An Objective Regional Cloud Mask Algorithm for GOES Infrared Imager With Regime-Dependent Thresholds

Xiaolei Zou, and Cheng Da
Department of Earth, Ocean and Atmospheric Sciences, Florida State University, USA

Abstract

A local, regime-dependent cloud mask (CM) algorithm is developed for isolating cloud-free pixels from cloudy pixels for Geostationary Operational Environmental Satellite (GOES) imager radiance assimilation using mesoscale forecast models. In this CM algorithm, thresholds for six different CM tests are determined by a one-dimensional optimization approach based on probability distribution functions of the nearby cloudy and clear-sky pixels within a 10°x10° box centered at a target pixel. It is shown that the optimized thresholds over land are in general larger and display more spatial variations than over ocean. The performance of the proposed CM algorithm is compared with Moderate Resolution Imaging Spectroradiometer (MODIS) CM for a one-week period from 19 to 23 May 2008. Based on MODIS CM results, the average Probability of Correct Typing (PCT) reaches 92.94% and 91.50% over land and ocean, respectively.

Data Description and Model Setting

In this study, infrared channels observations from GOES-12 Imager are utilized. Other information used in the LR-CM algorithm includes model simulations of brightness temperature, solar zenith angle, satellite zenith angle, surface types (land, ocean, or coastal), and terrestrial elevation.

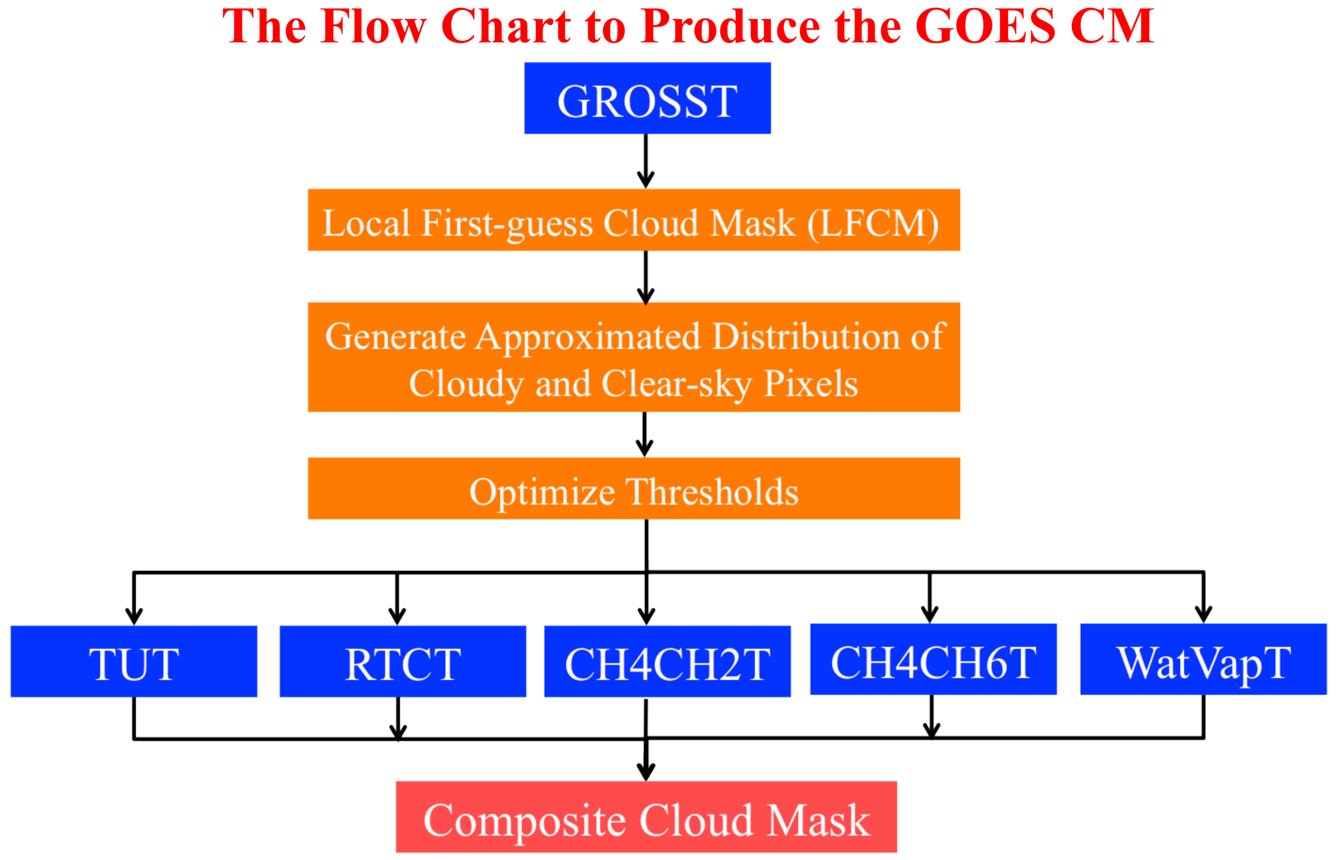
Model simulations of BT for GOES Imager channels 2, 3, 4, 6 required for the CM tests in the LR-CM algorithm are calculated by Community Radiative Transfer Model (Weng, 2007; Han et al., 2007). The Weather Research Forecast (WRF) Advanced Research Weather (ARW) model forecasts at 10-km resolution and 30-minute frequency are used as input to CRTM. The WRF-ARW forecasts are made with initial conditions being generated by the National Centers for Environmental Prediction (NCEP) Gridpoint Statistical Interpolation (GSI) data analysis system in which conventional observations and GOES-12 Imager radiance observations are assimilated (Zou et al., 2011; Qin et al., 2012).

