Dealing with Cloudy Data

by David Pescovitz

Looking up, it's easy to spot the clouds. The white fluff is strikingly contrasted by the blue sky. It's not so easy from space, especially above the Earth's poles. The clouds blend in against the vast expanses of snow and ice. This is a problem for scientists who use satellites to study clouds and climate. Recently though, UC Berkeley statistician Bin Yu, her graduate student Tao Shi, and their collaborators have devised a new algorithm that detects clouds even when the poles play tricks on the satellites' electronic eyes.

Professor Bin Yu's research is supported by National Science Foundation grants and by a Miller Research Professorship in 2004.

Understanding clouds is essential for scientists to computationally model Earth's current climate and make projections about the future. Historically, uncertainties about how clouds interact with radiation from the sun and the Earth have added to the difficulty of making accurate forecasts.

"Clouds have a cooling effect by reflecting the sun's radiation, but they also warm the Earth's surface by acting like a blanket," Yu says. "The tradeoff between cloud heating and cooling depends mainly on their thickness and height above the surface. So if you don't know where the clouds are and, as a result, the way they feedback on the global climate, the climate models don't really work."

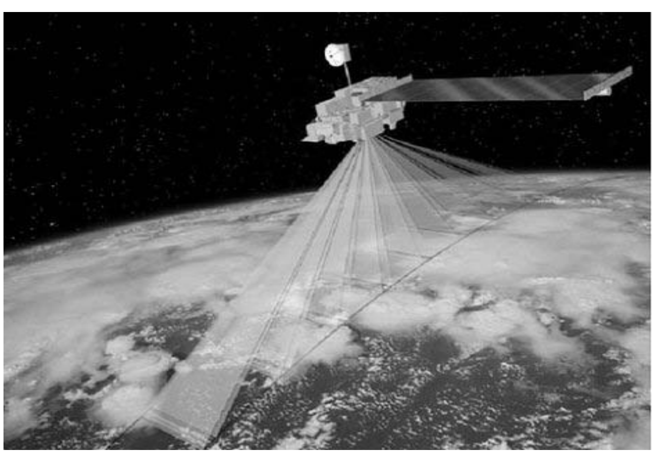


Illustration of both the NASA EOS Terra Satellite and the view directions of the nine MISR cameras. *(Courtesy the researchers)*

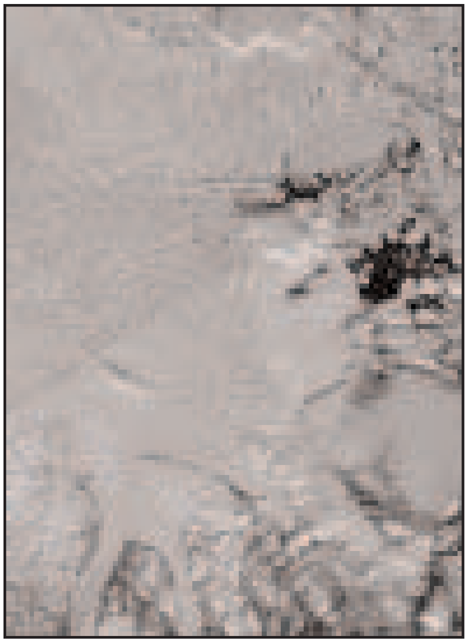
Prof. Bin Yu, whose work with cloudy data is described in this article, has recently been named to the Chang Jiang Chair Professorship at Peking University.

Today's cloud detection systems that sift through mountains of satellite measurements are notoriously error-prone when it comes to the polar regions. The detection schemes usually rely upon visual or thermal contrast between the clouds and the surfaces below them. But conventional approaches fail if such contrasts are lacking. This commonly happens in the polar regions, where the ground may be as bright and as cold as the cloud above it.

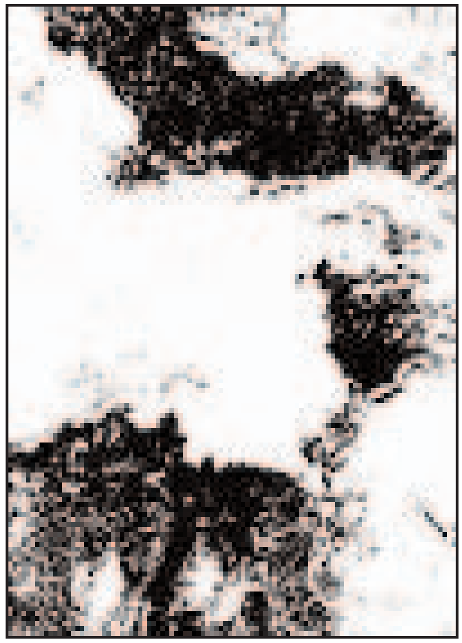
NASA's Earth Observing System (EOS) Terra satellite, launched in 1999, carries a suite of next-generation instruments that aim to revolutionize cloud detection from space. One of these devices is the Multi-Angle Imaging Spectroradiometer (MISR) that views Earth across several bands of the electromagnetic spectrum using cameras pointed at nine different angles. No data like MISR's have ever been collected from space before, opening up possibilities for new cloud detection approaches.

For example, the heights of reflecting surfaces can be derived from the multi-angle "stereo" views, and compared with the known elevation of the terrain below. If a surface is measurably higher than the terrain, it's classified as a cloud. When clouds are close to the ground though, as they often are at high latitudes, other approaches must be used.

Yu and Shi, in collaboration with Pennsylvania State University meteorologist Eugene Clothiaux



Before



After

To the left are two images that illustrate how the new algorithm clarifies cloudy data. The photo labeled "Before" is an MISR image collected over Greenland on June 20, 2001. This is a difficult scence for cloud detection due to the presence of the fork-shaped frozen river and low thin clouds which are invisible in this image. (The photo labeled "After" shows the classification results from the new algorithm. *(Photos courtesy the researchers.)*

and Amy Braverman, a statistician at NASA's Jet Propulsion Laboratory, developed an improved algorithm that leverages MISR's multi-angle capabilities in a novel way to dramatically improve cloud detection in polar regions.

"Instead of looking for clouds, we decided to look for ice," Yu says. "Since ice doesn't move, the data from MISR's multiple angles should be correlated when they're just viewing ice. The angles are filled with valuable information."

"If you don't know where the clouds are and . . . the way they feedback on the global climate, the climate models don't really work."

The researchers' new algorithm focuses on the reflective features of the surface. To distinguish icy terrain from clouds, this algorithm relies on correlations between such angular factors as the smoothness of the reflecting surface and its tendency to scatter light in a certain direction.

Using data from MISR, the researchers found their algorithm could classify polar-region clouds in the satellite data about as well as scientists laboriously identifying clouds in the images by hand, even for some of the most complex scenes.

The new algorithm will be among several to be tested extensively with MISR data in the coming year. When the results are in, MISR will take advantage of the best qualities of each. Until then, Yu and her colleagues are devising web software powered by the algorithm to encourage scientists to process images and put the algorithm through its paces along the way.

Meanwhile, Yu and Shi are considering other applications for MISR data. Already, they've launched a joint effort with Peking University and several environmental scientists in the U.S. to study the dynamics of air pollution in China.

"I'm fascinated by the huge amounts of data collected by new instruments that haven't been available before," Yu says. "The scientific fields that gather these data may be very different, but statistics is the common thread that connects them all together."

(Reprinted from Science Matters @ Berkeley, Volume 2, Issue 8, published by the College of Letters and Science.)