance matrices (S_e) in order to asses the impact on the retrieval. The reference run is based on the pre-launch AMSR-E radiometric error values. A different covariance matrix is constructed from the variances of the differences between observed and modelled brightness temperatures over the entire RRDP SIC1 data set. These variance values should include the modelling error as well as the errors of the forward model input parameters and the measurement errors of the instrument. These values are obviously larger than the pre-launch radiometric errors so that this covariance matrix is named L for short reference. In order to test the sensitivity of the retrieval to changes in the brightness temperature covariance matrix we also tested one version that uses all values in L multiplied by two and one that uses the same elements but divided by two. The results of comparing these retrieval runs against ASR are shown in Table III.

The brightness temperature covariance matrix has the largest impact on the retrieval out of all of the a priori constraints imposed on the method. By including the modelling and input parameter error into the S_e matrix, the modelled brightness temperatures are allowed to have larger deviations from the measurements. The increase in number of convergent cases

TABLE III
COMPARISON OF OEM RETRIEVAL AND ASR USING DIFFERENT
BRIGHTNESS TEMPERATURE COVARIANCE MATRICES.

Parameter	Stat	Ref	L	$2L$	$1/2L$
	Bias	-0.28	0.09	0.39	-0.23
TWV [mm]	RMSD	2.43	1.21	1.23	1.25
	R	0.21	0.26	0.22	0.31
	Bias	0.20	0.11	0.13	0.09
LWP [mm]	RMSD	0.27	0.12	0.13	0.11
	R	0.01	0.08	0.07	0.09
	Bias	15.21	7.27	5.61	9.17
IST [K]	RMSD	16.01	7.85	6.17	9.84
	R	0.69	0.91	0.93	0.85
	Convergent pts.	11681	15884	16428	14935
	Avg. iterations	13.91	11.87	8.92	11.78

as well as the decrease in average number of iterations was to be expected as practically the constraint on brightness temperature matching has been relaxed. Data points where the measurements could not be matched to within radiometer error before 50 iterations were reached will now achieve this match faster with the larger permitted TB error. This is evident in the differences between the $1/2L$, L and $2L$ versions where the larger variances result in more convergent points. This mechanism can also be seen in the variation in the number of average iterations needed for reaching convergence. The $2L$ version has the smallest number of average iterations with 8.92 while the $1/2L$ and L versions are very close to each other with 11.78 and 11.87 respectively. The better agreement of the L matrix OEM versions with ASR data data represents the trade-off for the optimal estimation approach. While it is necessary to account for the TB modelling error in the constraints to ensure an unbiased retrieval, and because the background constraints have not changed, it follows that the method output will move closer to the NWP data when relaxing the brightness temperature constraint.

It is important to note that the background state is static (from climatological means) while the comparison is done with collocated simultaneous data. Also of note is that the ASR data used in this comparison is not the same as the ECMWF data source used in calculating the background constraints and first guess values for the method, but the two are very similar. More details about the correlation between these two NWP data sources are given in the Results (Section IV).

D. The forward model

The purpose of the forward model is to translate the seven geophysical parameters into the twelve brightness temperatures corresponding to the AMSR-E radiometer. Here we use a modified version from [1] as forward operator which will be inverted in order to perform the retrieval. Over open ocean the original forward model has not been changed. Here we will overview the working principles in short and we refer to the original paper for a detailed description. The sea surface emissivity model takes into account water salinity, sea surface temperature and a geometric model for the wind roughened ocean surface to calculate the surface contribution. For the atmospheric component the forward model includes frequency specific absorption and scattering coefficients. In

the frequency range of the AMSR-E channels the main atmospheric absorbers are oxygen and water vapour so these values have been calculated from compiled radiosonde observations. The atmospheric emission is computed from the amount of atmospheric water vapour and the liquid water path through a list of regression equations. An important feature of the forward model is that the radiative transfer equation uses the absorption-emission approximation which excludes scattering from large rain drops and ice particles. This assumption is valid for clear and cloudy skies and for light precipitation conditions for the frequency range of 6 to 37 GHz. At the 89 GHz channels however, scattering by clouds is no longer negligible.

The original implementation of the forward model was designed to work over open ocean surfaces. In order to use it over the ice covered central Arctic the calculation of the surface emission has been modified in [4] to account for the different microwave emissivity of sea ice. Each pixel is considered as being formed of a mixture of thermal microwave emission from open water, first year and multi-year ice. The up-welling thermal contribution of the surface is:

$$T_S = C_{ow}E_{ow}T_{ow} + C_{fyi}E_{fyi}T_{fyi} + C_{myi}E_{myi}T_{myi} \quad (9)$$

where C, E and T are the concentration, emissivity and temperature of open water (ow), first year ice (fyi) and multi-year ice (myi) concentrations respectively. The sum of the three concentrations must be 1. The second equation that needs to be modified is the one responsible for calculating the reflection and scattering of the atmospheric down-welling radiation. The reflectivity R of a pixel will be calculated from individual reflectivities of each surface type:

$$R = 1 - E = 1 - (C_{ow}E_{ow} + C_{fyi}E_{fyi} + C_{myi}E_{myi}). \quad (10)$$

The frequency dependent emissivities are based on the values retrieved empirically by [17]. These represent monthly averages and the individual value for each pixel is calculated by interpolating in time between these monthly values.

Because sea ice dominates the microwave signal at the instrumental frequencies, it is important that the forward model can simulate unbiased brightness temperatures when compared to the measurements. Any bias in the simulations will influence the retrieved geophysical parameters because the optimal estimation method will try to compensate for the brightness temperature difference by adjusting the predicted state vector. In order to test whether a bias is present, we calculated the mean differences between the retrieval step simulated brightness temperatures and the AMSR-E measurements for the SIC1 dataset (Table IV). The highest discrepancies occur in the two 89 GHz channels and for now we have discontinued including these two channels in any retrieval scheme and also in the test cases described in the previous section. Excepting the 89 GHz channels, high biases are also present in the horizontal polarization channels while the lowest bias at 18.7V is still around 1 K. In order to compensate for these biases over sea ice, we assume that the difference in brightness temperatures is entirely caused by the surface component

through the prescribed emissivities. Under this assumption we calculate a set of corrected emissivities that minimise the channel wise difference between the simulated and observed brightness temperatures.