CM Tests in the GOES CM algorithm

Table 1: Six CM tests used by ABI, MODIS, Met Office, and EUMETSAT CM algorithms.

Name	Condition for cloudy pixels	Organization
GROSST	$B^{10.7\mu m} - O^{10.7\mu m} > \varepsilon_{GRST}$	Met Office CM
TUT	$\sigma^{10.7\mu m} - 3\gamma\sigma^{z} > \varepsilon_{TUT}$	ABI CM
RTCT	$\left(O_{max}^{10.7\mu m} - O^{10.7\mu m}\right) - 3\gamma\left(z_{max} - z\right) > \varepsilon_{RTCT}$	ABI CM
CH46T	$\left(O^{10.7\mu m} - O^{13.3\mu m}\right) - \left(B^{10.7\mu m} - B^{13.3\mu m}\right) < \varepsilon_{46T}$	EUMETSAT CM
CH42T	$O^{10.7\mu m} - O^{3.9\mu m} < \varepsilon_{42T}$ (daytime)	MODIS CM
	$O^{10.7\mu m} - O^{3.9\mu m} > \varepsilon_{42T}$ (nighttime)	
WtrVprT	$\left(O^{10.7\mu m} - O^{6.5\mu m}\right) - \left(B^{10.7\mu m} - B^{6.5\mu m}\right) < \varepsilon_{WVT}$	EUMETSAT CM

The Flow Chart to Produce the GOES CM



The First-Guess CM Datasets

In order to better detect cloud in a target region, the first-guess clear-sky and cloudy datasets are generated for each target GOES Imager pixel based on spatial and temporal variations within its 10°×10° area in the LR-CM algorithm. The 10°×10° size is chosen based on the largest horizontal scale of different types of clouds (Wang et al., 2013).
The thresholds are

