https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

The Polar Radiant Energy in the Far Infrared Experiment (PREFIRE) principal component-based cloud mask: A simulation experiment

Brian H. Kahn[1], Cameron Bertossa[2], Xiuhong Chen[3], Brian J. Drouin[1], Erin Wagner Hokanson[4], Xianglei Huang[3], Tristan S. L'Ecuyer[2,5], Kyle S. Mattingly[4], Aronne Merrelli[3], Tim Michaels[4], Nathaniel B. Miller[4,5], Federico Donat[6], Tiziano Maestri[6], and Michele Martinazzo[6]

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, 91109, USA
[2]Department of Atmospheric and Oceanic Sciences, University of Wisconsin–Madison, Madison, Wisconsin, USA
[3]Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, Michigan, USA
[4]Space Science and Engineering Center, University of Wisconsin–Madison, Madison, Wisconsin, USA
[5]Cooperative Institute for Meteorological Satellite Studies, Madison, Wisconsin, USA
[6]Physics and Astronomy Department, Alma Mater Studiorum – University of Bologna, Italy

Correspondence to: Brian H. Kahn (brian.h.kahn@jpl.nasa.gov)

Abstract. We describe a cloud mask simulation experiment developed for the PREFIRE mission. The basis of the cloud mask is a principal component (PC) methodology (PC-MSK) adapted from the algorithm heritage of the upcoming Far-infrared Outgoing Radiation Understanding and Monitoring (FORUM) mission. Simulated clear-sky and cloudy-sky PREFIRE radiances are calculated from the Goddard Earth Observing System (GEOS) meteorological fields and include a variety of complex cloud configurations. The simulation experiment is based on local training that is adjusted along segments of simulated orbits that mimic actual PREFIRE orbits. A numerically stable method of separating clear sky from cloudy sky is achieved using Otsu's binary classification method and requires no a priori thresholding estimate for multimodal histograms. Comparisons are made against a machine-learning cloud mask (ML-MSK) developed for the PREFIRE mission. The global hit rate of PC-MSK (92.6%) compares favorably to the hit rate of ML-MSK (95.3%). The Arctic hit rate of PC-MSK (86.7%) compares favorably to the hit rate of ML-MSK (89.4%) and both cloud masks are shown to meet mission requirements for PREFIRE cloud detection. The simulation experiment demonstrates the potential for accurate cloud masking with PREFIRE despite a low number of information-containing PCs compared to those obtained from hyperspectral infrared sounders. We conclude with a discussion about clear-sky and cloudy-sky training sets that are suitable for an operational version of PC-MSK and their development during the post-launch checkout time period.

1 Introduction

The Polar Radiant Energy in the Far InfraRed Experiment (PREFIRE; L'Ecuyer et al. 2021) is a satellite mission designed to reduce uncertainties in the Arctic energy balance. The focus is on the far-infrared (FIR) portion of the electromagnetic

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

spectrum — typically described as 100 to 600 cm^{-1} in the atmospheric sciences — that remains relatively unexplored (Harries et al. 2008). This spectral region is particularly significant because it contains about half of Earth's infrared emission and is strongly influenced by greenhouse gases such as water vapor (Turner and Mlawer 2010), cloud properties including amount, altitude, thermodynamic phase, microphysics (Saito et al. 2020; Peterson et al. 2022), and surface emissivity that globally impacts Earth's radiation budget and atmospheric circulation (Chen et al., 2014; Feldman et al. 2014). The baseline mission is comprised of two 6U CubeSat spacecraft with miniaturized spectrometers in two differently phased near-polar inclinations at altitudes between 470–650 km. The two CubeSats are capable of diurnal subsampling operation over one seasonal cycle (Kahn et al. 2020; Drouin et al. 2022) with the launches targeted for May 2024.

PREFIRE is expected to be a pathfinder mission for two future satellite observatories that will make observations at FIR frequencies. The first is the Far-infrared Outgoing Radiation Understanding and Monitoring (FORUM) (Palchetti et al. 2020; Sgheri et al. 2022) to fly as ESA's ninth Earth Explorer mission planned for launch in 2027. FORUM will fly in loose-formation with the Infrared Atmospheric Sounder Interferometer – New Generation (IASI-NG) covering the mid-infrared (MIR) and FIR portions of Earth's emission spectrum. The second is the Thin Ice Cloud in Far InfraRed Emissions (TICFIRE) (Libois et al. 2016) to fly as a key part of CSA's High-Altitude Aerosols, Water Vapor, and Clouds (HAWC) mission in 2031, and to be coordinated with NASA's Atmosphere Observing System (AOS) constellation. While PREFIRE, FORUM and TICFIRE will make observations in the FIR, the spectral resolution, orbital swath width, spatial resolution, and instrument characteristics are significantly different and are intended to enable different avenues of scientific exploration.

A first step in the algorithm flow requires the detection of clouds within a given footprint. The detection of cloud is rather challenging in the infrared spectrum because of the complexity added by thermal contrasts between surface and cloud, thermal and moisture inversions, unknown surface characteristics, and subpixel heterogeneity that is spatially and temporally variable. Many legacy cloud detection algorithms in the infrared use fixed thresholding or look-up tables that have known challenges in the polar regions (Liu et al., 2022). Recently, machine learning (ML) algorithms have emerged as effective tools for cloud detection and characterization using infrared radiances. Traditional threshold-based cloud detection algorithms have proven to be highly skillful for many types of clouds, but have limitations regarding universality as the thresholding is sensitive to spatial and temporal scene variability (Mahajan and Fataniya, 2020). The performance of cloud detection utilizing PREFIRE observations will ultimately determine which footprints are clear and cloudy for downstream algorithm processing.

Towards this end, we describe a principal component analysis (PCA)-based approach first developed by Maestri et al. (2019) (henceforth, M19) that was designed for hyperspectral FORUM radiances. In Section 2, the simulated data and modeling tools used are described. In Section 3, the methods of M19 are summarized and the assumptions used in the PREFIRE implementation are detailed. In Section 4, representative orbital segments are shown and global statistics are presented. In Section 5, expected challenges with an operational cloud mask are discussed, and in Section 6 the results and implications of this investigation are summarized.

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

2 Data and Modeling

2.1 Simulated Orbital Data

Simulated PREFIRE orbital swaths are created assuming an altitude of 540 km, a Mean Local Time for Ascending Node (MLTAN) of 0815 UTC, and a sensor zenith angle of zero degrees. All orbital data examined herein are limited to 01 January 2021. Auxiliary meteorological data ("Aux-Met") from the Goddard Earth Observing System Forward Processing for Instrument Teams archive (GEOS FP-IT; Lucchesi et al. 2015) using v5.12.4 of the GEOS model (Molod et al. 2015) are linearly interpolated to simulated Thermal IR Spectrometer (TIRS) footprint center points and observation times. Four separate GEOS FP-IT data collections comprise the content of the PREFIRE Aux-Met files: three-hourly instantaneous two-dimensional (2D) surface and column integrated quantities (inst3_2d_asm_Nx) and three-dimensional (3D) assimilated state on 72 model levels (inst3_3d_asm_Nv), along with one-hourly time-averaged 2D land surface diagnostics (tavg1_2d_lnd_Nx) and ocean surface diagnostics (tavg1_2d_ocn_Nx). To describe the meteorological context of simulated scenes contained within the orbital swaths, vertical profiles of temperature (T), specific humidity (q), cloud liquid water content (q_l), and ice water content (q_i) from inst3_3d_asm_Nv, and skin temperature (T_{skin}) from inst3_2d_asm_Nx are examined.

A sub-sampled set of simulation parameters representative of an orbital segment near the northeast coast of Greenland is depicted in Fig. 1. The orbital segment contains 2000 simulated TIRS footprints (250 along-track line segments and 8 cross-track footprints). The cloud top pressure (CTP) is defined as the uppermost pressure level with a non-zero value of either q_i or q_l. The natural logarithm of cloud optical depth (COD) is calculated from the vertical integral of q_i and q_l and cloud temperature profiles from the orbital simulation data, along with assumed effective radii following Bertossa et al. (2023; henceforth B23). A rudimentary classification of cloud phase follows the criteria described in Table 1. The orbital segment contains a complex mixture of clear sky, high and low clouds with ice and liquid phases, and ice, land and ocean surfaces. This region is one of the most challenging that PREFIRE will encounter. Liu et al. (2022) showed that existing satellite records of cloud fraction have the largest discrepancies over and in proximity to Greenland. Optically-thick, liquid-phase clouds with CTP between 500-900 hPa dominate the northern Greenland Sea. A mix of liquid and ice is found in a few footprints nearer to the Greenland coastline, while clear sky dominates along the coast and extends into the Greenland ice sheet (GrIS). An area of thin cirrus with a COD between 0.25–0.5 is found in the southern part of the orbital segment over the GrIS between 75°N and 77°N.