In order to minimise the influence of monthly variations throughout the winter season, the new corrected emissivities are calculated as winter averages, even though a slight improvement may be obtained by using temporally varying emissivities. After correction the differences are lowered across all channels with values between 0.15 and 1.7 K for the 6.9 - 37 GHz interval (Table IV, Column C). The bias in the 89GHz channels has also been lowered but still remains higher than 6.5 K. The initial and corrected set of emissivities are shown in Appendix A (Table XIII).

TABLE IV
MEAN DIFFERENCES (IN K) BETWEEN SIMULATED AND OBSERVED BRIGHTNESS TEMPERATURES. COLUMN A: INITIAL SETUP THAT USES EMISSIVITIES INTERPOLATED BETWEEN MONTHLY MEANS. COLUMN B: EMISSIVITIES AVERAGED OVER THE ENTIRE WINTER DATASET FOR EACH ICE TYPE. COLUMN C: DIFFERENCES AFTER CORRECTING THE AVERAGED WINTER EMISSIVITIES.

Channel	A	B	C
SIC1 winter pixels			
6.9V	2.335	4.776	1.711
6.9H	4.998	6.408	1.46
10.7V	3.094	5.133	1.451
10.7H	6.011	6.751	1.719
18.7V	0.982	1.936	1.046
18.7H	3.039	3.145	1.589
23.8V	1.033	1.507	0.587
23.8H	2.436	2.531	0.588
36.5V	1.566	0.554	-0.236
36.5H	2.372	1.657	-0.156
89V	-18.461	-15.428	-6.678
89H	-21.315	-16.365	-7.689
Multi-year ice pixels			
6.7V	3.957	5.337	2.015
6.7H	4.178	3.092	1.196
10.7V	3.724	5.183	1.692
10.7H	2.836	1.893	1.384
18.7V	-0.831	0.618	0.957
18.7H	-2.703	-2.332	0.451
23.8V	-2.944	-1.56	-0.226
23.8H	-3.432	-2.673	-1.157
36.5V	-7.462	-6.267	-1.948
36.5H	-5.939	-4.927	-2.168
89V	-26.249	-24.227	-7.303
89H	-24.492	-22.127	-8.136
First year ice pixels			
6.9V	3.191	5.611	1.344
6.9H	7.301	10.072	1.181
10.7V	3.708	5.873	1.047
10.7H	8.266	10.651	1.096
18.7V	1.304	2.643	0.735
18.7H	4.971	6.814	1.326
23.8V	1.580	2.871	0.638
23.8H	4.048	5.769	0.917
36.5V	3.731	4.137	0.814
36.5H	4.789	6.188	1.224
89V	-12.672	-6.959	-4.055
89H	-16.918	-8.506	-5.197

E. OEM retrieval setup

Following the sensitivity to a priori constraints presented above, we chose a retrieval set-up that includes the reanalysis derived background covariance matrix and their climatological

mean as background values. The first guess is given by collocated ERA-Interim fields for the atmospheric parameters and NASA Team derived sea ice concentration and multi-year ice fraction. After applying the corrected sea ice emissivities, the brightness temperature covariance matrix is calculated from the variances of the channel wise differences between simulations and observations. The values for for the background error (S_p) and the brightness temperature (S_e) covariance matrices are shown in Appendix A (Tables XII, XI).

IV. RESULTS

For developing this method into an operational retrieval rigorous validation against in-situ or other independent reference data is necessary. When we are dealing with seven different geophysical parameters, finding adequate validation data for each of them is a challenging task. Field experiments in the Arctic naturally take place during the boreal summer when the sea ice is more likely to be affected by the presence of melt-ponds and surface melt that greatly affect the microwave signature of the surface. Also because the current working resolution of the OEM depends on that of its lowest frequency channel, 56 km, localized radiosonde and surface data can also prove problematic in that they don't represent a large scale average. While a more thorough validation effort for the output parameters of the OEM is planned for future studies, in the current work we want to focus on testing the working principles of the method. An example of how the retrieval output looks like for one day (January 11th, 2006) is shown in Figure 1. Most parameters look realistic with the notable exception of LWP and IT. Both parameters show values that seem too large for the Arctic with a few regions approaching 0.3 mm of LWP and a large swath of the Central Arctic showing ice temperatures above 270 K. All parameters are presented in more detail below, with a description of possible error sources in the discussion (Section V).