$$\varepsilon_{GRST} = \begin{cases} 4.5K, & land \\ 2.5K, & ocean \end{cases}$$

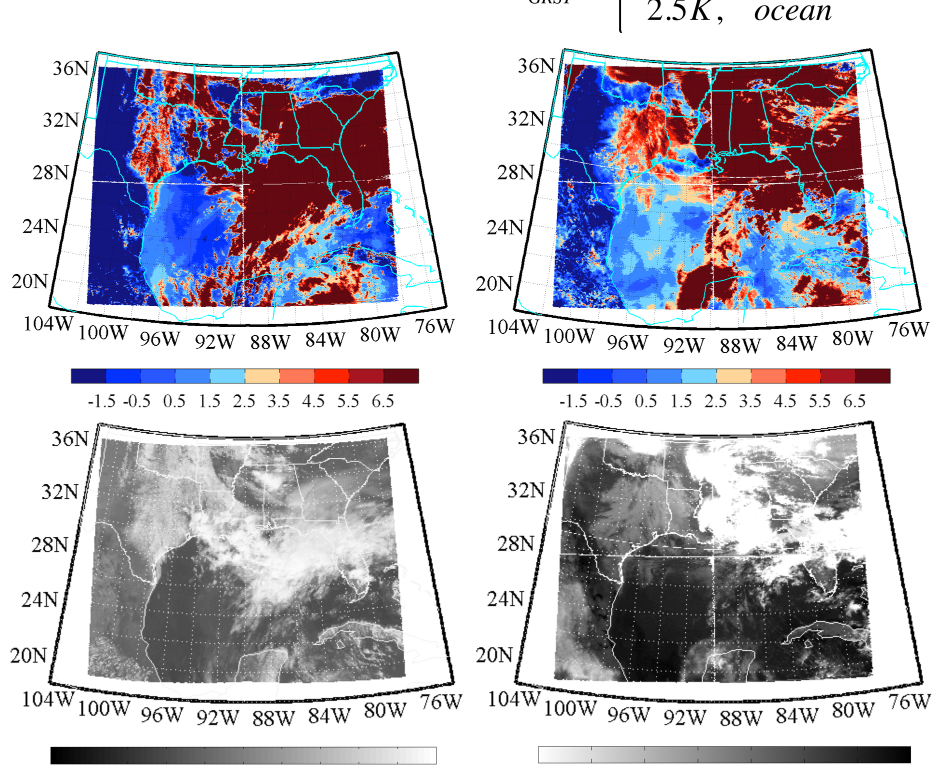


Fig. 1: (a)-(b) Differences of brightness temperature of GOES imager channel 4 (10.7μm) between model simulations and observations (B-O) at (a) a daytime period of 1745-1747 UTC May 22, 2008 and (b) a nighttime period of 0633-0635 UTC May 23, 2008. (c) Reflectance of GOES-12 imager channel 1 (visible channel, 0.65μm) during the daytime period of (a). (d) Brightness temperature distribution of GOES-12 imager channel 1 (infrared channel 2, 3.9μm) during the nighttime period of (b).

Probability of Occurance of Clear-sky and Cloudy Pixels

The probability of occurrence of clear-sky and cloudy pixels detected by the i^{th} CM test based on the first-guess CM datasets are defined respectively as follows

$$P_i^{clr}(\varepsilon_j \le \varepsilon < \varepsilon_j + 0.1) = \frac{N(G^{clr} \cap T_i(\varepsilon_j \le \varepsilon < \varepsilon_j + 0.1))}{N(G^{clr} \cap T_i)}$$

$$P_i^{cld}(\varepsilon_j \le \varepsilon < \varepsilon_j + 0.1) = \frac{N(G^{cld} \cap T_i(\varepsilon_j \le \varepsilon < \varepsilon_j + 0.1))}{N(G^{cld} \cap T_i)}$$