The T_{skin} and integrated precipitable water vapor (PWV) are shown in Fig. 2. Except for a small part of the ocean where T_{skin} is near freezing, most of the orbital segment is subfreezing. Values of PWV range between 5–10 mm over the warmest T_{skin}, with additional elevated values of PWV between 4–6 mm near the band of thin cirrus over the ice sheet. Elsewhere, the orbital segment is characterized with very low values of PWV less than 3–4 mm, values that are consistent with semitransparent atmospheric windows in the FIR (Turner and Mlawer, 2010).

Selected T and q profiles for clear and cloudy footprints are shown in Fig. 3. This orbital segment is characterized by a mixture of surface-based and elevated T inversions with tropopause pressures that vary between 200–300 hPa. Additionally, profiles of q are highly variable with occurrences of low-level q inversions, and mid-to upper-tropospheric values of q that vary by an

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

order of magnitude within a fixed pressure level. Despite the significant spatial variability, the clear and cloudy T and q profiles have significant overlap in their vertical distributions suggesting that the thermodynamic profiles are not particularly distinct from each other. The meteorological conditions and complex topography contained within the scene illustrated in Figs. 1–3 suggest an overall challenging region for cloud detection.

2.2 PCRTM

The Principal Component-based Radiative Transfer Model (PCRTM) v3.4 (Liu et al. 2006) provides radiance calculations used in the PC-MSK development. Every TIRS observation has both clear-sky and cloudy-sky simulated radiances; clear-sky is simulated by skipping the available cloud properties. Principal component (PC) scores are determined from a pre-computed look-up table of PCs that depend on atmospheric state. Chen et al. (2013) demonstrated that PCRTM v2.1 is sufficiently accurate with a root-mean square error (RMSE) of 0.67K while being capable of computational speeds of 1000s of times faster than benchmark calculations from the Line-By-Line Radiative Transfer Model (LBLRTM; Clough et al. 2005). PCRTM v3.4 is an improved version of PCRTM v2.1. In addition, these simulations have incorporated realistic surface spectral emissivity for all 18 International Geosphere Biosphere Programme (IGBP) surface types based on the global surface emissivity dataset developed by Huang et al. (2016). Chen et al. (2013) developed a simulator that generates user-selected instrument radiances using PCRTM with numerical weather prediction (NWP) or reanalysis model profiles of temperature, humidity, clouds, and surface properties. PCRTM output at 0.5 cm^{-1} sampling that covers the thermal infrared (50–2760 cm^{-1}) is convolved with the TIRS instrument spectral response functions (SRFs) to produce simulated PREFIRE radiances using GEOS NWP data.

2.3 TIRS Instrument and Channel Characteristics

Each PREFIRE cubesat carries a Thermal Infrared Spectrometer (TIRS) with 54 channels ranging between 4 to 53 μm with a spectral sampling of 0.84 μm, and an additional broadband channel that extends from 0 to 53 μm. Estimates of the noise equivalent delta radiance (NEdR) before thermal vacuum testing are illustrated in Fig. 1 of Miller et al. (2023) and generally range between 0.01–0.04 W m^{-2} sr^{-1} μm^{-1}. The signal-to-noise ratio (SNR) peaks between 50–250 in the 8.5 to 29 μm spectral region with substantially reduced values at longer and shorter wavelengths. This initial algorithm implementation does not take into consideration the dependence of NEdR in the cross-track direction, nor the differences between the two TIRS instruments. The 23 selected channels are limited to those available between 8.5 to 29 μm to (1) minimize impacts from low SNR at shorter wavelengths, (2) remove effects of solar reflected radiation at shorter wavelengths, (3) remove channels with strong water vapor absorption at longer wavelengths (e.g., Xie et al. 2022), and (4) maximize spectral sensitivity to cloud thermodynamic phase (Peterson et al. 2022), cloud temperature and extinction efficiency dependence (Saito et al. 2020) near 20 μm.

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

Figure 4 shows TIRS radiance spectra used for clear-sky and cloudy-sky training sets for the locations depicted in Fig. 1. The clear and cloudy spectra have significant overlap in magnitude, implying a more challenging cloud detection target than elsewhere in which clear and cloudy radiances can greatly differ (e.g., the Tropics). For cloudy footprints, a majority of the spectra were randomly drawn from liquid clouds as they dominate the cloud statistics in the scene. Figure 1 also confirms that the liquid clouds are located at lower altitudes (higher CTPs) than ice clouds. Despite the coarse spectral resolution of TIRS, the spectra appear to contain a muted spectral slope in the 8.5 to 11 μm window due to index of refraction differences between ice and liquid (Peterson et al. 2022). When observations are made post-launch, clouds are expected to contain additional variability in the atmospheric window from variations in particle size, optical depth, thermodynamic phase, and cloud overlap variability that are not present in this pre-launch cloudy simulation data set. Therefore, the TIRS channels between 8.5 to 29 μm are likely to have additional capability for cloud characterization that cannot as of yet be quantified because of limitations in the cloudy simulation data set.

Gaussian noise that is consistent with the NEdR of each TIRS channel was randomly added to clear and cloudy radiances. The noise is independently calculated for clear and cloudy spectra within the same footprint. Tests with correlated noise between clear and cloudy radiances for the same footprints increases the rate of cloud detection failure (not shown).

2.4 Machine learning-based cloud mask (ML-MSK)

B23 developed a neural network (NN)-based cloud detection algorithm for PREFIRE and showed that it is capable of detection rates in excess of 90% in the Arctic region. The algorithm uses the Tensorflow Keras NN framework (https://keras.io/) and employs 3-km FV3 global non-hydrostatic model simulations produced by Geophysical Fluid Dynamics Laboratory (GFDL) as the basis for training and testing of simulated PREFIRE radiances. The training data were obtained in discrete latitude bands. The fact that B23 was trained with GFDL simulations, while PC-MSK was trained with GEOS simulations, ensures each algorithm's independence. Having two independent cloud masks provides robustness for clear-sky detection which is especially important for surface emissivity retrievals (Xie et al. 2022). The cloud mask of B23 is labeled as ML-MSK.

3 Methodology

3.1 Training matrices

The PC-MSK follows the mathematical framework of M19 but there are key differences in the construction of the training sets. The clear-sky training matrix ($\mathbf{TR}_{\mathbf{clr}}$) and cloudy-sky training matrix ($\mathbf{TR}_{\mathbf{cld}}$) (1) have dimensions of 20 samples and 23 channels (20×23) between 8.5 to 29 μm (e.g., Fig. 4), (2) are locally defined for PREFIRE orbital segments of 2000 TIRS spectra that typically extend across 10–15° latitude, and (3) are updated with each successive orbital segment. A local training approach differs from the global approach typically used in the literature that divides training data into a few, broad zonal latitude bands (Huang and Antonelli, 2001). For instance, Hultberg et al. (2017) explored a hybrid approach of local and global

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

training data sets for wildfire detection, but the importance of including global data in the training is emphasized. In the development of PC-MSK, zonally-symmetric bands were used for training in a similar fashion to B23 and M19 (tropics, midlatitudes, poles). While some success was achieved with hit rates over 90% in the tropics using a zonally-symmetric training set, the PC-MSK led to worse performance in mid- and high latitudes. The local training approach taken in this simulation study is consistent with the reduced information content of TIRS compared to hyperspectral data such as IASI (Hultberg et al., 2017) and FORUM (M19).