A. Wind speed

The retrieval of atmospheric parameters is conditioned by the relationships through which the forward model relates these parameters with the top of the atmosphere brightness temperatures. In the microwave spectrum the surface roughness together with the surface temperature and water salinity determine the ocean surface emissivity. The ocean salinity is prescribed in the model from climatology sources since it has a negligible impact for frequencies above 5 GHz. In the case of wind speed the connection is made through the modelled surface emissivity of the ice-free ocean. This means that although the method will retrieve wind speed values as part of the state vector even in sea ice covered regions, these values reproduce the climatological background and are not considered a result of the retrieval and are not analysed further. The RRDP SIC0 data set is especially well suited for evaluating the wind speed retrieval because it offers a large set of open ocean data points free of sea ice contamination. When using the 2006 data set only data points below 15% SIC were compared to the ASR data. The RSS product is also included in a comparison as a sanity check for the OEM retrieval. A
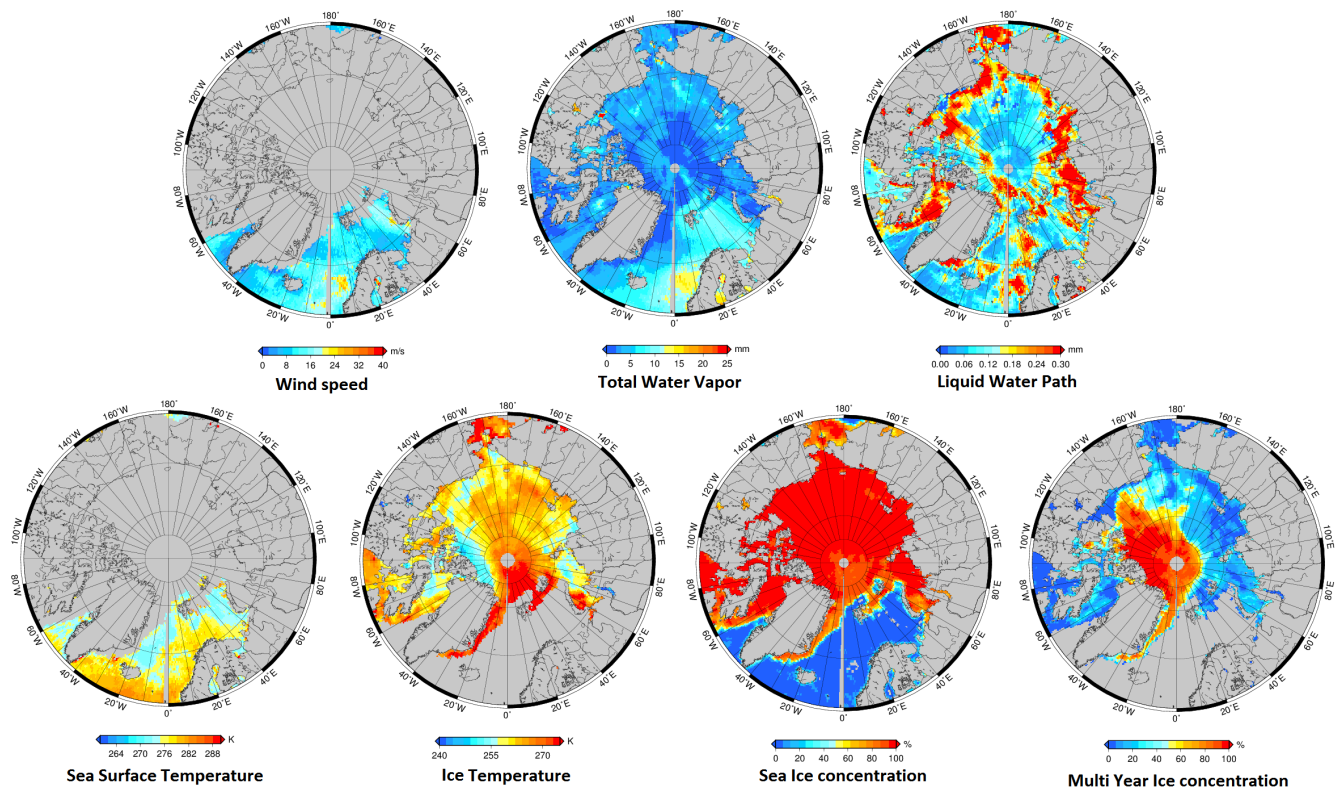
Fig. 1. OEM retrieval example of seven parameter maps for 11th of January 2006.

big advantage of this product is that because it uses the same AMSR-E instrument perfect collocation is possible. The RSS product uses a newer version of the Wentz forward model than the one implemented in the OEM, but the retrieval over ocean is based on TB regressions and does not use the supplementary constraints of the OEM described in Section III-C.

TABLE V
COMPARISONS OF ERA-INTERIM AND OEM VS ASR WIND SPEED (M/S) OVER SIC0. OEM VS RSS WIND SPEED PRODUCT COMPARISON DONE OVER THE 2006 DATA SET.

Scenario	Bias	Stdev Diff	R
ERA vs ASR, SIC0	-0.47	2.49	0.82
OEM vs ASR, SIC0	-1.09	3.50	0.62
OEM vs RSS, 2006 DS	-2.01	2.95	0.76

Table V shows the results of comparing the optimal estimation retrieval with the ASR model and with the RSS data product. Over SIC0 scenes we have established a baseline relationship between the ASR product as reference and ERA-I data used as the first guess in the iterative process. There is no significant bias between the ERA-I wind speed product and the ASR data, together with a relatively small standard deviation of the difference at 2.5 m/s and a moderately high correlation of 0.82. For the same dataset the OEM retrieval exhibits a larger negative bias at -1.09 m/s and larger spread of the differences to the ASR model with a standard deviation of 3.5 m/s while the correlation is decreased to 0.62. The comparison with the RSS product over a larger dataset presents a much better agreement between the two retrieval products while the OEM underestimation trend is confirmed. The larger

correlation between the RSS product and the OEM retrieval are expected because of the use of a common instrument and similarities between the forward models used.

The behaviour of the retrieval over SIC0 compared to the ASR model data is exemplified in Figure 2. The two data products agree better at the lower range of wind speed values but start to diverge more in higher wind speed conditions. The forward model parametrization of wind speed effects is susceptible to larger errors at higher parameter values [1], which could explain the outliers that appear at the high end of wind speed range in the OEM retrieval. Because this parameter is highly variable in time and space, collocation differences also play a role in the disagreement between the model and the retrieval.

B. Total water vapour

The distribution of water vapour is not uniform throughout the atmosphere and this is visible in our comparison data sets. The central Arctic has a lower water vapour load than the lower latitudes and in general sea ice covered regions will be drier than open ocean ones. We compared the first guess ECMWF value with the ASR reference for both the SIC0 and SIC1 sets. Besides comparing the OEM with ASR data, we also included the AMSU-B *TWV* product over RRDP SIC1 pixels. The sanity test against RSS product is done over ice free ocean regions from the large 2006 data set.

For the SIC0 scenario, the two models match well with a low negative bias for ERA-I and a high correlation of 0.94 (Table VI). The OEM retrieved *TWV* matches fairly well
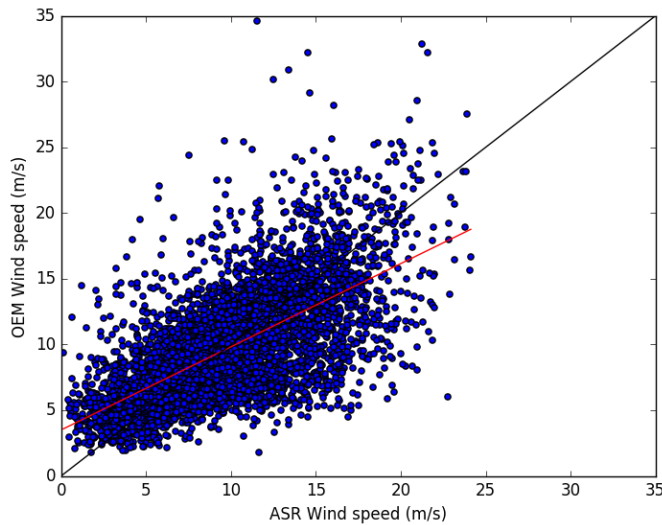
Fig. 2. Scatterplot of ASR data vs OEM retrieved wind speed for the SIC0 dataset. The identity line is shown in black and the fit of the data is show in red.
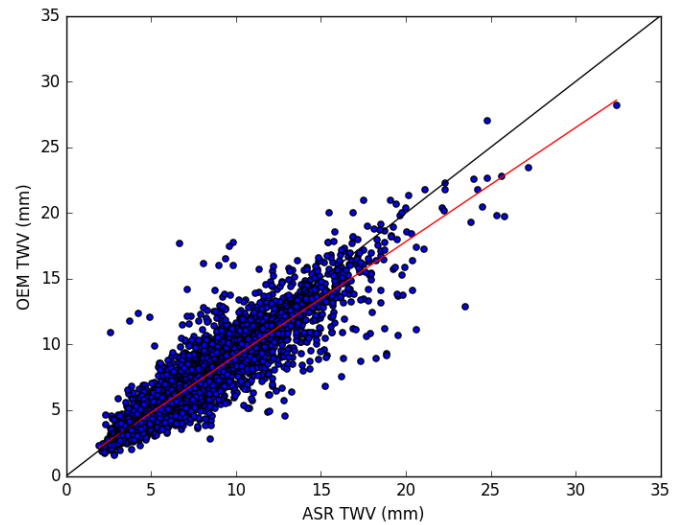

