Combining AMSR-E and QuikSCAT image data to improve sea ice classification

Peter Yu and David A. Clausi
Systems Design Engineering
University of Waterloo
{*p5yu, dclausi*}*@uwaterloo.ca*

Roger De Abreu and Tom Agnew
Environment Canada
{*Roger.DeAbreu, Tom.Agnew*}*@ec.gc.ca*

Abstract

The benefits of augmenting AMSR-E image data with QuikSCAT image data for supervised sea ice classification in the Western Arctic region are investigated. Experiments compared the performance of a maximum likelihood classifier when used with the AMSR-E only data set against the combined data and examined the preferred number of features to use as well as the reliability of training data over time. Adding QuikSCAT often improves classifier accuracy in a statistically significant manner and never decreased it significantly when enough features are used. Combining these data sets is beneficial for sea ice mapping. Using all available features is recommended and training data from a specific date remains reliable within 30 days.

1. Introduction

The mapping of sea ice is an important task for understanding global climate and for safe navigation. The Canadian Ice Service (CIS) is tasked with creating operational ice charts that map the sea ice around Canada. These charts identify different ice classes geographically and are created with the help of remotely sensed data, such as RADARSAT images. To help improve sea ice mapping, information from other satellite systems needs to be evaluated.

In this work, combining AMSR-E passive microwave and QuikSCAT SeaWinds scatterometer data is considered. Passive microwave data has been used for ice typing [12] and QuikSCAT shows promise as well [5]. Recent work has shown that combining passive microwave and scatterometer data is useful for sea ice classification and mapping [9, 11]. However, little work has specifically combined AMSR-E and QuikSCAT data or compared the performance of using the multisensor data set against the individual data sets.

These issues are examined here with a pattern recognition framework to generate quantitative results that can lead to developing an automated ice mapping system. In particular, the following research questions are posed:

1. Does adding QuikSCAT to AMSR-E provide a significant improvement in sea ice classification?
2. How does the number of preferred features used affect classification?
3. How reliable is the training data over time?

The results will indicate whether the AMSR-E and QuikSCAT sensor combination should be investigated in more detail for use in sea ice classification.

2. Data

The AMSR-E data used is the daily average Level 3, 12.5 km brightness temperature product [2]. This product has a polar stereographic projection and consists of 8 bands: 18 GHz, 23 GHz, 36 GHz and 89 GHz in both horizontal and vertical polarizations. The QuikSCAT data used is the near real-time daily average H-pol Arctic σ^o product (NHEAVEH) obtained online from NOAA/NESDIS [3]. The projection of the AMSR-E data matches the QuikSCAT data so the only registration needed is scaling to a common resolution (done via nearest neighbour interpolation to preserve data values). The dates in 2004 to 2005 were chosen since they provide enough ground truth samples for training ($\geq$ 100 per class for 9 features, following the guideline of $\geq$ 10 samples per feature [6]) in the study area of the Western Arctic around the Beaufort sea.

Ice charts produced by the CIS are used as ground truth. The ice charts define regions which CIS operators have manually classified. The CIS defined ice types

978-1-4244-2653-9/08/$25.00 ©2008 IEEE

were aggregated into five classes for these experiments by grouping similar CIS stage of development codes. This ensures that there are enough samples for each class. Table 1 maps the CIS codes to the class names used. Since the ice chart regions can have mixed ice types, only regions with $\geq$ 70% concentration of one of the classes are considered ground truth to ensure that only representative samples are used.

Table 1. CIS codes mapped to class name.

CIS Code	Class
1, 2	Thin Ice
3, 4, 5	Medium (Med) Ice
6, 7, 4., 1., 8, 9	First Year (FY) Ice
7., 8., 9.	Multiyear (MY) Ice
Open Water	OW

3. Methods

A maximum likelihood classifier with multivariate Gaussian class distributions was used for all experiments. The intent is to evaluate the data sources under a common classifier rather than designing the best classifier for the data. AMSR-E and QuikSCAT data have different units so the bands were normalized to $[0, 1]$ before being used as features, ensuring that each feature is weighted fairly in distance calculations [6]. Except for tests in Section 4.3 (explained later), each date was trained and validated independently with random selection of training and testing samples by consulting the CIS chart for that date.