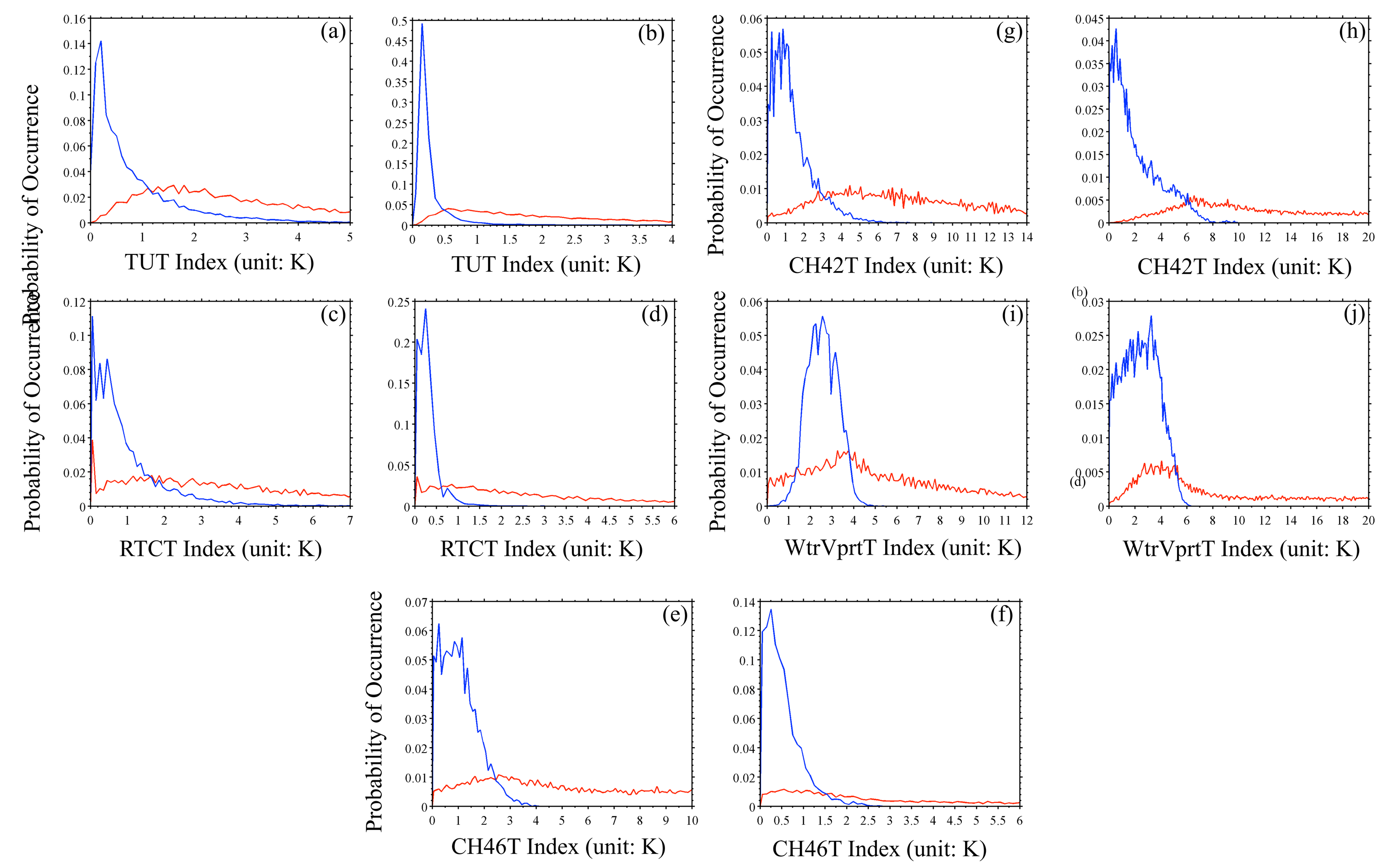


Fig. 2: Probability of occurrences of clear-sky pixels (blue) and cloudy pixels (red) determined by the first guess CM near a pixel over land (left panels) or ocean (right panel) for (a)-(b) TUT, (c)-(d) RTCT, (e)-(f) CH46T, (g)-(h) CH42T, and (i)-(j) WtrVprT indices.

Mathematical Formulation of Threshold Optimization

An objective function is defined for each CM test as follows:

$$f_i^{PCT}(\varepsilon) = \frac{N\left[G^{clr} \cap T_i^{clr}(\varepsilon)\right] + N\left[G^{cld} \cap T_i^{cld}(\varepsilon)\right]}{N\left[T_i\right]}$$

The objective function $f_i^{PCT}(\varepsilon)$ is maximized subject to the following two constraints:

$$f_i^{RFAR}(\varepsilon) = \frac{N\left[G^{clr} \cap T_i^{cld}(\varepsilon)\right]}{N\left[T_i^{cld}(\varepsilon)\right]} \le \alpha_1 \qquad (i = 1,2,...,5)$$

$$f_i^{RLR}(\varepsilon) = \frac{N\left[G^{clr} \cap T_i^{cld}(\varepsilon)\right]}{N\left[G^{clr}\right]} \le \alpha_2 \qquad (i = 1,2,...,5)$$