Fig. 3. Scatterplot of ASR data vs OEM retrieved *TWV* for the SIC0 dataset. The red line shows the fit of the data, while the black line represents the one to one correspondence.

TABLE VI
COMPARISONS OF ERA-INTERIM) VS ASR *TWV* (MM) OVER SIC0 AND SIC1. ERA-INTERIM AND OEM COMPARISON VS AMSU-B *TWV* OVER SIC1. OEM VERSUS RSS WATER VAPOUR PRODUCT COMPARISON DONE OVER THE 2006 DATA SET.

Scenario	Bias	StDev Diff	R
ERA vs ASR, SIC0	-0.57	1.26	0.94
OEM vs ASR, SIC0	-0.54	1.38	0.92
ERA vs ASR, SIC1	0.08	0.29	0.98
OEM vs ASR, SIC1	1.18	1.23	0.37
ERA vs AMSU-B, SIC1	-0.94	1.68	0.94
OEM vs AMSU-B, SIC1	-0.61	3.13	0.71
OEM vs RSS, 2006 DS	-0.11	0.75	0.97

TABLE VII
COMPARISONS OF ERA-INTERIM AND OEM WITH ASR *LWP* (G/M^2) OVER SIC0 AND SIC1. OEM VERSUS RSS LIQUID WATER PATH PRODUCT COMPARISON DONE OVER THE 2006 DATA SET.

Scenario	Bias	Std diff	R
ERA vs ASR, SIC0	18.8	62.9	0.60
OEM vs ASR, SIC0	87.8	75.5	0.39
ERA vs ASR, SIC1	1.7	6.3	0.67
OEM vs ASR, SIC1	145.9	64.9	0.06
OEM vs RSS, 2006 ds	58.4	48.8	0.76

with the ASR model over SIC0 with bias, difference standard deviation and correlation very close to the comparison values between the two models

This comparison can also be viewed as a scatterplot in Figure 3. We consider this comparison to be a sanity check for the OEM because in these conditions the retrieved output should not deviate too strongly from the NWP *TWV* data in conditions of no sea ice cover.

The agreement between ASR and ERA-Interim is seen over SIC1 scenes as well, with a high correlation of 0.98. The reference comparison between the OEM first guess from ERA-Interim and the AMSU-B retrieval over SIC1 shows good correlation but a high standard deviation for the very low water vapour values encountered in this region. The OEM retrieval again performed differently from the first guess data for the comparison with AMSU-B *TWV*, scoring a slightly lower bias but higher standard deviation. The moderately high correlation of 0.71 with a dedicated Arctic retrieval product is considered a good result. The best match for the OEM is again over open water with the RSS product for bias, standard deviation and correlation.

C. Liquid water path

Liquid water path has a strong impact on atmospheric absorption at the higher microwave frequencies, and the AMSR-E data thus contains valuable information on this variable.

However, this is a difficult parameter to represent in models both because of our incomplete knowledge about the complex physics involved as well as the low amounts of non precipitating liquid water present in the Arctic atmosphere. Over SIC0 for this parameter the ERA-I shows moderate agreement with ASR, with a positive bias and a moderate correlation of 0.6 (Table VII). The OEM displays a larger positive bias and similar stdev versus the ASR model together with a small correlation of 0.37. Over SIC1 the two models (ASR and ERA-Interim) show better agreement with lower values for both bias and standard deviation than over open water. The correlation coefficient increases slightly to 0.67. As in the case of the SIC0 comparison, the OEM retrieval shows large discrepancies to the ASR product with high bias and standard deviation and a very low correlation coefficient at 0.05. We need to emphasize that disagreement with the model data is not viewed as a negative result because we cannot expect models to represent correctly clouds with their high spatial and local variability at the time, place and resolution of the satellite footprint [9].

TABLE VIII
COMPARISONS OF ERA-INTERIM AND OEM WITH ASR SKIN TEMPERATURE (K) SIC0 AND SIC1. OEM VERSUS RSS SEA SURFACE TEMPERATURE COMPARISON IS DONE OVER THE 2006 DATA SET.

Scenario	Bias	Std diff	R
ERA vs ASR, SIC0	-0.12	0.29	0.97
OEM vs ASR, SIC0	-1.43	1.82	0.18
ERA vs ASR, SIC1	2.78	2.39	0.94
OEM vs ASR, SIC1	10.15	4.99	0.71
OEM vs RSS, 2006 DS	-1.09	2.18	0.69

D. Surface temperature

This parameter is included in the surface emission simulation and thus in estimating the reflection and scattering of the downwelling radiation by the surface. Because of the modifications applied to the forward model in order to include sea ice covered surfaces, temperature has been split into two parameters. One refers to the temperature of the ice free ocean surface (SST - sea surface temperature). The other refers to the sea ice temperature (IT) and is in turn connected to the empirical sea ice emissivities. The retrieval method will return both values for each pixel and in the forward calculations these temperatures are applied to the percentage of cover corresponding to each surface type. This means that retrieved ice temperatures over 0% sea ice reproduce the background. This is also the case for sea surface temperatures for data points with 100% *SIC*. In the case of partially ice covered pixels, we calculated *IT* statistics only for pixels with SIC above 75% and *SST* statistics only for pixels with SIC below 15%.

