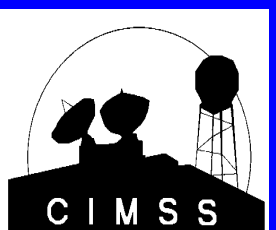
MODIS Cloud Classification via the Local Region of Interest (LROI) Scheme

Bormin Huang and Steven A. Ackerman
Cooperative Institute for Meteorological Satellite Studies,
University of Wisconsin–Madison

W. Paul Menzel
NOAA, National Environmental Satellite, Data, and Information Service



Background

- NWP centers assimilate clear radiances.
- Pixels containing undetected cloud radiances contaminate the geophysical product.

  For example, only a few percent of cloud in a FOV can introduce unacceptable errors in SST.
- Need good cloud detection/classification algorithms!!

MODIS Cloud Classification

- The MODIS (Moderate Resolution Imaging Spectroradiometer) is a key instrument aboard the NASA EOS Terra and Aqua satellites, viewing the entire Earth's surface every 1 to 2 days.
- It provides 36 infrared and visible bands ranging from 0.4 to 14.4 μm, with high spatial resolution from 250 m to 1 km.
- The high spatial resolution of MODIS observation provides valuable information for cloud classification !!!

MODIS Cloud Classification Schemes

- The unsupervised maximum likelihood (ML) scheme (*Li. et al., 2003*)
- The operational MODIS cloud mask algorithm (*The MODIS team at CIMSS, Univ. of Wisconsin, 2003*)
- The multicategory support vector machine (MSVM) scheme (*Lee, Wahba, Ackerman, 2004*)
- **The local region of influence (LROI) scheme (*new!*)**

Data Description

- 1536 MODIS scenes over the Gulf of Mexico in July 2002 were classified as clear, ice, or water clouds by a satellite expert.
- There were 256 clear scenes, 952 ice clouds and 328 water clouds.
- Each of the three categories were randomly divided in half. The first halves are used as a training set, and the second halves as a testing set.
- We used the same data used in the MSVM paper by *Lee, Wahba, and Ackerman (2004).*

Testing Data (Truth)