The optimal threshold ε is determined which maximizes the objective function $f_i^{PCT}(\varepsilon)$ under the above two constraints

$$f_i^{PCT}(\varepsilon) = \max\{f_i^{PCT}(\varepsilon)\}$$

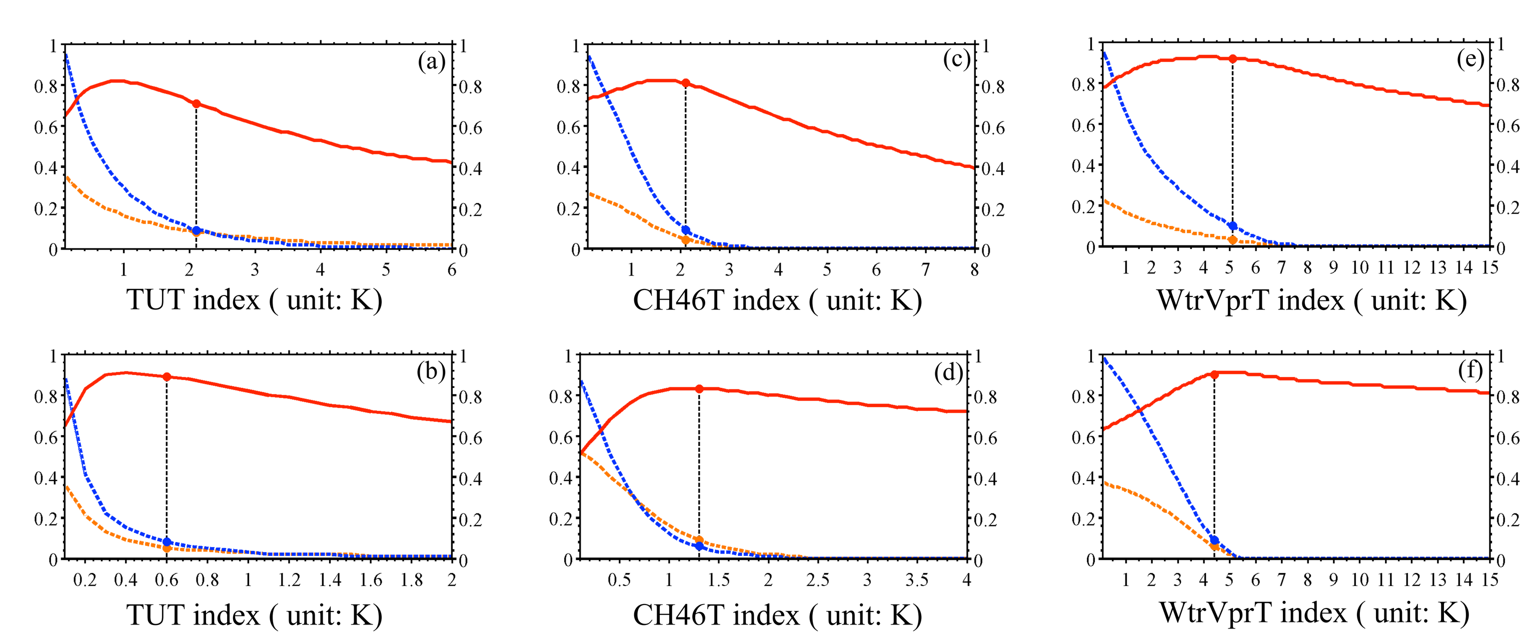
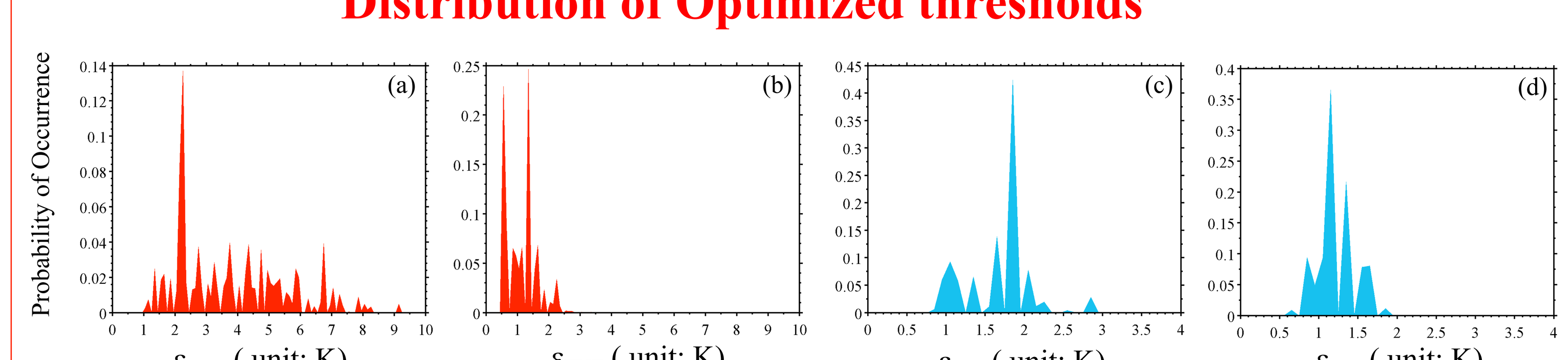