The 20 randomly-selected TIRS footprints are drawn from quartiles of 11 μm radiances, the boundaries of which are dynamically set by the minimum and maximum radiances found within each orbital segment. Five spectra are drawn from each quartile of clear and cloudy radiances. This ensures that (1) samples are more evenly distributed in radiance space that more faithfully capture the true spectral variance in the scene, and (2) samples are not disproportionately weighted towards warm or cold areas that may dominate the scene. The clear sky samples are drawn from the clear-sky simulated spectra with no dependence on whether clouds were present in the interpolated GEOS data. Cloudy samples were drawn from the cloudy-sky simulated spectra from only those observations that contained clouds. If the clear spectra are only drawn from outside of clouds, the PC-MSK performance notably degrades in some orbital segments. This degradation may be related to cloud state dependence of temperature and moisture in particular cloud states (e.g., Yue et al., 2013) as simulated clear-sky spectral variability may vary within and outside of clouds. The cloud state dependence is generally small in the tropics but is frequently substantial in the extratropics (Fetzer et al., 2006).

3.2 Calculation of the similarity index (SI)

A pair of 21×23 extended training matrices are formed for each of the 2000 TIRS footprints in the orbital segment: the first for clear sky ($\mathbf{ETR_{clr}}$) and the second for cloudy sky ($\mathbf{ETR_{cld}}$), each containing the 20 training spectra plus each spectrum within the orbital segment of 2000 TIRS footprints. Then, PCs (eigenvalues) and eigenvectors are separately calculated once per orbital segment for $\mathbf{TR_{clr}}$ and $\mathbf{TR_{cld}}$, and then for $\mathbf{ETR_{clr}}$ and $\mathbf{ETR_{cld}}$ for each of the 2000 TIRS footprints. Generally-speaking, if a given TIRS spectrum is clear, the eigenvectors of $\mathbf{TR_{clr}}$ and $\mathbf{ETR_{clr}}$ should be similar. Likewise, if a given TIRS spectrum is cloudy, the eigenvectors of $\mathbf{TR_{cld}}$ and $\mathbf{ETR_{cld}}$ should be similar. The approach of M19 operates within this general framework.

The eigenvectors are contained in four eigenmatrices ($\mathbf{TREM_{clr}}$, $\mathbf{TREM_{cld}}$, $\mathbf{ETREM_{clr}}$ and $\mathbf{ETREM_{cld}}$) that have a specific number of PCs that are attributable to a limited number of ranked geophysical spectral signals P_0 from the set of eigenvectors. The eigenvectors are shown in Fig. 5 for the clear-sky and cloudy-sky training data sets up to p=[1,…,5]; where p is a counting variable for the eigenvectors/PCs. The eigenmatrices that contain eigenvectors up to a value of P_0 are used to calculate a similarity index (SI) for clear sky,

$$\mathrm{SI_{clr}(j)} = 1 - \frac{1}{2\,P_0} \sum_{p=1}^{P_0} \sum_{\nu=1}^{23} \left| \mathbf{ETREM}_{\mathrm{clr,j}}(\nu, \mathrm{p})^2 - \mathbf{TREM}_{\mathrm{clr,j}}(\nu, \mathrm{p})^2 \right|, \quad (1)$$

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

where the second summation is performed on the total number of channels accounted for. Similarly, for cloudy-sky,

$$\mathrm{SI_{cld}(j)} = 1 - \frac{1}{2\,P_0} \sum_{p=1}^{P_0} \sum_{\nu=1}^{23} \left| \mathbf{ETREM}_{\mathrm{cld,j}}(\nu, \mathrm{p})^2 - \mathbf{TREM}_{\mathrm{cld,j}}(\nu, \mathrm{p})^2 \right| . \quad (2)$$

Following Turner et al. (2006) and M19, the value of P_0 is defined by minimization of an indicator function (IND),

$$\mathrm{IND(p)} = \frac{\mathrm{RE(p)}}{(P-p)^2}, \quad (3)$$

and the real error (RE),

$$\mathrm{RE(p)} = \sqrt{\frac{\sum_{i=p+1}^{P} \lambda_i}{T_i (P-p)}}, \quad (4)$$

where λ is the *i*th eigenvalue, T_i=20 is the number of spectra in the training sets, P=23 is the number of spectral channels, and P_0 is obtained when IND is minimized in Eqn. (3). The value of P_0 is independently determined for clear and cloudy skies. Calculations of P_0 vary between three to seven, depending on the orbital segment, with the highest values in more complex scenes. Clear-sky and cloudy-sky values of P_0 can differ by one or two about half of the time depending on the complexity of the orbital segment. We note that P_0 is much smaller than that typically found with the hyperspectral Atmospheric Emitted Radiance Interferometers (AERI) in Turner et al. (2006), where P_0=100 is typical of the ARM Department of Energy's (DOE) Atmospheric Radiation Measurement (ARM) program's North Slope of Alaska (NSA) observatory, and P_0=350 is typical of the ARM Tropical Western Pacific (TWP) observatories.

Fixing P_0 to the same value for clear-sky and cloudy-sky training sets as done in M19 was tested. That led to a degradation in cloud detection performance relative to the use of independent values of P_0. However, this cannot be considered a general rule because there are 100s of training spectra and ~3000 channels in FORUM. The PC-MSK is sensitive to P_0 using TIRS radiances since there are only a few PCs that contain geophysical signal. Even an offset of 1 for P_0 from the most optimal value calculated from Eqns. (3) and (4) may degrade PC-MSK cloud detection accuracy. Examination of spatial maps of PCs (not shown) confirms that PCs > P_0 contain patterns of speckled noise, while PCs ≤ P_0 contain patterns of robust spatial variability. Figure 5 shows the five leading eigenvectors for clear and cloudy radiances where P_0=4 according to Eqns. (3) and (4). Values of P_0 are recalculated for each orbital segment of 2000 TIRS spectra because (i) the magnitude of geophysical variability is strongly dependent on the location along the PREFIRE ground track, and (ii) the inherent sensitivity of PC-MSK performance to P_0.

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

Once Eqns. (1) and (2) are calculated for every TIRS footprint within the orbital segment, the SI difference (SID) is calculated using the following relation defined in M19,

$$\mathrm{SID(j)} = \mathrm{SI_{cld}(j)} - \mathrm{SI_{clr}(j)}\ , \tag{5}$$

where j=[1,…,2000]. A histogram of SID is shown in Fig. 6 for the orbital segment depicted in Figs. 1 and 2. For optimal performance, the histogram of SID should separate into two distinct modes, one related to clear (on the left), and the other to cloud (on the right). M19 discusses two methods to threshold the SID histogram into clear and cloud. The first method termed the *elementary approach* fixes the difference in clear and cloud at SID=0. Clearly, this is not desirable as SID is not guaranteed to be centered at zero and strongly deviates from it along the orbital track (not shown). The second method termed the *distributional approach* allows for a more accurate classifier. Cossich et al. (2021) used a training set of lidar-identified cloud-top thermodynamic phase to determine the value of SID that minimizes the overlap of phase misclassifications by allowing for a dynamic shift of the threshold in SID from zero (the *distributional approach*). Algorithm supervision and a sufficient diversity of training data sets that are valid for all conditions encountered along-orbit are required for a global implementation. This method was deemed impractical for PC-MSK as the training matrices are locally adjusted along the orbit, requiring a different approach.

3.3 Automated thresholding for clear and cloud

An unsupervised, nonparametric classification developed by Otsu (1979) is used for the thresholding of two modes in a bimodal distribution. This method is computationally efficient and is able to separate SID histograms into clear and cloudy regions even if the modes are overlapping. For clarity, Fig. 6 is sorted into bin widths of SID=0.05, whereas, Otsu's method is applied to raw unbinned values of SID with several significant digits. Otsu's method was compared against other commonly used techniques, namely, thresholding using an iterative minimum error method (Kittler and Illingworth, 1986), an iterative selection methodology (Ridler and Calvard, 1978), a moment-preserving method (Tsai, 1978), and a method that leverages the entropy contained in histograms (Kapur et al., 1985). Glasbey (1993) as well as Sezgin and Sankur (2004) demonstrated that the minimum error method of Kittler and Illingworth (1986) is most accurate for a diverse variety of sample images. However, as the minimum error method is iterative, an ill-suited initial value may lead to a non-convergent solution (Xue and Zhang, 2012) and is therefore not preferrable for a fully automated cloud mask algorithm. The testing of the Kittler and Illingworth (1986) method confirmed that the PC-MSK algorithm was non-convergent for a few orbital segments. While Otsu's method was successful for all orbital segments tested, the method is occasionally challenged when (1) the two modes have very unequal counts in the distributions, such as an orbital segment that is virtually clear with only a few cloudy footprints, or (2) when two modes exhibit large differences in variances (Xue and Zhang, 2012). To conclude, Sezgin and Sankur (2004) showed that Otsu

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

(1979) ranks well against most thresholding methods and is suitable for use in an operational implementation of PC-MSK because of its numerical stability.