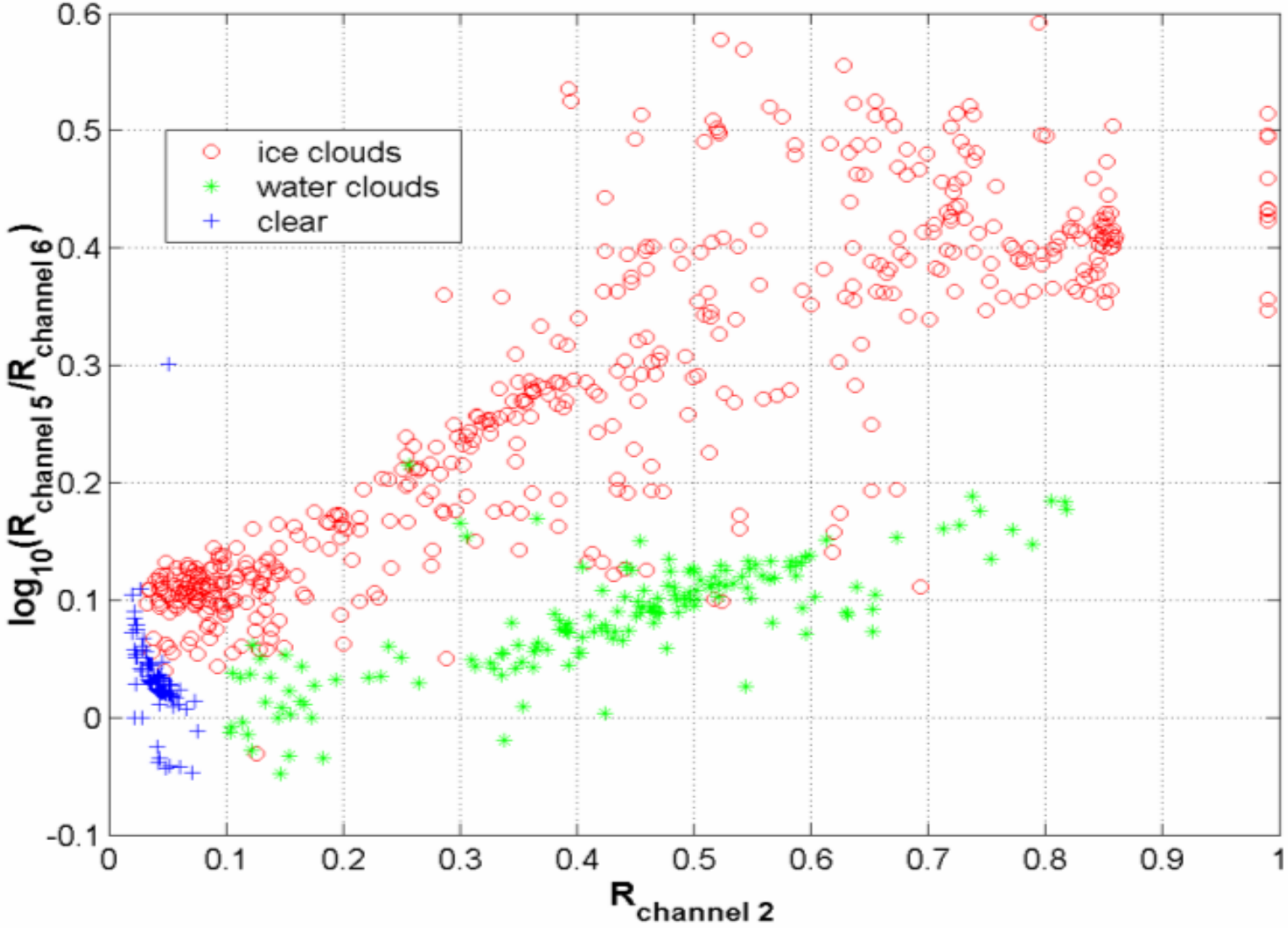


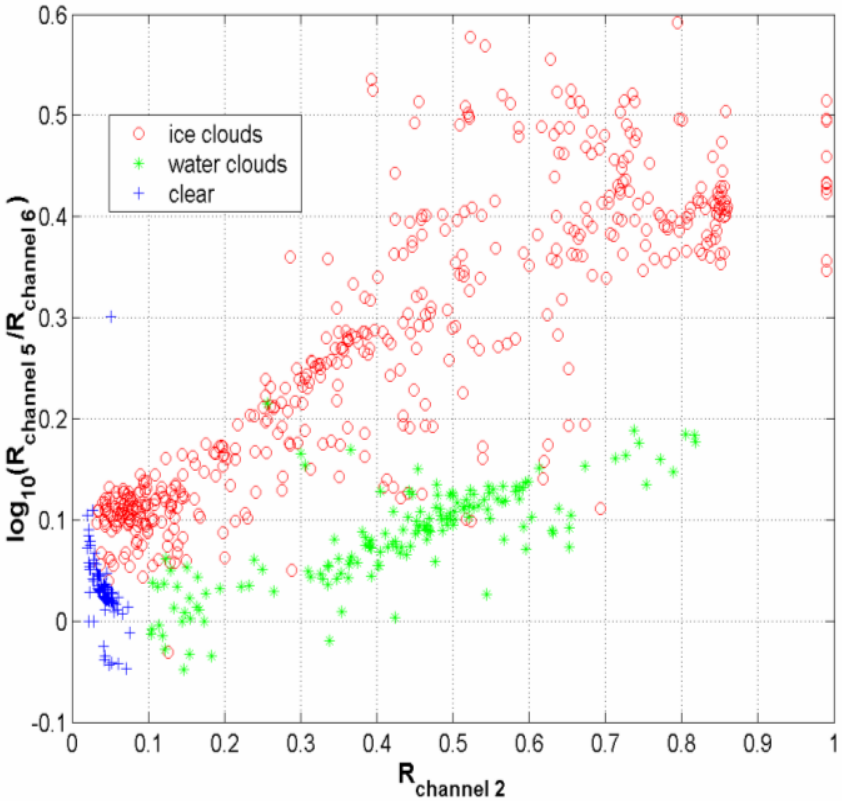
Fig. 1. Scattereplot of $\log_{10}$ ($R_{\text{channel 5}}$ / $R_{\text{channel 6}}$) vs. $R_{\text{channel 2}}$.

The unsupervised ML classification method

Not good for the nonconvex boundary!

Table 1 Distribution of predicted category vs. true category after the ML classification. The misclassification error rate is 481/768 = 62.63%.