Fig. 3: Variations of $f_i^{PCT}(\varepsilon)$ (orange), $f_i^{RFAR}(\varepsilon)$ (blue), and $f_i^{PCT}(\varepsilon)$ (red) at a target pixel over land (upper panels) and a target pixel over ocean (lower panel) for (a)-(b) TUT, (c)-(d) CH46T, and (e)-(f) WtrVprT indices. The "optimal" thresholds are indicated by vertical dotted line (black).

Distribution of Optimized thresholds



Distribution of Optimized thresholds (Cont')

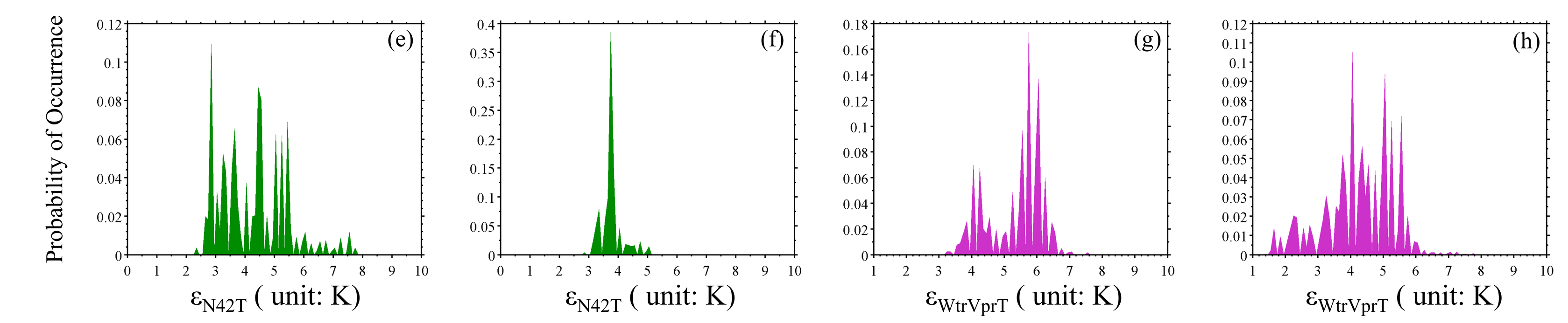
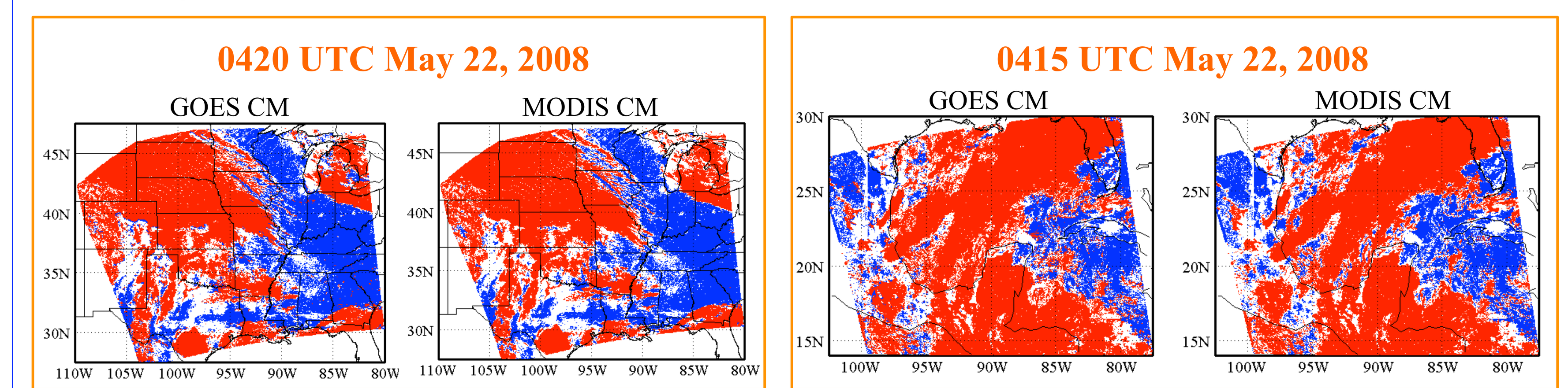