3.4 Pre-launch algorithm

Table 2 summarizes the key differences between M19 and the PC-MSK implementation for PREFIRE. Training matrices use "random draws" from each orbital segment of 2000 TIRS footprints, but with a constraint regarding the random draws taken within quartiles of 11 μm radiance bins. Random draws for clear skies are performed from the entire set of 2000 footprints, while random draws for cloudy skies are performed only for footprints with COD≥0.25; the value of COD is identical to the ML-MSK training threshold (B23). As discussed in Section 3.1, clear-sky samples should not be limited to actual clear sky, but should include the effects of clear-sky signatures in semi-transparent clouds which can impact semi-transparent cloudy spectral signatures.

The local random sampling is robust and performs well in a majority of geographical regions. However, there are particular regions where the cloud detection performance is acutely sensitive to the particular random draw of the training set. Thus, we employed an added layer of brute-force stochastic sampling for all of the simulated orbital data. The clear and cloudy spectra are sampled independently five times, and are combined together for a total of 25 combinations of training sets. PC-MSK is then executed 25 times using each of the training set combinations. The performance for each of the 25 PC-MSK simulations is then compared against the COD-MSK "truth" (wherever COD>0.25) to determine which of the 25 yields the best agreement. That combination is then used for the final PC-MSK estimate. For an operational implementation, our intent is to use either (1) a more sophisticated and nuanced training set that is optimized for accuracy and furthermore does not require a brute-force approach because of the obvious problem of choosing the best of the 25 combinations, or (2) develop a static training set that is fixed for particular latitude and longitude boxes, separated by land, ocean, and ice surfaces, with variations over the seasonal cycle. The primary purpose of this pre-launch algorithm is to describe a flexible tool capable of using any type of training set. In Section 5, a more detailed set of options will be described that outline our strategy for optimizing the formation of clear-sky and cloudy-sky training data sets.

4 Results

4.1 Selected orbital segments

The three cloud masks (COD-MSK, ML-MSK, and PC-MSK) and SID calculated for PC-MSK are shown in Fig. 7 for the orbital segment depicted in Fig. 1. PC-MSK and ML-MSK show similar skill in detecting (i) the cirrus over the ice sheet, (ii) the area of low cloudiness over ocean, and (iii) the clear sky in between the two distinct cloud regimes. In particular the detection of thin ice cloud over the ice sheet is encouraging as melt regions over the GrIS will be primary observational targets for PREFIRE (L'Ecuyer et al. 2021). For this particular orbital segment, the hit rate for PC-MSK is 88.9% and for ML-MSK

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

is 86.2%. This orbital segment is somewhat atypical in that PC-MSK slightly outperforms ML-MSK by 2.7%, whereas the global statistics suggest ML-MSK outperforms PC-MSK by 2.7% (see Section 4.2). Overall, these results suggest a large degree of consensus between the two independent cloud mask algorithms that are capable of detecting clouds within very complex meteorological environments.

The next orbital segment is located over the heart of the GrIS and is shown in Fig. 8. The hit rate for ML-MSK is 94.3% and for PC-MSK is 82.3%. The vast majority of misclassifications for PC-MSK (16.9%) are clear footprints labeled as cloudy. As discussed in Section 3, the PC-MSK is sensitive to the training data used, and is particularly sensitive within complex scenes that may not have sufficient sampling for training spectra. Further work on an optimal sampling strategy is warranted post-launch once TIRS spectra are available (see Section 5). In Fig. 8, the ML-MSK and PC-MSK misassign different footprints of clear or cloud rather than a common misidentification. This is evidenced by the vast majority of misclassifications for ML-MSK (5.6%) are cloudy footprints that are assigned as clear. This behavior is also atypical in the simulated day investigated (01 January 2021) as other orbital segments (not shown) have a significant number of common footprints that are misclassified between the two masks.

A tropical orbital segment in proximity to the Philippines is shown in Fig. 9 in which both ML-MSK and PC-MSK have identical hit rates of 99.8%. This scene has a uniform T_{skin} with clear and cloudy radiances that substantially differ in magnitude from each other (not shown), unlike the radiances shown for the orbital segment near northeast Greenland (Fig. 4). The PC-MSK appears to be insensitive to the sampling in this tropical scene. Out of the 25 random draws of clear and cloud performed in this region, the lowest hit rate obtained with PC-MSK is ~98%. The deep tropics were generally insensitive to the random draws used for the training data sets.

4.2 Global and polar statistics

The global hit rates for ML-MSK and PC-MSK using COD-MSK as truth are shown in Fig. 10. The global hit rate for ML-MSK over all surfaces is 95.3% while for PC-MSK it is 92.6%, with ML-MSK outperforming PC-MSK by 2.7%. As expected, the hit rates are highest over oceans with ML-MSK at 96.8% and PC-MSK at 94.5%. The most challenging category for both masks is a mix of land and ocean along coastlines or near inland bodies of water. In these cases, the hit rates for ML-MSK is 92.9% and for PC-MSK it is 88.3%.

The hit rates in the Arctic (Fig. 11) and the Antarctic (Fig. 12) are reduced relative to the global averages but suggest both cloud masks are able to meet PREFIRE cloud detection requirements of 80–90% of clear-sky occurrences. The Arctic hit rate for ML-MSK over all surfaces is 89.8%, confirming the results of B23, while for PC-MSK it is 86.7%, with ML-MSK outperforming PC-MSK by 3.1%. As with the case of global hit rates (Fig. 10), the most challenging category for both masks is a mix of land and ocean along coastlines or near inland bodies of water. In these cases, the hit rate for ML-MSK is 87.1% and for PC-MSK it is 84.8%. Interestingly, the hit rates over land surfaces (which includes the GrIS) are very similar for both masks in the Arctic and Antarctic with ML-MSK at 89.8% and PC-MSK at 89.6%. The hit rates over ocean are very similar

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

to that of land for ML-MSK at 89.6% and PC-MSK at 85.5%. This suggests that some of the most difficult clouds to detect are located over the Arctic Ocean and sea ice, especially low clouds that are capped by thermal inversions (B23).

Hit rates are higher than the global average in both midlatitude bands (not shown), but particularly in the Tropics (not shown) where hit rates are 98.8% for ML-MSK and 96.6% for PC-MSK over all surfaces. Both cloud detection approaches perform better in scenes in which the clear and cloudy radiances are much different in magnitude. As expected, this simulation study confirms that the biggest challenges for validating and improving the cloud masks will be in the polar regions.

5 Discussion

Channel-dependent instrument noise has been included in this simulation study although it is fixed to a constant value. Each of the eight cross-track footprints in the two PREFIRE CubeSats will have unique noise and channel SRF characteristics that will require adjustments to the PC-MSK algorithm. The algorithm will be refined once TIRS data is available post-launch and the scene-dependent noise impacts on the PC-MSK will be accounted for.

The effects of mixed surface types are not considered within single TIRS footprints with surface emissivity fixed to the nearest neighbor IGBP surface type data set (Huang et al., 2016). Sub-footprint variability is likely to exhibit spectral signatures in TIRS-observed radiances that will require careful examination in conjunction with Visible Infrared Imaging Radiometer Suite (VIIRS) imagery. Further examination of the cloud mask performance will be quantified using microwave (MW)-derived surface types using the Advanced Technology Microwave Sounder (ATMS) following the approach in Kahn et al. (2020). Additionally, annually-updated VIIRS surface type will be available in Aux-Met, along with daily-updated sea ice and snow cover from VIIRS, the Advanced Microwave Scanning Radiometer (AMSR), and the Special Sensor Microwave Imager/Sounder (SSMIS) that are available in the PREFIRE Aux-Sat product.