Depending on the frequency, the microwave emission of sea ice can come from the ice layer 20 cm deep at 6.9 GHz or at most a few cm deep for the 37 GHz channels in the case of first year ice. This penetration depth varies with snow cover, ice type and season. The ice temperature (IT) retrieved by the method will represent the emission of the ice and snow volume within the penetration depth. The ASR model data is not the ideal comparison tool for this parameter as the skin temperature included in the ASR model will frequently differ from the temperature of the emitting layer inside the ice because of the temperature gradient inside the snow and ice pack [17].

Over open water the penetration depth of microwaves is small enough that we can approximate the emission layer temperature for the sea surface temperature. While the two models agree well with each other (Table VIII), the OEM retrieved *SST* shows a higher negative bias of around 1.4 K and a corresponding standard deviation of 1.8 K when compared to the ASR data. Like in the case of *LWP* the retrieval shows a very low correlation with the ASR model. Unlike the comparison with the model data, the sanity test with the RSS *SST* product shows similar levels of bias and standard deviation but significantly better correlation at 0.7. Over SIC1 there is a larger discrepancy in terms of bias and standard deviation between models (ASR and ERA-Interim) but a high correlation coefficient of 0.94. The microwave emission layer temperature retrieved by the OEM shows a large positive bias and standard deviation versus the ASR skin temperature prod-

uct but a moderately high correlation at 0.7. This behaviour is expected as the emission layer inside the snow and ice volume is typically warmer than the atmosphere interface layer due to thermal insulation, while the skin temperature is closer to the atmospheric air temperature.

The systematic difference between ASR skin temperature and OEM retrieved sea ice temperature is shown in Figure 4. Following the explanation that the microwave emitting layer temperature will be significantly higher than that of the interface layer between snow/sea ice and air, there is a trend in how the OEM retrieval differs from the ASR parameter. At the lower range of values the difference between the two can be as high as 20 K while at the higher ranges of temperatures the difference decreases to within 5 K. This behaviour is consistent with the observations in [17]. During the winter season when the air temperature and the closely related skin temperature are at their lowest, the difference with the temperature deeper inside the sea ice is greatest due to the strong insulation effect of the snow layer. As this snow layer decreases and the skin temperature increases, the temperature gradient inside sea ice decreases as well.
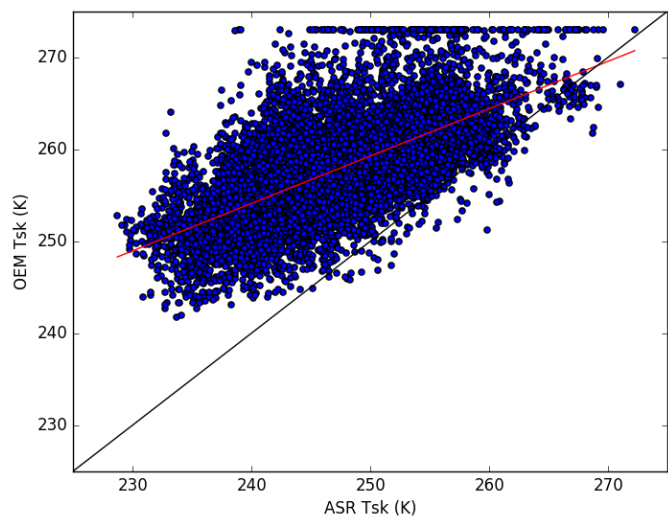

Fig. 4. Scatterplot of ASR skin temperature vs OEM retrieved sea ice temperature for the SIC1 dataset. The fit line is represented in red, while the black line is the one to one fit.

E. Sea ice concentration

Sea ice is formed from the freezing of the surface ocean water. An important consequence of this process is that new sea ice includes brine pockets which influence its microwave emission. Depending on the age of the sea ice, this salt content may vary, youngest ice having generally the highest amount and old ice the lowest. Besides sea ice salinity, the sea ice surface roughness and snow cover also influence the surface component of the microwave emission. As a simple although not perfect reference and background state for sea ice concentration, we use the NASA Team algorithm which provides both total concentration and multi-year ice fraction.

This comparison was done by looking at the mean retrieved ice concentration value as well as the standard deviation of

Scenario	Mean	Std
NT, SIC0	11.94	14.09
OEM, SIC0	1.30	1.98
NT, SIC1	94.97	3.99
OEM, SIC1	97.79	2.02

Scenario	Mean	Std
NT, SIC1	34.79	24.95
OEM, SIC1	42.19	30.53
NT-OEM, SIC1	-7.40	10.66

the data over the validated 100% and 0% scenes in the SIC1 and SIC0 datasets respectively (Table IX). For the SIC0 set, the OEM shows a significantly lower mean retrieval value of 1.3% compared to the first guess NASA Team result of almost 12%. In terms of standard deviation, the OEM retrieval shows much lower variability for SIC0 at just below 2% standard deviation, while the initial guess is a lot more noisy at 14% standard deviation. Over SIC1 scenes, the differences between the NASA Team product and the OEM retrieval are smaller in terms of mean retrieved value with 97.8% for OEM and 95% for NASA team. While the standard deviation of NASA Team first guess data is lower for 100% SIC scenes, it is still twice as large as the value for OEM SIC at 4% and 2% respectively. Compared to the other results presented in this study, the SIC1 and SIC0 results for ice concentration retrieval can be considered validation results as the RRDP consists of verified scenes of pure surface types. For an operational product however, validation with intermediary ice concentration data is still necessary. We would like to point out that the NASA Team algorithm implementation used here does not reflect the quality of the operational product based on the same algorithm. Here we used a simplified version of this method that does not include any weather filters which explains the large mean value and standard deviation exhibited over SIC0. The purpose of this comparison is to show that the OEM retrieval can improve on a lower quality first guess SIC value both in terms of absolute value and retrieval noise.

F. Multi-year ice fraction

During the melt season, most of the salt in the ice drains out and ice that remains will thus have a very low salinity. Old ice therefore has a significantly different emissivity compared to young ice. For this reason the empirical emissivities contain a set of values that apply to the fraction of old ice detected in the satellite footprint. The first guess as well as the background value of the multi-year ice fraction is again provided by the NASA Team algorithm.

Over SIC1 the NASA Team retrieved old ice fraction is lower and it shows less variability than the OEM retrieval (Table X). On their own these values are difficult to evaluate

because the reference dataset is not divided by ice type. This comparison thus serves to showcase how the OEM diverges from the NASA Team data used as first guess.