Fig. 4: Distribution of optimal thresholds over land (left panel) and ocean (right panel) for (a-b) TUT, (c-d) CH46T, (e-f) Negative CH42T, and (g-h) WtrVprT for a case study at 0415 UTC May 22, 2008.

Comparison with MODIS Cloud Mask

The GOES CM are compared with MODIS CM during a five-day period from May 19 to May 23, 2008. Totally 5,616,090 GOES-12 Infrared Imager Pixels are collocated.



Average performance over 1 week

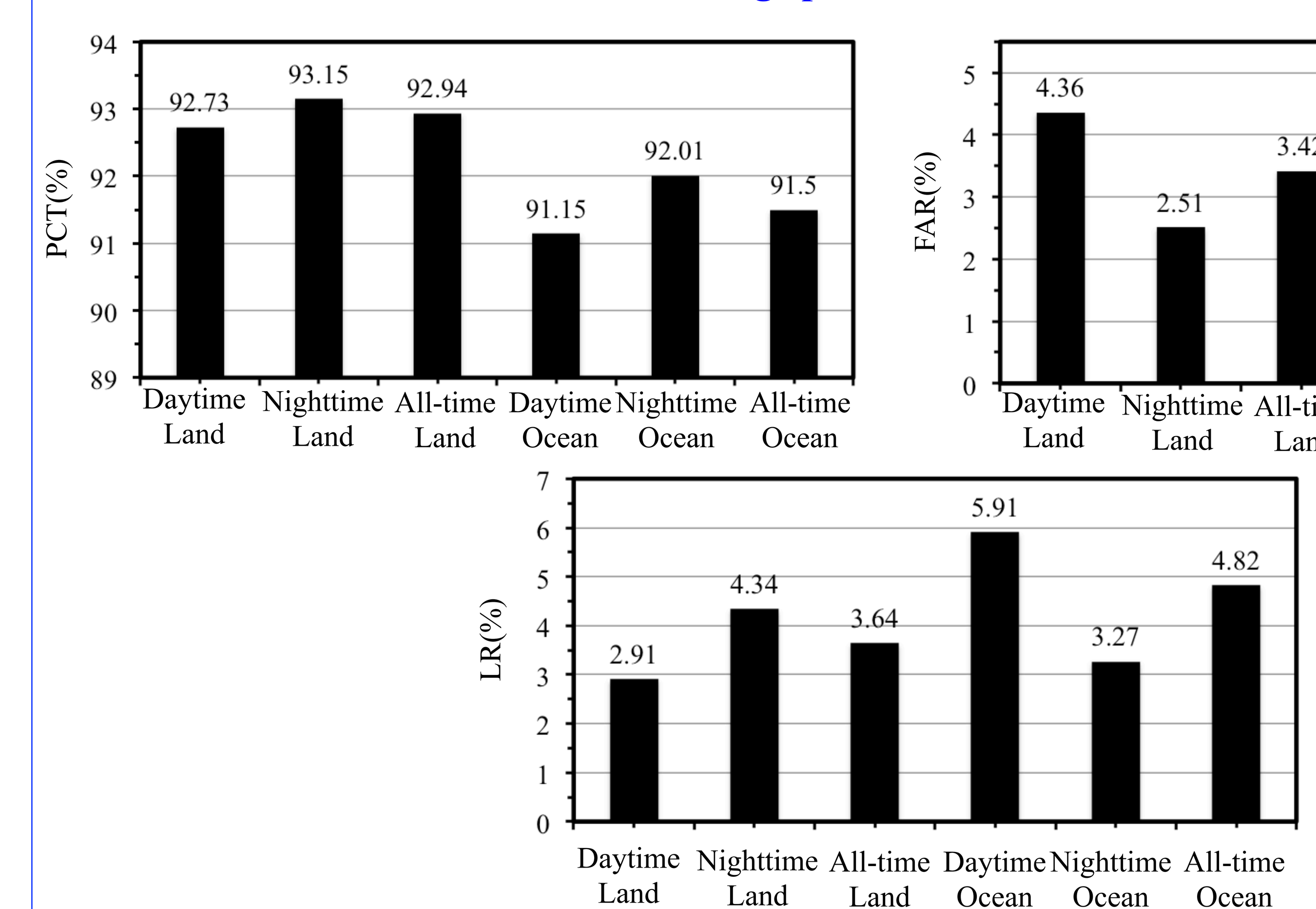


Fig. 5: (a) The PCT, (b) FAR, and (c) LR of the CM determined by the LR-CM algorithm for all the GOES-12 imager pixels collocated with MODIS data during a five-day period from May 19 to May 23, 2008.

Summary and Conclusions

A geostationary satellite provides a time-continuous evolution of weather phenomena over its instrument observing domain. Direct assimilation of geostationary radiance in recent studies proves to improve the forecast skill of numerical models. Removing cloud-contaminated radiance data is a critical step in geostationary radiance assimilation since the current data assimilation system can only ingest clear-sky radiances. The LR-CM algorithm is developed to remove cloud-contaminated observations in a quality control step prior to data assimilation. It employs an optimal procedure for determining a set of dynamic thresholds where local temporal and spatial variations of clouds are introduced. Specifically, the LR-CM algorithm determines implicit dynamic threshold based on the local distribution of clear-sky and cloudy pixels. Firstly, a GROSST is used for generating an approximated distribution of clear-sky and cloudy pixels. The thresholds utilized in each CM tests are then objectively determined by a