The experiments required selecting feature subsets to use in classification. Sequential Floating Forward Search (SFFS) [8] was implemented but was modified to sequential forward search (SFS) which adds unselected features one by one, choosing the one that forms the best subset with the already selected features. Differences in classification results between the combined data set and the individual data sets are easier to interpret with SFS since they are directly due to including QuikSCAT. For example, if QuikSCAT is not in the n features chosen from the combined data set, the selected bands are the same n bands chosen from AMSR-E alone. This is not true for SFFS: a different set of n AMSR-E bands may be chosen that can have a different classification result. The accuracy difference between SFS and SFFS was not statistically significant for the data used here, so SFS was chosen. To evaluate subsets, the transformed divergence [10] between the least separable classes is maximized.

Kappa (κ) and its confidence interval σ [1] are used to evaluate classifier accuracy. When converted to Z-values [4], they can indicate whether the difference between two classification rates is statistically significant.

4. Results

4.1. Question 1: Improvements due to adding QuikSCAT

Table 2 shows Z-values which compare the classification using only AMSR-E data and using combined AMSR-E + QuikSCAT data at each feature subset size. It compares the "best" (as chosen by SFS) n features of AMSR-E against the best n features of AMSR-E + QuikSCAT (the last column compares all eight AMSR-E with all nine AMSR-E + QuikSCAT bands). For the most part, adding QuikSCAT either makes a statistically significant (at the 95% level) improvement over AMSR-E alone or does not hurt the performance. A few rare cases show a significant decrease but these are for smaller feature set sizes, which (as shown later) perform worse than using all features. The last column shows that when all nine features are used, adding QuikSCAT improves the classification rate, with more than half showing a statistically significant increase.

Table 2. Z-values comparing AMSR-E + QS and AMSR-E at each feature set size.

	Feature Set Size								
Dates	1	2	3	4	5	6	7	8	9vs8
20040906	**-2.00**	-0.99	-0.52	1.23	1.09	0.53	-0.09	0.61	0.38
20041004	**-8.80**	**4.40**	**3.25**	**6.35**	**4.25**	**2.00**	**3.39**	**3.81**	**2.66**
20041025	**23.25**	**3.51**	1.71	**2.26**	0.59	1.41	0.56	1.49	1.61
20041108	-	-	-	-	-	-	1.39	1.89	**2.31**
20050301	-	**-2.23**	-0.68	0.89	-1.29	-0.08	0.45	0.09	-0.09
20050601	**6.23**	**7.02**	**6.58**	**5.25**	**3.31**	**4.83**	**2.56**	**3.56**	**3.90**
20050613	**6.27**	-0.65	**-2.62**	-1.75	0.03	-0.19	0.87	-0.05	0.90
20050620	-	**8.66**	**8.66**	**10.14**	**11.01**	**9.70**	**10.12**	**10.23**	**10.27**
20050822	**11.33**	-1.12	-0.67	1.79	-0.48	**3.01**	0.67	1.10	**2.11**
20050829	-	-	-	-	-	-	**2.98**	**2.50**	**3.37**
20050815	-	-	-	-	0.95	0.03	-0.70	0.15	1.00

Bold black = AMSR-E + QS better; **Bold red** = AMSR-E better

Unbolded = Not significant; Blank = QS not chosen by SFS.

Table 3 shows κ for three different cases: using all eight AMSR-E, using QuikSCAT alone and using all nine combined bands. This gives an idea of the classification performance via κ and shows that the combined data set is better than either AMSR-E or QuikSCAT alone. Also of note is that QuikSCAT alone cannot match the accuracy of AMSR-E.

Table 4 shows the classification (producer's) accuracy of the individual ice classes for each date, comparing the results from all 8 AMSR-E bands and all 9 combined bands. It shows consistent improvements due to QuikSCAT (highlighted in green) for all classes, pre-

Table 3. κ values for full feature sets.

Date	κ AMSR-E	QS	AMSR-E+QS
20040906	0.87	0.82	0.87
20041004	0.78	0.64	**0.81**
20041025	0.62	0.60	0.64
20041108	0.57	0.47	**0.61**
20050301	0.88	0.87	0.88
20050601	0.63	0.59	**0.69**
20050613	0.59	0.39	0.60
20050620	0.40	0.41	**0.57**
20050822	0.62	0.40	**0.66**
20050829	0.66	0.45	**0.71**
20050815	0.71	0.33	0.73

Bold = AMSR-E+QS better (statistically significant)

senting a strong case that QuikSCAT is a good complement to AMSR-E.

Table 4. Accuracy of each class using AMSR-E (A) vs. combined (A+QS) data.