In order to better asses the skill of *MYIF* retrieval we can compare the two products above with collocated ASCAT backscatter data which is a measure of the surface and volume scattering from sea ice. Here a higher *MYIF* is expected to correlate with higher backscatter values.
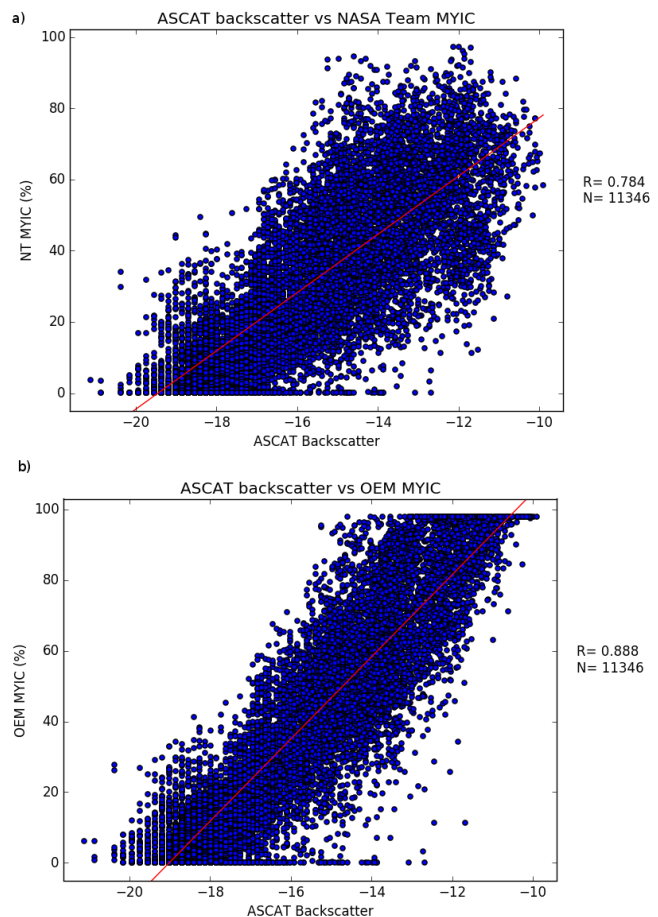

Fig. 5. ASCAT backscatter vs NASA Team (a) and OEM (b) *MYIF* retrievals.

In Figure 5 it is shown that both NASA Team and OEM *MYIF* retrievals agree with the scatterometer data at low *MYIF* values. As the backscatter value increases however, the scatter of the NASA Team product increases, resulting in an overall correlation of 0.78. The OEM retrieval agrees better with the backscatter values throughout the whole range of values and scores a better correlation at 0.89.

V. DISCUSSION

The results presented in this study aim to characterize the behaviour of the OEM versus the NWP background and first guess that are necessary components of the optimal estimation approach. The comparison does not aim to validate retrieval skill from the OEM but is a test of the ability to retrieve complementary data using a consistent passive microwave based retrieval. Validation using independent sources is a

future goal for this project, once the basic principles have been thoroughly proven.

The integrated retrieval depends on how well the forward model can account for the contribution of each of the seven geophysical parameters to the final signal. The cost function which is minimized in order to reach convergence, balances the penalty of moving away from the background values with the penalty of moving away from the observed brightness temperatures. This mechanism ensures that the method will retrieve the optimal ensemble of state vector parameters for which the forward modelled brightness temperatures are constrained by the modelling error of the measured brightness temperatures. Besides matching the equivalent brightness temperature as it would be seen by the satellite instrument, the forward model has to attribute correctly the contribution of each geophysical parameter to the composite signal. The simplifications included in the forward model will have an adverse influence on the retrieval. The two NWP products, ASR and ERA-Interim are well correlated but significantly distinct datasets which is why we used the first as a verification reference and the other as start guess value.

The relevant limitations in the case of wind speed are connected to the approximations used in the forward model. For a detailed overview of these error sources we refer the reader to the forward model description in [1]. Wind speed is one of the most variable geophysical parameters in both time and space and, besides the known deficiencies of the forward model, collocation discrepancies can also contribute to the lack of correlation between the models and the OEM. Over a large dataset and with perfect collocation because of the common instrument the OEM retrieval shows good agreement with the RSS wind speed product (Table V).

The surface temperature is used to determine the cloud temperature necessary for liquid water path calculations as well as for the water vapour regression equation. This is where a significant difference between the NWP and the MW remote sensing approach is worth mentioning. The skin temperature present in NWP data refers to the temperature of the interface layer between the surface and the atmosphere. Over sea ice this would represent the temperature of the top most layer of the snow or ice which is close to the temperature of the air above it. For the microwave radiometer however, the ice temperature represents the temperature of an emitting layer of variable thickness that depends on the different penetration depths of the channel frequencies. The penetration depth difference for AMSR-E frequencies varies between 1 cm at 89 GHz and 7 cm at 6.9 GHz for first-year ice and between 4 cm and 30 cm for multi-year ice [17]. Because of the temperature gradient inside the ice pack, the difference between the surface air temperature and the emitting layer temperature will increase with decreasing channel frequency. Because each channel "sees" a slightly different surface temperature the final output temperature is the value that minimizes the cost function for that particular data point and can be understood as a composite value depending on the different channel specific emission layer temperatures and the background ice surface temperature (Table VIII).

The individual contribution from the other two atmospheric

parameters of total columnar water vapour and liquid water path are both connected in the forward model with the surface temperature. The bias in retrieved ice temperature can be connected to the positive bias in water vapour retrieval (Table VI) through the water vapour - effective air temperature regression equations used in the forward model. Because the effective air temperature is derived from the surface temperature and water vapour is likewise dependent on the effective air temperature it is realistic to expect the surface temperature bias to be transmitted to the water vapour values. When comparing the OEM water vapour retrieval with the AMSU-B product it is important to note that the AMSU-B instrument uses three channels around the strong water absorption line at 183 GHz. The AMSR-E instrument used by the OEM has two channels around the weak water absorption line at 22 GHz, so that the information content available to the two methods is very different, with higher sensitivity to low water vapor values available to the AMSU-B method. The comparison between the two retrievals shows relatively good correlation and small bias. The match is however worse than between ERA-Interim and AMSU-B *TWV*. As a separate observation to the difference in sensitivity to low *TWV* values between the two methods it is worth mentioning that the forward model used by the OEM was tuned using global water vapour ranges, which are generally higher than those present in the Arctic.