For the operational version of PC-MSK, clear spectra will be calculated for every TIRS footprint including cloudy TIRS footprints. The clear training sets will be formed from combinations of (1) simulated clear-sky TIRS spectra, and (2) observed clear-sky TIRS spectra detected by ML-MSK, the VIIRS cloud mask, and the Cross-track Infrared Microwave Sounder Suite (CrIMSS) effective cloud fraction (ECF). The simulated clear sky is not likely to have the richness and complexity of surface variability that will be present in the observed TIRS radiance spectra. This could be an advantage for cloud detection as some spectral behavior reflected in the PCs may be unambiguous indicators of complex surface behavior observed in clear sky.

The spectra used in the training sets could be drawn from simulations or observations, or some combination thereof. (i) Spectra can be drawn from GEOS FP-IT clear-sky simulated spectra available for every TIRS footprint, and our offline archive of cloudy-sky simulated spectra available for the first day of every month in 2021. (ii) Spectra can be drawn from observed TIRS footprints that are identified as clear or cloudy using VIIRS and CrIMSS coincidences. (ii-a) For close temporal coincidences, the VIIRS cloud mask and the CrIMSS ECF will be used to identify unambiguously clear or cloudy TIRS footprints to form a global database of spectra. Then, random draws can be used to form training sets. (ii-b) The ML-MSK can be used to identify clear and cloudy spectra that will be used to form a random draw for the training. However, a drawback from this approach is

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

a reduced independence of PC-MSK from ML-MSK. (ii-c) A hybrid combination of (i) and (ii) is attempted. While a close temporal coincidence between JPSS and PREFIRE satellites is required for an unambiguous determination of clear-sky and cloudy-sky TIRS radiance spectra, the training set itself does not necessarily require a close temporal coincidence. All training methods will be compared and their quantitative performance will be assessed by surface and scene type during post-launch check-out before the data release, with the focus on the mission requirement for clear-sky detection.

The clear-sky and cloudy-sky training sets will be operationally tested using three approaches. (1) The global methodologies outlined by Turner et al. (2006) and M19 will be used to form training sets for latitudinal bands or large geographical regions that subdivide land and ocean, and are furthermore a function of the seasonal cycle. (2) The local methodology outlined in this simulation experiment will be applied. (3) A hybrid local and global approach will be tested (e.g., Hultberg et al., 2017). Adaptable and flexible training sets could emphasize (1) in the tropics where the global approach works well with PREFIRE, and (2) in more complex scenes such as the Arctic.

Initial validation of PC-MSK and ML-MSK will be performed against the VIIRS cloud mask and CrIMSS ECF for orbital segments that are not used in the formation of the training data sets. As with the identification of ideal training set radiances, the most optimal comparisons will be limited to small time differences (~minutes) between PREFIRE and JPSS. Inevitably, the time restriction is a function of variable atmospheric conditions, implying a more (less) restrictive time difference, for instance, when wind speed is higher (lower). The meteorological context will leverage the Aux-Met files to examine temporal differences between PREFIRE and JPSS coincidences as a function of both wind speed and direction shear. This initial validation effort will play a crucial role in optimizing the training data for the operational version of PC-MSK.

Lastly, we recommend using PC-MSK and ML-MSK together to ensure clear-sky detection with the highest level of confidence. The ML-MSK reports an uncertainty estimate, while PC-MSK reports the value of SID with respect to the threshold determined from Otsu's method. These additional quantities will offer further information regarding the likelihood of clear-sky or cloudy-sky detected within a given TIRS footprint. The PREFIRE Science Team will investigate the synergistic use of PC-MSK and ML-MSK for downstream products during the post-launch checkout phase.

6 Summary

This work describes a cloud mask simulation experiment for the PREFIRE mission. Each of the two PREFIRE cubesats will have a TIRS instrument with 54 channels that span the range of 4 to 53 μm with a nominal full-width, half-max SRF of 0.84 μm. Cloud detection in the infrared spectrum is challenging because of pervasive and wide-ranging complexity that includes, but is not limited to, spatial and temporal variations in cloud sizes and types, highly diverse vertical structures of temperature and moisture, and surface types with poorly known characteristics such as surface emissivity. The detection is further complicated by the relatively coarse spectral width that convolves fine gas absorption line features with broad spectral features of cloud and surface properties. The PREFIRE retrieval algorithm flow starts with clear-sky and cloudy-sky detection for every TIRS footprint. To ensure a robust cloud mask, the PREFIRE team developed two cloud masks: the first is a machine learning-

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

based cloud mask (ML-MSK) described by Bertossa et al. (2023), and the second described in this work is a principal component-based cloud mask (PC-MSK) based on the algorithm developed for the upcoming FORUM mission.

A set of TIRS radiance spectra are calculated from GEOS FP-IT meteorological and surface data on 01 January 2021, and are used in the PCRTM v3.4 radiative transfer model that is configured to the TIRS prelaunch SRFs. The PC-MSK is locally trained along simulated PREFIRE orbits, unlike previous studies that have used global training sets. A similarity index difference (SID) is calculated from information-containing eigenvectors and PCs for segments of 2000 TIRS footprints that typically span 10–15° in latitude. The SID is formed into multimodal histograms with separate but occasionally overlapping modes for clear and cloudy skies. A numerically stable method of separating the clear-sky mode from the cloudy-sky mode is achieved using Otsu's method of binary classification. Unlike previous works that use the SID methodology, Otsu's method requires no a priori estimate for establishing a threshold between the histogram modes.

Hit rates of ML-MSK and PC-MSK against "truth" are divided into categories by latitude band, land, ocean, and coastline. Global hit rates for PC-MSK (92.6%) compare well to hit rates for ML-MSK (95.3%). More specifically, Arctic hit rates for PC-MSK (86.7%) compare favorably to hit rates for ML-MSK (89.4%). Both of these hit rates meet mission requirements for PREFIRE cloud detection — 80–90% of Arctic observations. The two cloud masks are entirely independent of each other in this simulation study. Since the simulated training data set for ML-MSK is based on 3-km FV3 GFDL simulations, while PC-MSK is based on GEOS FP-IT simulations, the two algorithms have a fair degree of independence. The results of this experiment are encouraging, as they reveal a high level of agreement between these independent approaches. Despite a low number of information-containing PCs in TIRS spectra compared to those obtained from hyperspectral infrared sounders, PREFIRE should be capable of meeting the mission requirements for clear-sky detection.

Code and Data Availability

The simulated PREFIRE radiances and associated GEOS FP-IT model output are part of a pre-launch development database and are available directly from the lead author. The code used to produce the cloud mask is also part of the pre-launch development code base. The final PC-MSK code will be made publicly available once PREFIRE data is publicly available a few months after launch in May 2024.

Author Contributions

BHK designed the experiment, led the writing of the manuscript and creation of the tables and figures, and led the adaptation of the algorithm based on the publication of Maestri et al. (2019). CB provided the ML-MSK data and expertise. XC and XH developed the simulation package of PREFIRE spectra based on PCRTM and GEOS FP-IT data. BJD provided specifics on the TIRS instrument and estimated NEdT levels. EHW, TM, TSL, FD, TM, and MM helped with the conceptual design and

writing of the manuscript. AM developed the orbit simulator flown through gridded GEOS FP-IT data and simulated PREFIRE radiances. NBM provided clear-sky simulated radiances. All authors contributed to writing of the manuscript.

Competing Interests

The contact author has declared that none of the authors has any competing interests.

Acknowledgments

A portion of this research was carried out at the Jet Propulsion Laboratory (JPL), California Institute of Technology, under a contract with the National Aeronautics and Space Administration. This work was supported by NASA under the Earth Ventures-Instrument (EV-I) program's Polar Radiant Energy in the Far-Infrared Experiment (PREFIRE) mission. The simulated TIRS radiances were generated at the Space Science and Engineering Center (SSEC) at the University of Wisconsin–Madison. BHK is grateful to Qing Yue for conversations that helped refine the algorithm technique for PC-MSK.

References

Bertossa, C., L'Ecuyer, T., Merrelli, A., Huang, X., and Chen, X.: A neural network-based cloud mask for PREFIRE and evaluation with simulated observations, J. Atmos. Ocean Tech., 40, 377–396, https://doi.org/10.1175/JTECH-D-22-0023.1, 2023.