Dates	Data	Accuracy (%) Thin	Med	FY	MY	OW
20040906	A	-	-	79	91	96
	A+QS	-	-	78	91	97
20041004	A	70	87	-	85	92
	A+QS	74	88	-	88	93
20041025	A	72	60	89	78	96
	A+QS	73	65	90	78	96
20041108	A	-	78	82	80	-
	A+QS	-	82	83	82	-
20050301	A	-	-	94	95	-
	A+QS	-	-	94	94	-
20050601	A	-	-	85	73	95
	A+QS	-	-	86	78	96
20050613	A	-	-	69	77	97
	A+QS	-	-	70	79	97
20050620	A	-	-	51	71	95
	A+QS	-	-	77	70	96
20050822	A	-	-	88	67	96
	A+QS	-	-	89	70	97
20050829	A	-	-	79	74	95
	A+QS	-	-	85	78	97
20050815	A	-	-	77	80	96
	A+QS	-	-	82	80	97

AMSR-E + QS better AMSR-E better

Spatially, the AMSR-E 8 band and AMSR-E + QuikSCAT 9 band results are compared in Figure 1. Both data sets allow correct identification of many pixels (gray) and incorrect identification of others (red). However, AMSR-E + QuikSCAT is more often correct (blue) than AMSR-E alone (yellow). QuikSCAT again appears to be complementary to AMSR-E.

4.2. Question 2: Number of preferred features

This section examines how the number of features used affect classification with the AMSR-E + QuikSCAT data set. Table 5 shows Z-value comparisons between classifications using all 9 bands and those using subsets of 1 to 8 bands. The full feature set never has a statistically significant reduction in accuracy com-

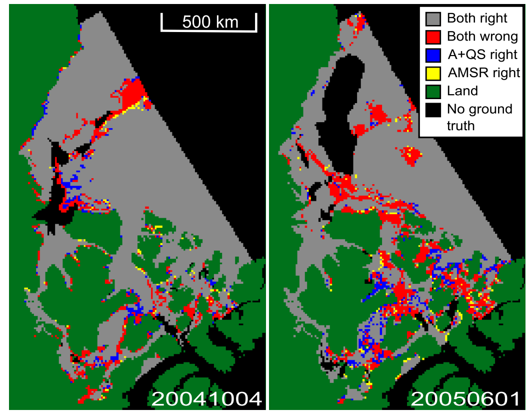

Figure 1. Spatial comparison of accuracy; colours show how each data set performed when classifying each pixel.

pared to any of the feature subsets. With a reduced number of features, there are more cases of significant increases in accuracy due to using the full feature set. Table 5 indicates that the data is not suffering from the "curse of dimensionality" [7] and that the 9 band feature set results in improvements over the smaller subsets. Therefore, the full feature set should be used to take maximum advantage of all the information available. This is important because feature selection does not always choose the QuikSCAT band even though it offers additional information, as shown earlier by the dashes in Table 2.

Table 5. Z-values comparing the full AMSR-E + QS feature set with its subsets.

Dates	Feature Subset Size 1	2	3	4	5	6	7	8
20040906	**4.45**	**2.56**	**1.98**	0.86	0.70	0.21	0.27	-0.23
20041004	**13.63**	**4.99**	**4.10**	0.36	-0.07	-0.68	-0.95	-1.15
20041025	**3.29**	**4.92**	**2.42**	**2.30**	**2.30**	1.33	1.13	0.12
20041108	**11.62**	**4.46**	1.68	**2.62**	**3.03**	**3.82**	1.10	0.42
20050301	-1.41	0.17	0.13	0.25	0.80	-0.72	-0.49	-0.17
20050601	**6.87**	**4.90**	**4.96**	**3.77**	**4.63**	**3.03**	**2.29**	0.35
20050613	**13.17**	**9.67**	**11.02**	**6.25**	**3.35**	**2.45**	1.04	0.95
20050620	**23.08**	**8.92**	**6.22**	**3.63**	**2.89**	**2.21**	-0.03	0.04
20050822	**16.87**	**12.34**	**8.68**	**5.88**	**6.25**	1.28	**2.03**	1.00
20050829	**22.39**	**10.58**	**12.95**	**8.19**	**6.08**	**3.63**	0.62	0.87
20050815	**3.36**	**7.15**	**3.39**	**2.10**	1.47	1.80	1.68	0.84

Bold black = Full feature set better; Unbolded = Not significant.