True category	Predicted category			Total
	Clear sky	Ice clouds	Water clouds	
Clear sky	128	0	0	128
Ice clouds	173	150	153	476
Water clouds	32	123	9	164



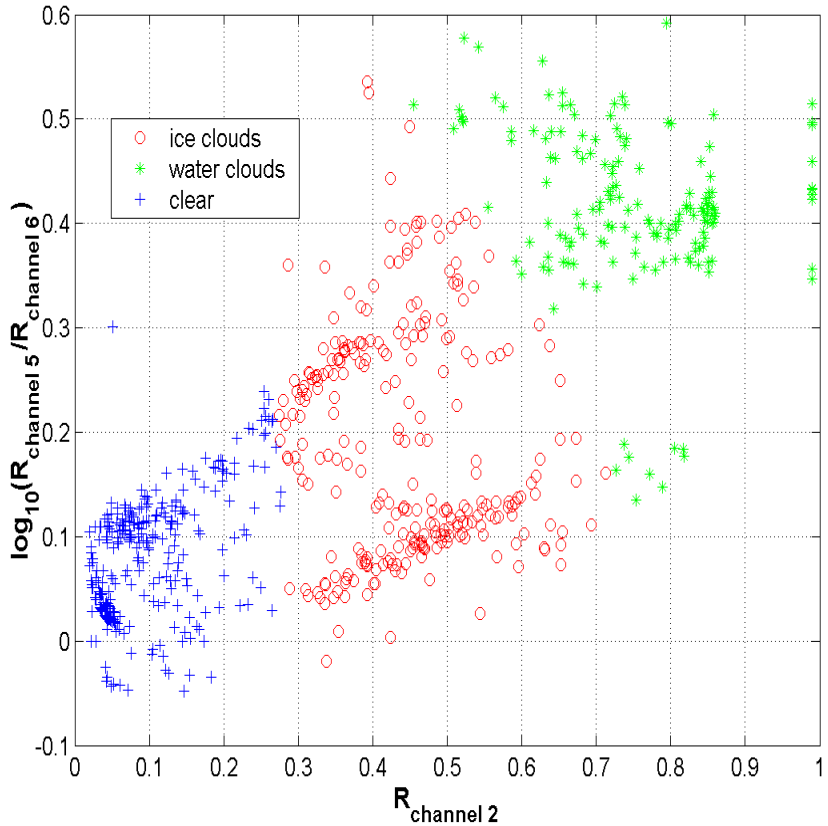


Fig. 1. Scattereplot of $\log_{10}(R_{channel\ 5} / R_{channel\ 6})$ vs. $R_{channel\ 2}$.

The LROI scheme (*new!*)

1). Given an observation in the test set and its LROI, find all the neighboring members in the training set whose distances are within the LROI. If there is no training data in the LROI, keep increasing the radius of LROI.

2). Compute the center of each class within LROI as the weight average of all training members of the same categories with respect to each new testing member:

$$\vec{r}_{class_k} = \frac{\sum_{\forall j \in class_k} \vec{r}_{ij} * e^{-|\vec{r}_{ij}|^2 / \min(|\vec{r}_{ij}|^2)}}{\sum_{\forall j \in class_k} e^{-|\vec{r}_{ij}|^2 / \min(|\vec{r}_{ij}|^2)}}$$

where $\vec{r}_{ij}$ is the distance vector from the testing sample to the training sample, and $\vec{r}_{class_k}$ is the distance vector from the testing data to the center of the class k.

3) Calculate the probability of each cloud type and classify the data type. In this step, the probabilities of the testing samples for each cloud type are calculated by

$$P_{class_k} = \frac{e^{-|\vec{r}_{class_k}|^2 / \min(|\vec{r}_{class_k}|^2)}}{\sum_{k=1}^{\#\ of\ class} e^{-|\vec{r}_{class_k}|^2 / \min(|\vec{r}_{class_k}|^2)}}$$

In this study the testing sample will be assigned to the category for which it has the maximum probability.

P_{class_k} is the cloud fraction of cloud type k within a FOV!!