Chen, X., Huang, X., and Liu, X., 2013: Non-negligible effects of cloud vertical overlapping assumptions on longwave spectral fingerprinting studies, J. Geophys. Res. Atmos., 118, 7309–7320, https://doi.org/10.1002/jgrd.50562, 2013.

Chen, X., Huang, X., and Flanner, M. G.: Sensitivity of modeled far-IR radiation budgets in polar continents to treatments of snow surface and ice cloud radiative properties, Geophys. Res. Lett., 41, 6530–6537, doi:10.1002/ 2014GL061216, 2014.

Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., et al., 2005: Atmospheric radiative transfer modeling: A summary of the AER codes. J. Quant. Spectrosc. Radiative Trans., 91, 233–244, https://doi.org/10.1016/j.jqsrt.2004.05.058, 2005.

Cossich, W., Maestri, T., Magurno, D., Martinazzo, M., Di Natale, G., Palchetti, L., Bianchini, G., and Del Guasta, M.: Ice and mixed-phase cloud statistics on the Antarctic plateau, Atmos. Chem. Phys., 21, 13811–13833, https://doi.org/10.5194/acp-21-13811-2021, 2021.

Drouin, B. J., Kahn, B., Lim, B., Merrelli, A., Nelson, E., Quinn, G., Nagle, F., and L'Ecuyer, T.: Orbital Trade Study for the PREFIRE Mission, IEEE Aerospace Conference (AERO), Big Sky, MT, USA, 1-7, https://doi.org/10.1109/AERO53065.2022.9843312, 2022.

Feldman, D. R., Collins, W. D., Pincus, R., Huang, X., and Chen, X.: Far-infrared surface emissivity and climate. Proc. Natl. Acad. Sci., 111, 16297– 16302. https://doi.org/10.1073/pnas.1413640111, 2014.

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

Fetzer, E. J., Lambrigtsen, B. H., Eldering, A., Aumann, H. H., and Chahine, M. T.: Biases in total precipitable water vapor climatologies from Atmospheric Infrared Sounder and Advanced Microwave Scanning Radiometer, J. Geophys. Res., 111, D09S16, doi:10.1029/2005JD006598, 2006.

Glasbey C.A., An analysis of histogram-based thresholding algorithms, CVGIP: Graphical models and image processing, 55, 532–537, https://doi.org/10.1006/cgip.1993.1040, 1993.

Harries, J., Carli, B., Rizzi, R., Serio, C., Mlynczak, M., Palchetti, L., Maestri, T., Brindley, H., and Masiello, G.: The far-infrared Earth, Rev. Geophys., 46, RG4004, https://doi.org/10.1029/2007RG000233, 2008.

Huang, H.-L., and Antonelli, P.: Application of principal component analysis to high-resolution infrared measurement compression and retrieval, J. Appl. Meteor., 40, 365–388, https://doi.org/10.1175/1520-0450(2001)040<0365:AOPCAT>2.0.CO;2, 2001.

Huang, X., Chen, X., Zhou, D. K., and Liu, X.: An Observationally Based Global Band-by-Band Surface Emissivity Dataset for Climate and Weather Simulations. J. Atmos. Sci., 73, 3541–3555, https://doi.org/10.1175/JAS-D-15-0355.1, 2016.

Hultberg, T., August, T., and Lenti, F.: Local or global? How to choose the training set for principal component compression of hyperspectral satellite measurements: a hybrid approach, Proc. SPIE 10423, Sensors, Systems, and Next-Generation Satellites XXI, 104231G (29 September 2017); https://doi.org/10.1117/12.2278349, 2017.

Kahn, B. H., Drouin, B. J., and L'Ecuyer, T. S.: Assessment of sampling sufficiency for low-cost satellite missions: Application to PREFIRE, J. Atmos. Ocean. Tech., 37, 2283–2298, https://doi.org/10.1175/JTECH-D-20-0023.1, 2020.

Kittler, J., and Illingworth, J.: Minimum Error Thresholding, Pattern Recognition, 19, 41–47, https://doi.org/10.1016/0031-3203(86)90030-0, 1986.

Kapur, J. N., Sahoo, P. K., and Wong, A. K. C.: A new method for gray-level picture thresholding using the entropy of the histogram, Graph. Models Image Process., 29, 273–285, 1985.

L'Ecuyer, T. S., Drouin, B. J., Anheuser, J., Grames, M., Henderson, D. S., Huang, X., Kahn, B. H., Kay, J. E., Lim, B. H., Mateling, M., Merrelli, A., Miller, N. B., Padmanabhan, S., Peterson, C., Schlegel, N.-J., White, M. L., and Xie, Y.: The Polar Radiant Energy in the Far Infrared Experiment, Bull. Amer. Met. Soc., 102, 1431–1449, https://doi.org/10.1175/BAMS-D-20-0155.1, 2021.

Libois, Q., Proulx, C., Ivanescu, L., Coursol, L., Pelletier, L. S., Bouzid, Y., Barbero, F., Girard, É., and Blanchet, J.-P.: A microbolometer-based far infrared radiometer to study thin ice clouds in the Arctic, Atmos. Meas. Tech., 9, 1817–1832, https://doi.org/10.5194/amt-9-1817-2016, 2016.

Libois, Q., and Blanchet, J.-P.: Added value of far-infrared radiometry for remote sensing of ice clouds, J. Geophys. Res. Atmos., 122, 6541–6564, https://doi.org/10.1002/2016JD026423, 2017.

Liu, X., Smith, W. L., Zhou, D. K., and Larar, A.: Principal component-based radiative transfer model for hyperspectral sensors: theoretical concept, Appl. Opt., 45, 201–209, https://doi.org/10.1364/AO.45.000201, 2006.

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

Liu, X., He, T., Sun, L., Xiao, X., Liang, S., and Li, S.: Analysis of daytime cloud fraction spatiotemporal variation over the Arctic from 2000 to 2019 from multiple satellite products, J. Climate, 35, 3995–4023, https://doi.org/10.1175/JCLI-D-22-0007.1, 2022.

Lucchesi, R.: File Specification for GEOS-5 FP-IT. GMAO Office Note No. 2 (Version 1.3), 60 pp, available from https://gmao.gsfc.nasa.gov, 2015.

Maestri, T., Cossich, W., and Sbrolli, I.: Cloud identification and classification from high spectral resolution data in the far infrared and mid-infrared, Atmos. Meas. Tech., 12, 3521–3540, https://doi.org/10.5194/amt-12-3521-2019, 2019.

Mahajan, S., and Fataniya, B.: Cloud detection methodologies: variants and development—a review. Complex Intell. Syst., 6, 251–261, https://doi.org/10.1007/s40747-019-00128-0, 2020.

Miller, N. B., Merrelli, A., L'Ecuyer, T. S., and Drouin, B. J.: Simulated clear-sky water vapor and temperature retrievals from PREFIRE measurements, J. Atmos. Ocean Tech., https://doi.org/10.1175/JTECH-D-22-0128.1, 2023.

Molod, A., Takacs, L., Suarez, M., and Bacmeister, J.: Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2, Geosci. Model Dev., 8, 1339–1356, https://doi.org/10.5194/gmd-8-1339-2015, 2015.

Otsu, N.: A threshold selection method from gray-level histograms, *IEEE Transactions on Systems, Man, and Cybernetics*, 9, 62–66, https://doi.org/10.1109/TSMC.1979.4310076, 1979.

Palchetti, L.; Brindley, H.; Bantges, R.; Buehler, S. A.; Camy-Peyret, C.; Carli, B.; Cortesi, U.; Del Bianco, S.; Di Natale, G.; Dinelli, B. M.; Feldman, D.; Huang, X. L.; C.-Labonnote, L.; Libois, Q.; Maestri, T.; Mlynczak, M. G.; Murray, J. E.; Oetjen, H.; Ridolfi, M.; Riese, M.; Russell, J.; Saunders, R.; Serio, C.: FORUM: unique far-infrared satellite observations to better understand how Earth radiates energy to space, Bull. Amer. Met. Soc., 101, E2030–E2046 https://doi.org/10.1175/BAMS-D-19-0322.1, 2020.