A similar situation exists in the case of liquid water path where the cloud temperature is calculated by the forward model as the mean between surface and freezing temperatures. An overestimation of the surface temperature will lead to an overestimation of the cloud temperature which implies an overestimation in liquid water path (Table VII) who's absorption characteristics are inversely correlated with cloud temperature [18]. A second error source related to the cloud temperature is that the current calculation technique will always lead to cloud temperatures below the freezing point over sea ice regions. While supercooled clouds are frequent, there is no modeling in the forward model about the percentage to which they are frozen. About ice clouds the model assumes negligible absorption and scattering. This leads to brightness temperature errors at higher frequencies. A possible solution for the cloud temperature parametrization is also suggested in [18], by using data from the Moderate Resolution Imaging Spectroradiometer (MODIS) cloud top temperature product to adapt the parametrization to typical cloud temperatures expected in the retrieval region.

For all atmospheric parameters retrieved over open ocean, the OEM retrieval matches moderately well to very well with the RSS product because both are matching observed brightness temperatures from the same instrument. This comparison shows that the OEM is more sensitive to the brightness temperature measurements than to the NWP a priori data which is encouraging as the measurements contain better information than the models. While the RSS product is a standard in passive microwave remote sensing, it can only be applied over open ocean areas while the aim of the OEM is to provide atmospheric retrieval in pure sea ice and mixed surface conditions as well.

The sea ice concentration retrieval provides a reasonable

level of accuracy for absolute values of sea ice cover in the SIC1 and SIC0 datasets (Table IX). The corrected set of averaged sea ice emissivities represents a compromise solution for use over the SIC1 dataset. The overall differences between simulated and observed brightness temperatures has been significantly reduced compared to the initial version that uses the uncorrected emissivity set (Table IV), however at a local level the mean emissivity values might not be representative for a particular sea ice region and this discrepancy will affect the retrieved sea ice concentration value as in all other sea ice concentration algorithms using fixed tie-points. The inclusion of only two ice types with fixed signatures in time and space is a necessary simplification that will lead to errors when the real sea ice signature is significantly different from these. Ideally a more elaborate snow/ice emissivity model should be included that can adequately represent the surface variability. For the moment the large number of parameters required by such a model make it an impractical approach.

From the comparison with ASCAT backscatter values it is apparent that the OEM *MYIF* retrieval provides a better agreement with scatterometer data than the NASA Team product (Figure 5). While this is not a validation for the true multi-year ice fraction, differences between young ice and old ice scattering properties represent one criterion to distinguish between the two ice types. The OEM results are more stable throughout the whole *MYIF* range of values while the NASA Team result shows a larger scatter for high ASCAT backscatter.

VI. CONCLUSIONS

The aim of this study was to investigate the behaviour of the OEM retrieval over pure surface types and determine its sensitivity to different a priori constraints. We have shown that the background error covariance matrix has a strong influence over the rate of convergence. Using climatological sources for the diagonal elements of this matrix results in more iterations necessary to reach convergence and lowers the overall number of convergent cases. Using a full covariance matrix that includes off diagonal elements derived from the same climatological source has a negative effect on the retrieval, while also slightly reducing the number of convergent cases.

The differences between results using a static versus a collocated NWP start guess for each data point are small, demonstrating the stability of the results against different starting conditions. The difference in number of convergent cases is minimal, with a slight advantage for the static case together with a small increase in the average number of iterations required to reach convergence.

Testing different versions of the brightness temperature covariance matrix proved a large influence both on the retrieval and on the convergence statistics. After including the channel wise combined modelling error, forward model parameter error and the measurement error into this matrix, the agreement with ASR improved. This is expected behaviour because by increasing the variances for the brightness temperature the constraint of matching the measurements is relaxed. Without implementing the modelling error however the retrieval would be subjected to potential bias contamination from the forward model.

Using the initial empirical emissivities from [17] for retrieval over sea ice areas resulted in large biases between the simulated and observed brightness temperatures. We modified the emissivities towards minimising these biases resulting in a new set of winter sea ice emissivities. The corrected emissivities have decreased the differences between model and measurements to an average of 1 K from an average of 2.8 K.

We have compared OEM retrieval results over pure 100% and 0% sea ice concentration scenes with ASR model data. Despite the fact that the first guess data from ERA-Interim and the ASR product match very well, in sea ice covered regions the OEM does not always agree with the models. The fact that the sanity test comparison over ice free areas against both ASR model data and the RSS product show good agreement suggests that the OEM is strongly constrained to match the measured brightness temperatures and not the NWP data used as a priori. This approach shows promise once the deficiencies of the sea ice emissivity modeling are reduced.

There is a discrepancy between the surface temperature product from NWP and the sea ice emission layer temperature that can be retrieved from passive microwave radiometers. Over sea ice, this discrepancy is quite large and is expected to influence the retrieval of liquid water path and water vapour. This difficulty may be overcome in the case of liquid water path by employing a new parametrization for cloud temperature, based on the MODIS cloud top temperature product as is suggested in [18]. For water vapour, the regression curves used currently by the forward model can be tested against radiosonde profiles measured in the Arctic and updated accordingly. The passive microwave retrieved surface temperature has to be seen as different from the surface physical temperature as it is connected to the emitting layer temperature. This parameter can yield more useful information by separating the current unique value by channel, depending on frequency penetration depth, by using the empirical regression equations from [17]. A list of consistent, frequency dependent, emission layer temperatures can be retrieved like this and included in an improved sea ice forward model.

For simplicity, in this study we only used an OEM setup that used the ten AMSR-E channels between 6.9 and 36.5 GHz resampled at the common resolution of 56 km. In order to improve the product resolution in the future the 6.9G Hz, 10.7 GHz and 18.7 GHz channels can be excluded from the retrieval and the method can produce output at the 21 km spatial resolution of the 23.8 GHz channels. Depending on the number of desired parameters, different combinations of channels can be used for balancing frequency dependent sensitivity with retrieval resolution. For determining these optimal pairings between measurement input and retrieval output, a channel wise parameter sensitivity study is planned. Connected to this, but with the aim of improving atmospheric sensitivity, the 89 GHz channels can be included in the retrieval once the forward model underestimation is corrected. Preliminary testing has shown that both excluding the low frequency channels as well as including the 89 GHz channels are useful approaches that give reasonable results.

APPENDIX

A priori constraints and ancillary data used in the OEM

TABLE XI
DIAGONAL ELEMENTS OF BACKGROUND ERROR COVARIANCE MATRIX - S_p. SHOWN HERE ARE THE STANDARD DEVIATIONS.