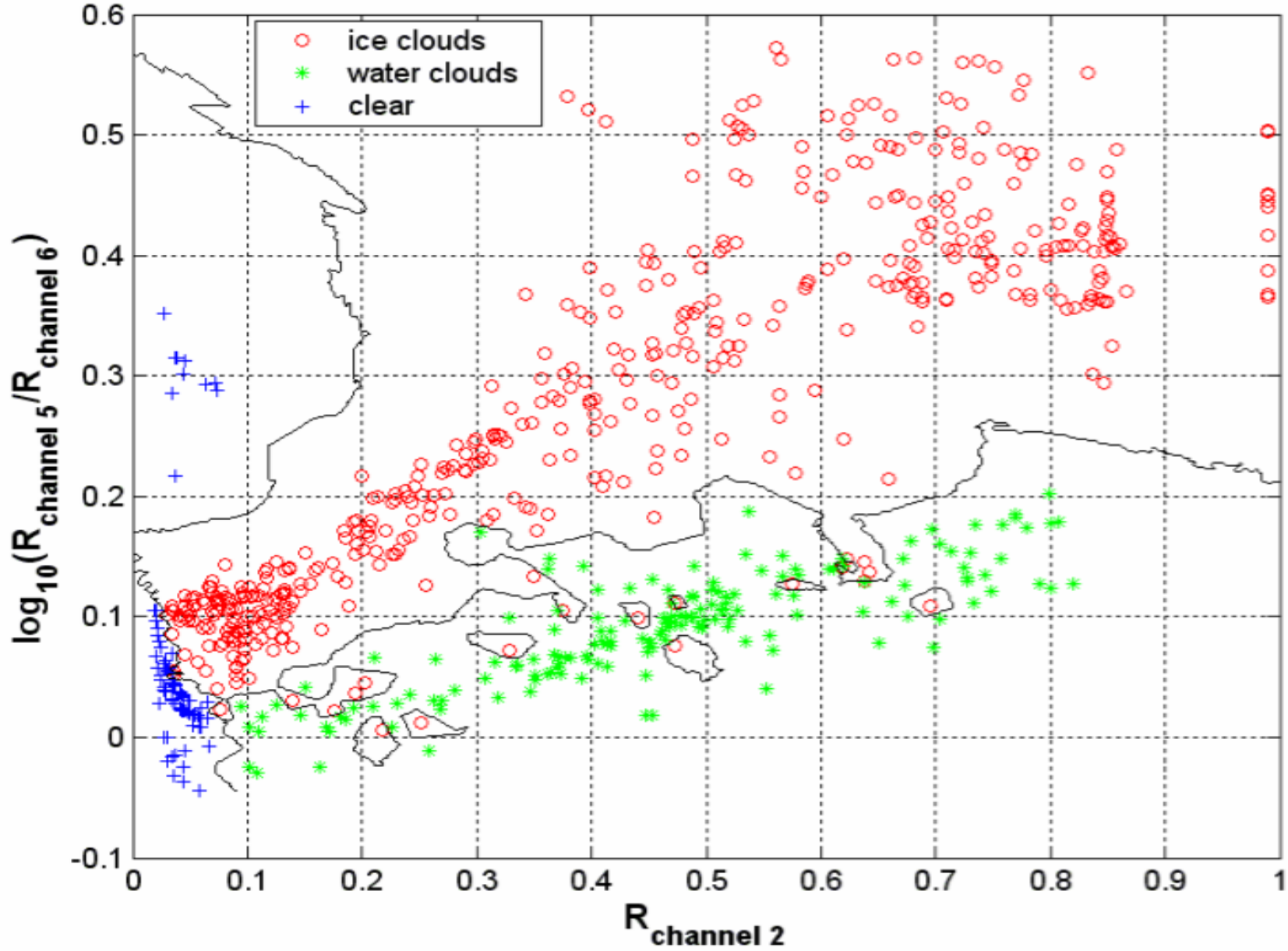


Fig. 3. Classification boundaries on the complete training set based on the LROI scheme.