Peterson, C. A., Huang, X., Chen, X., and Yang, P.: Synergistic use of far- and mid-infrared spectral radiances for satellite-based detection of polar ice clouds over ocean. Journal of Geophysical Research: Atmospheres, 127, e2021JD035733, https://doi.org/10.1029/2021JD035733, 2022.

Ridler, T., and Calvard, S.: Picture Thresholding Using an Iterative Selection Method, IEEE Transactions on Systems, Man, and Cybernetics, 8, 630–632, https://doi.org/10.1109/TSMC.1978.4310039, 1978.

Saito, M., Yang, P., Huang, X., Brindley, H. E., Mlynczak, M. G., and Kahn, B. H.: Spaceborne middle-and far-infrared observations improving nighttime ice cloud property retrievals. Geophys. Res. Lett., 47, e2020GL087491, https://doi.org/10.1029/2020GL087491, 2020.

Sezgin, M., and Sankur, B.: Survey over image thresholding techniques and quantitative performance evaluation, J. Electron. Imag., 13, 146–168, https://doi.org/10.1117/1.1631315, 2004.

Sgheri, L., Belotti, C., Ben-Yami, M., Bianchini, G., Carnicero Dominguez, B., Cortesi, U., Cossich, W., Del Bianco, S., Di Natale, G., Guardabrazo, T., Lajas, D., Maestri, T., Magurno, D., Oetjen, H., Raspollini, P., and Sgattoni, C.: The FORUM end-to-end simulator project: architecture and results, Atmos. Meas. Tech., 15, 573–604, https://doi.org/10.5194/amt-15-573-2022, 2022.

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

Tsai, W.: Moment-preserving thresholding: a new approach, Computer Vision, Graphics, and Image Processing, 29, 377–393, 1985.

Turner, D. D., Knuteson, R. O., Revercomb, H. E., Lo, C., and Dedecker, R. G.: Noise reduction of Atmospheric Emitted Radiance Interferometer (AERI) observations using principal component analysis, J. Atmos. Ocean Tech., 23, 1223–1238, https://doi.org/10.1175/JTECH1906.1, 2006.

Turner, D. D., and Mlawer, E. J.: The Radiative Heating in Underexplored Bands Campaigns. Bull. Amer. Meteor. Soc., 91, 911–924, https://doi.org/10.1175/2010BAMS2904.1, 2010.

Xie, Y., Huang, X., Chen, X., L'Ecuyer, T. S., Drouin, B. J., and Wang, J.: Retrieval of surface spectral emissivity in polar regions based on the optimal estimation method, Journal of Geophysical Research: Atmospheres, 127, e2021JD035677, https://doi.org/10.1029/2021JD035677, 2022.

Xue, J.-H., and Zhang, Y.-J.: Ridler and Calvard's, Kittler and Illingworth's and Otsu's methods for image thresholding, Pattern Recognition, 33, 793–797, https://doi.org/10.1016/j.patrec.2012.01.002, 2012.

Yue, Q., Fetzer, E. J., Kahn, B. H., Wong, S., Manipon, G., Guillaume, A., and Wilson, B.: Cloud-State-Dependent Sampling in AIRS Observations Based on CloudSat Cloud Classification, J. Climate, 26, 8357–8377, https://doi.org/10.1175/JCLI-D-13-00065.1, 2013.

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

Cloud Phase	COD	q_i	q_l
Clear	<0.25	n/a	n/a
Liquid	≥0.25	=0.0	>0.0
Ice	≥0.25	>0.0	=0.0
Liquid+Ice	≥0.25	>0.0	>0.0

Table 1: Categorization of cloud phase used in this evaluation. If COD is between 0.0 and 0.25, the footprint is classified as "clear" following the training threshold used in B23. All criteria requiring 0.0 apply to all vertical levels, while criteria requiring >0.0 can be true for either a single level or multiple levels, but must be true at cloud top.

Maestri et al. (2019)	Present Study
Separate training for clear and four different cloud types (liquid phase, mixed phase, ice phase, and thin cirrus)	Training data sets limited to two classes: clear and cloud (no differentiation by cloud type)
Uses the *distributional method* that requires an independent training set to optimize SID for clear and cloud differentiation	Uses an unsupervised, bimodal class separation method of Otsu (1979) that does not require optimization of SID to differentiate clear and cloud
Training data from 70 clear and 30 cloud spectra	Training data from 20 clear and 20 cloud spectra
Training data sets developed regionally (tropics, midlatitudes, poles) and seasonally	Training data based on randomly drawn spectra from instantaneous orbital segments of 2000 TIRS footprints
FORUM has 1000s of high-spectral-resolution channels	TIRS has 54 low-spectral-resolution channels, 23 of which are used in the present study
The upper limit of information-containing PCs may exceed 100 for large ensembles of observed spectra	The number of information-containing PCs ranges between 3–7 using simulated TIRS radiances

Table 2: List of key characteristics that differ between M19 and the present study.

	COD-MSK	ML-MSK	PC-MSK
Cloudy	467,365 (45.1%)	463,531 (44.7%)	447,790 (43.2%)
Clear	569,571 (54.9%)	573,405 (55.3%)	589,146 (56.8%)

Table 3: Total numbers and percentages of clear-sky and cloudy-sky TIRS footprints for the COD-MSK, ML-MSK, and PC-MSK used in this study. The total number of footprints is 1,036,936 over 16 complete orbits. The totals for COD-MSK do not represent the true cloud frequency as each simulated TIRS footprint is defined as clear if COD < 0.25.

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

General Category	Land Fraction Value in AUX-JPSS	Total Counts (Percentages)
Ocean	≤ 0.1	635,468 (61.3%)
Land	≥ 0.9	311,462 (30.0%)
Coastline	> 0.1 and < 0.9	90,006 (8.7%)

Table 4: Total numbers and percentages of TIRS footprints over ocean, land, and coastline. Ocean includes geographical areas of sea ice. Land includes geographical areas covered in snow and glacial ice. Coastline includes TIRS footprints with mixtures of ocean and land, and this category also includes inland bodies of water.

Region	Latitude	Total Counts (Percentages)
Arctic	60°N–90°N	168,400 (16.2%)
NH Midlatitudes	30°N–60°N	176,095 (17.0%)
Tropics	30°S–30°N	346,736 (33.4%)
SH Midlatitudes	60°S–30°S	176,622 (17.0%)
Antarctic	90°S–60°S	169,083 (16.3%)

Table 5: Total numbers and percentages of TIRS footprints in different latitude bands.