Parameter	$\sqrt{\sigma^2}$
WS [m/s]	3.5
TWV [mm]	3.3
LWP [g/m^2]	142.8
SST [K]	4.9
IT [K]	4.9
SIC [%]	31.6
MYIF [%]	54.7

TABLE XII
DIAGONAL ELEMENTS OF THE BRIGHTNESS TEMPERATURE COVARIANCE MATRIX - S_e. SHOWN HERE ARE THE STANDARD DEVIATIONS (K).

Channel	$\sqrt{\sigma^2}$
6.9V	1.68
6.9H	3.46
10.7V	1.53
10.7H	3.71
18.7V	1.31
18.7H	3.27
23.8V	0.98
23.8H	2.57
36.5V	1.81
36.5H	2.52

TABLE XIII
MEAN WINTER EMISSIVITIES FOR ALL AMSR-E CHANNELS. COLUMN A: INITIAL SET OF VALUES DERIVED FOR SIC1. COLUMN B: CORRECTED SIC1 EMISSIVITIES.

Channel	A	B
First year ice		
6.9V	0.954	0.976
6.9H	0.854	0.893
10.7V	0.953	0.976
10.7H	0.860	0.902
18.7V	0.964	0.975
18.7H	0.875	0.903
23.8V	0.960	0.972
23.8H	0.875	0.900
36.5V	0.936	0.955
36.5H	0.851	0.880
89V	0.864	0.817
89H	0.809	0.747
Multi-year ice		
6.9V	0.955	0.977
6.9H	0.861	0.874
10.7V	0.930	0.952
10.7H	0.840	0.848
18.7V	0.884	0.887
18.7H	0.808	0.798
23.8V	0.848	0.841
23.8H	0.775	0.764
36.5V	0.761	0.732
36.5H	0.699	0.675
89V	0.774	0.656
89H	0.731	0.622

TABLE XIV
ELEMENTS OF THE BACKGROUND STATE VECTOR. THE VALUES FOR ICE TEMPERATURE, SEA ICE CONCENTRATION AND MULTI-YEAR ICE FRACTION ARE NOT FIXED BUT EVALUATED ONCE FOR EVERY DATA POINT.

Parameter	$\sqrt{\sigma^2}$
WS [m/s]	4.11
TWV [mm]	2.86
LWP [mm]	0.16
SST [K]	274.5

REFERENCES

[1] F. J. Wentz and T. Meissner, "AMSR ocean algorithm," Algorithm Theoretical Basis Document (ATBD). version 2, nov 2000.

[2] L. T. Pedersen, "Merging microwave radiometer data and meteorological data for improved sea ice concentrations," *EARSeL Advances in Remote Sensing*, vol. 3, no. 2-XII, pp. 81–89, 1994.

[3] C. Kongoli, S.-A. Boukabara, B. Yan, F. Weng, and R. Ferraro, "A new sea-ice concentration algorithm based on microwave surface emissivitiesapplication to amsu measurements," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 175–189, 2011.

[4] C. Melsheimer, G. Heygster, and L. T. Pedersen, "Integrated retrieval of surface and atmospheric parameters over the Arctic from AMSR-E satellite microwave radiometer data using inverse methods," in *IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium*, vol. 4.   IEEE, 2008, pp. IV–986.

[5] P. Ashcroft and F. Wentz, "Algorithm theoretical basis document, AMSR Level 2A algorithm," *RSS, Santa Rosa, CA, Tech. Rep*, vol. 121, 2000.

[6] D. Dee, S. Uppala, A. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. Balmaseda, G. Balsamo, P. Bauer *et al.*, "The era-interim reanalysis: Configuration and performance of the data assimilation system," *Quarterly Journal of the royal meteorological society*, vol. 137, no. 656, pp. 553–597, 2011.

[7] L. T. Pedersen and R. Saldo, *Sea Ice Concentration (SIC) Round Robin Data Package*, November 2012, doc Ref:SICCI-RRDP-06-12, Version: 1.0.

[8] N. Ivanova, L. Pedersen, R. Tonboe, S. Kern, G. Heygster, T. Lavergne, A. Sørensen, R. Saldo, G. Dybkjær, L. Brucker *et al.*, "Inter-comparison and evaluation of sea ice algorithms: towards further identification of challenges and optimal approach using passive microwave observations," *The Cryosphere*, vol. 9, 2015.

[9] D. H. Bromwich, A. B. Wilson, L.-S. Bai, G. W. Moore, and P. Bauer, "A comparison of the regional Arctic System Reanalysis and the global era-interim reanalysis for the arctic," *QJR Meteorol. Soc*, vol. 142, pp. 644–658, 2016.

[10] C. Melsheimer and G. Heygster, "Improved retrieval of total water vapor over polar regions from AMSU-B microwave radiometer data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 8, pp. 2307–2322, 2008.

[11] P. Ashcroft and F. J. Wentz, "Amsr-e/aqua l2a global swath spatially-

resampled brightness temperatures, version 3," *Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center*, 2003.

[12] F. J. Wentz, "A well-calibrated ocean algorithm for Special Sensor Microwave / Imager," *J. Geophys. Res.*, vol. 102(C4), pp. 881–894, 1997.

[13] D. K. Smith, C. Mears, K. A. Hilburn, and L. Ricciardulli, "A 25-year satellite microwave mean total precipitable water data set for use in climate study," EUMETSAT Meteorological Satellite Conference, 2013.

[14] M. Szczodrak, P. J. Minnett, and C. Gentemann, "Comparison of AMSR-E Retrievals of Total Water Vapor over the Ocean with Ship based Measurements," 10th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS), 2006.

[15] C. D. Rodgers, *Inverse methods for atmospheric sounding: theory and practice*. World scientific, 2000, vol. 2.

[16] T. Markus and D. J. Cavalieri, "An enhancement of the nasa team sea ice algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 3, pp. 1387–1398, 2000.

[17] N. Mathew, G. Heygster, and C. Melsheimer, "Surface emissivity of the arctic sea ice at amsr-e frequencies," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 12, pp. 4115–4124, 2009.

[18] C. Seethala and Á. Horváth, "Global assessment of AMSR-E and MODIS cloud liquid water path retrievals in warm oceanic clouds," *Journal of Geophysical Research: Atmospheres*, vol. 115, no. D13, 2010.