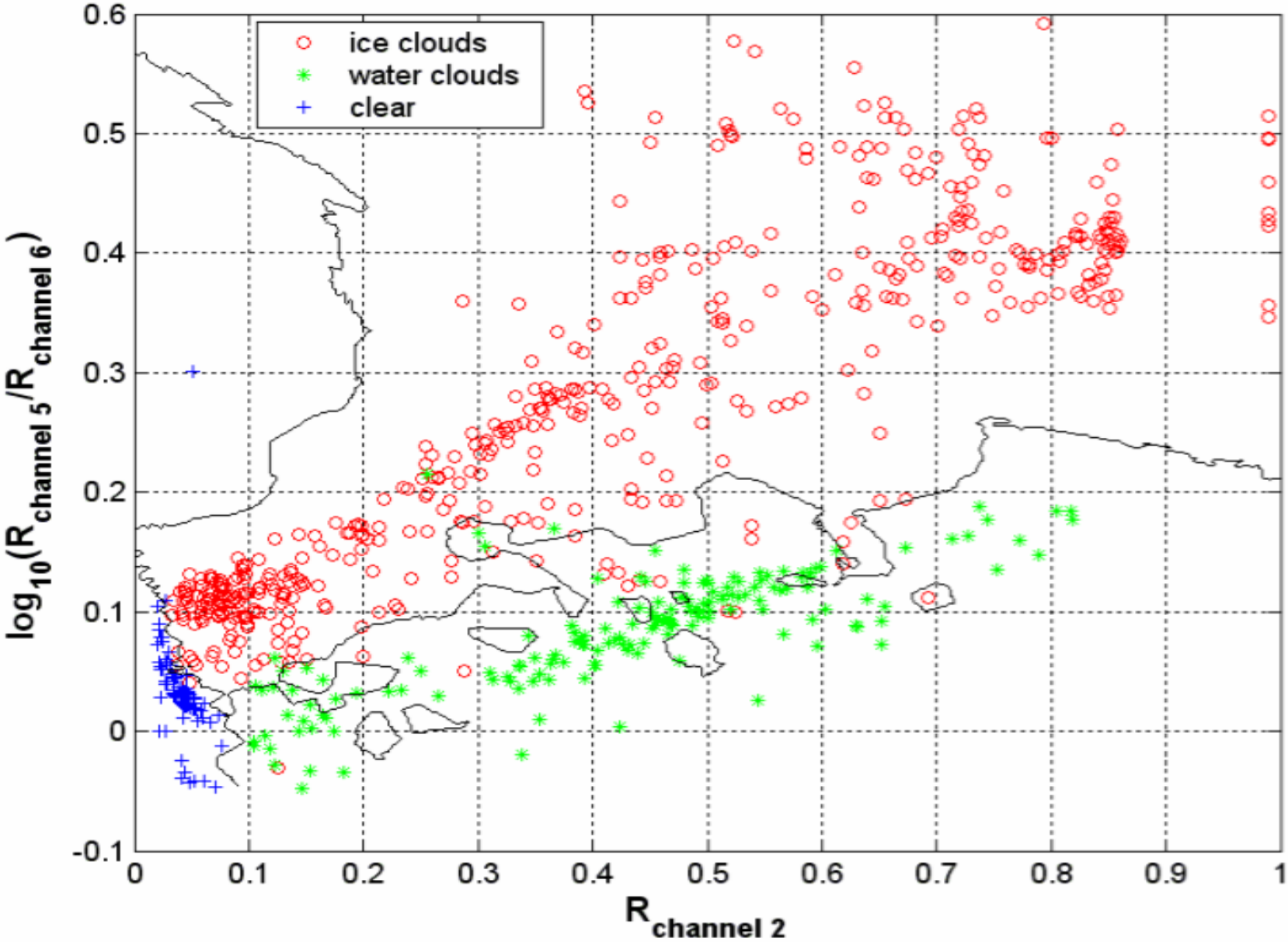


Fig. 4. Classification boundaries on the complete testing set based on the LROI scheme.

Distribution of predicted category vs. true category using the LROI classification. The misclassification error rate is 33/768 = 4.30%

True category	Predicted category			Total
	Clear sky	Ice clouds	Water clouds	
Clear sky	127	1	0	128
Ice clouds	2	456	18	476
Water clouds	0	12	152	164

Comparison of the misclassification error rates:

the LROC scheme = 4.30%.

the MSVM scheme = 4.6875%.

the operational MODIS cloud mask scheme = 18%.

the ML scheme = 62.63%.

Summary

1. The local region of influence (LROI) scheme is developed for the MODIS cloud classification.

2. The LROI scheme outperforms
   the maximum likelihood (ML) scheme,
   the multicategory support vector machine (MSVM), and
   the operational MODIS cloud mask algorithm,
   in terms of the misclassification error rate.

3. Unlike the other schemes, the LROI scheme also provides "cloud fraction" of each class (e.g. ice cloud and water cloud) within each FOV, a desired parameter for cloudy remote sensing!!!

International TOVS Study Conference, 14th, ITSC-14, Beijing, China, 25-31 May 2005. Madison, WI, University of Wisconsin-Madison, Space Science and Engineering Center, Cooperative Institute for Meteorological Satellite Studies, 2005.