4.3. Question 3: Reliability of training data over time

In the previous experiments, each date was trained independently. In this section, training is done with all samples from one date and applied to classify other dates in 2004 to 2005, with one date being classified at a time. The test is repeated by training with each date that has enough training samples. This experiment reveals the applicability of the training data over time. The classification results are binned by the absolute number of days between the date of the training data and the date being classified (ΔDays). The bin size is 15 days. Figure 2 shows the mean κ and its standard deviation in each bin as a function of ΔDays. As expected, accuracy decreases with increasing ΔDays. The AMSR-E + QuikSCAT data set has higher average κ for larger ΔDays, suggesting more time invariance but with large ΔDays, κ has a higher spread with unacceptable negative κ values. This test shows that training data within 30 days of the date being classified is acceptable, an important consideration if a database of training data is to be created for automated classification.

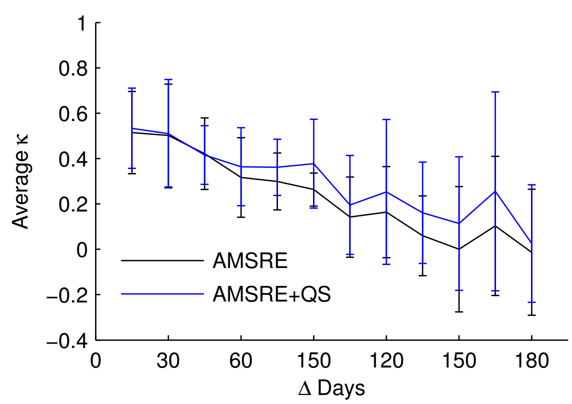

Figure 2. Average κ obtained for dates that are ΔDays from training dates.

5. Conclusions

This work has shown that the combined data set is statiscally better than either AMSR-E or QuikSCAT data alone for sea ice classification. The data appear to be complementary, providing improvements in all ice types. Using the full 9 features of AMSR-E + QuikSCAT is recommended since this uses all of the available data without suffering from dimensionality problems. Training data for this data set appear to be valid for only about a month, so training databases will have to be designed accordingly. Although the findings here are strictly for the western Arctic area, preliminary tests that consider the entire Arctic show similar results. This will be investigated in more detail.

Future work should look at improving the absolute classification rate. This may be obtained by designing a different classifier or by using another method of data fusion. QuikSCAT provides additional information; the work that remains is how to make the best use of it.

6. Acknowledgements

The authors thank CIS staff and Stephen Howell for data and advice. The Natural Sciences and Engineering Research Council of Canada and the Canadian Federal Government's International Polar Year are thanked for financially supporting this project.

References

[1] T. Bishop, S. Fienberg, and P. Holland. *Discrete Multivariate Analysis - Theory and Practice*. MIT Press, Cambridge, 1975.

[2] D. Cavalieri and J. Comiso. *AMSR-E/Aqua Daily L3 12.5-km Brightness Temperature (Tb) Polar Grids*. National Snow and Ice Data Center., Boulder, CO, 2004-2005.

[3] P. Chang. QuikSCAT ice page, 2000. http://manati.orbit.nesdis.noaa.gov/cgi-bin/qscat_ice.pl.

[4] R. G. Congalton, R. G. Oderwald, and R. A. Mead. Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogram. Eng. Remote Sens.*, 49(12):1671–1678, Dec. 1983.

[5] R. de Abreu, K. Wilson, M. Arkett, and D. Langlois. Evaluating the use of QuikSCAT data for operational sea ice monitoring. In *Proc. Int'l Geosci. and Remote Sensing Symposium*, Toronto, Canada, June 2002.

[6] R. Duda, P. Hart, and D. Stork. *Pattern Classification*. Wiley, New York, second edition, 2001.

[7] A. Jain and B. Chandrasekaran. Dimensionality and sample size considerations in pattern recognition practice. In P. Krishnaiah and L. Kanal, editors, *Handbook of Statistics*, volume 2, pages 835–855. North-Holland, Amsterdam, 1982.

[8] P. Pudil, J. Novovičová, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15(11):1119–1125, Nov. 1994.

[9] Q. P. Remund, D. G. Long, and M. R. Drinkwater. An iterative approach to multisensor sea ice classification. *IEEE Trans. Geosci. Remote Sensing*, 38(4):1843–1856, July 2000.

[10] R. Schowengerdt. *Remote Sensing, models and methods for image processing*. Academic Press, San Diego, second edition, 1997.

[11] R. Tonboe and L. Toudal. Classification of new-ice in the Greenland Sea using satellite SSM/I radiometer and SeaWinds scatterometer data and comparison with ice model. *Rem. Sens. Env.*, 97(3):277–287, Aug. 2005.

[12] M. Wensnahan, G. Maykut, T. Grenfell, and D. Winebrenner. Passive microwave remote sensing of thin ice using principal component analysis. *J. Geophysical Research*, 98(C7):12453–12468, July 1993.