one-dimensional optimization approach based on the information provided by the first-guess datasets. Two constraints are imposed to ensure the correctness of the LR-CM algorithm under extreme sky conditions (e.g., overcast clouds, and large portion of clear-sky). A pixel is identified as cloudy if it is flagged as cloudy by any CM tests in the algorithm. It is observed that the distribution of optimal thresholds over land possesses more variations than that over ocean. Besides, the average threshold values over land are also greater than the averages over ocean for most CM tests.

A total of 5,616,090 GOES-12 infrared Imager pixels are collocated with MODIS CM during a five-day period from May 19 to May 23, 2008, in order to evaluate the performance of the LR-CM algorithm. A high PCT (above 92%) and a low FAR (below 4%) and a low LR (4%) are achieved during daytime or nighttime over land or ocean. The proposed CM algorithm can easily be implemented for other infrared imager data such as SEVIRI and ABI onboard geostationary satellites.

This is our third study on GOES imager radiance assimilation following the work by *Zou et al.* [2011] and *Qin et al.* [2012], who investigated the assimilation of GOES imager radiance observations at a rather coarse resolution through data thinning. The LR-CM algorithm will be incorporated into the GSI system to remove cloud-contaminated radiance data. It is part of the work toward realizing the full potential of GOES high temporal (3-15 minutes) and spatial resolutions (4-8 km) data.

Reference:

Zou X., and C Da, 2013: An objective Regional Cloud Mask Algorithm for GOES Infrared Imager with Regime-Dependent Thresholds, Journal of Geophysical Research. (Under review)