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

560

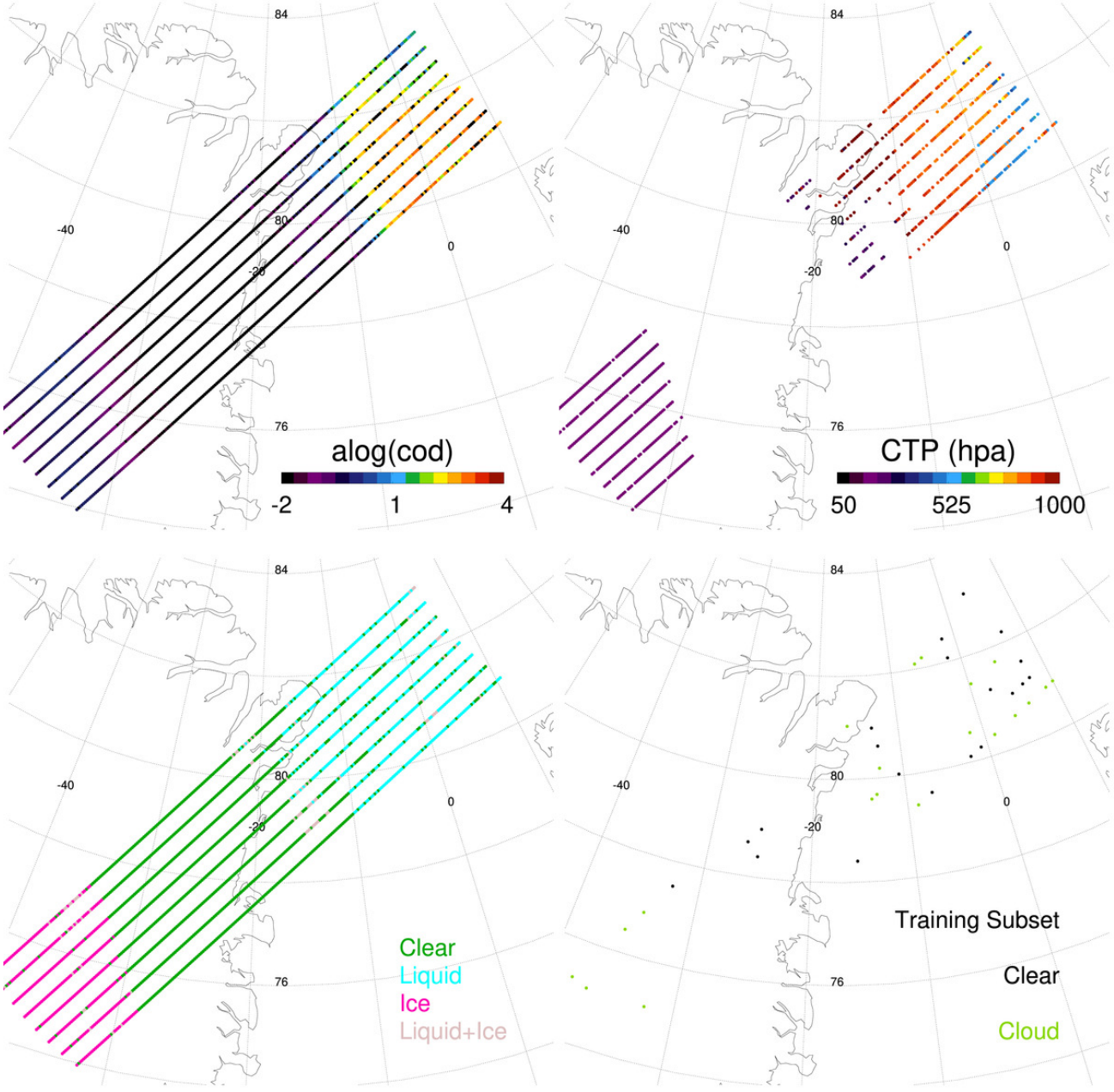


Figure 1: Upper left: Cloud optical depth (COD). Upper right: cloud top pressure (CTP). Lower left: cloud top thermodynamic phase defined in Table 1. Lower right: locations of randomly-sampled clear and cloudy training data. Data are from GEOS FP-IT meteorological fields on 01 January 2021.

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

565

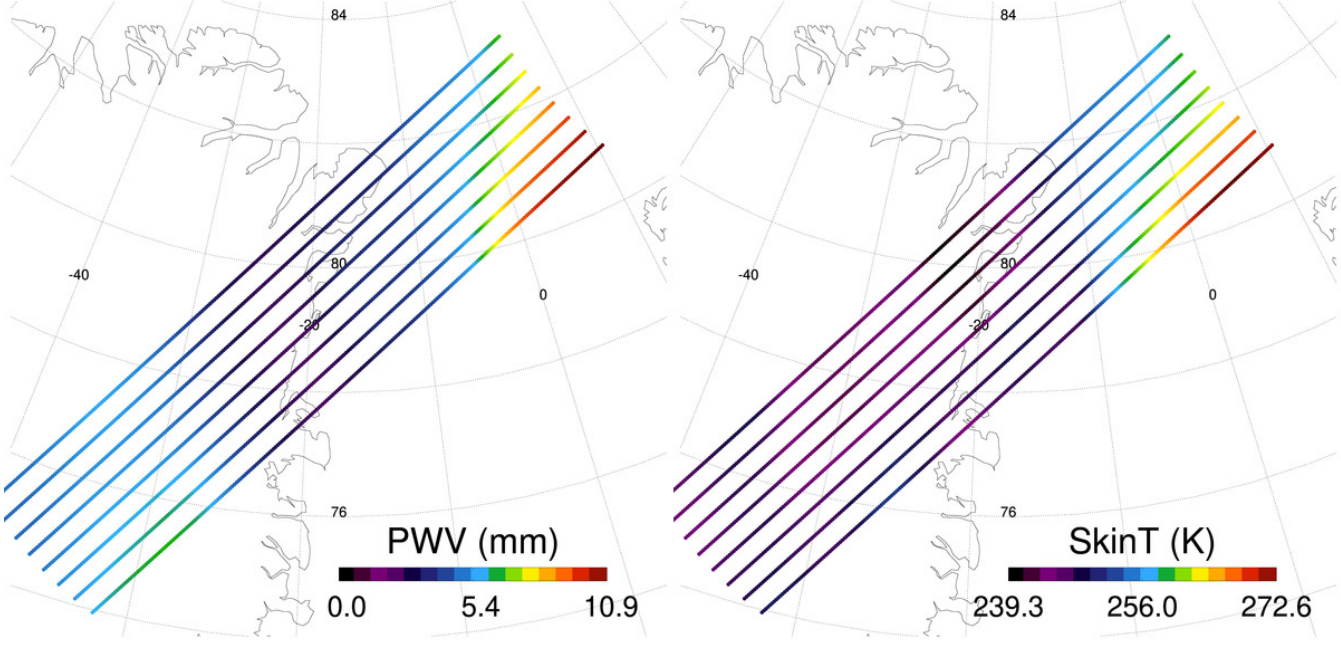


Figure 2: Left: precipitable water vapor (PWV) in millimeters (mm). Right: skin temperature (T_{skin}) in K.

570

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

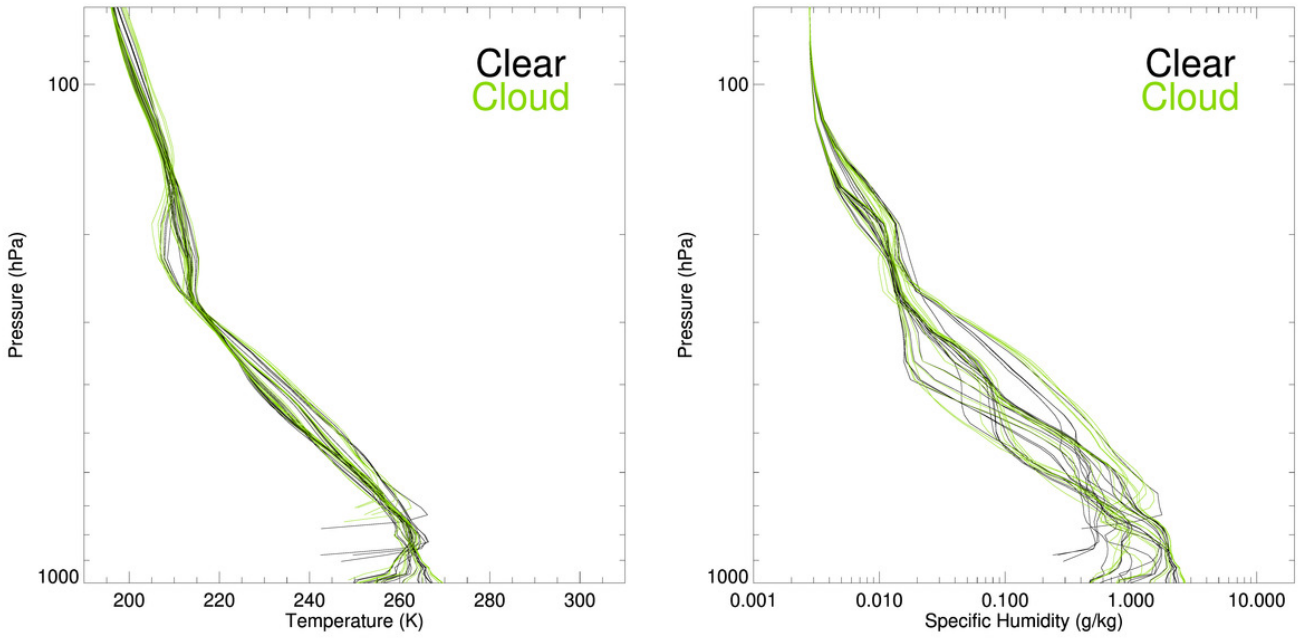


Figure 3: Left: temperature (T) profiles. Right: specific humidity (q) profiles. These profiles are from the clear and cloudy training data shown in Fig. 1.

575

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

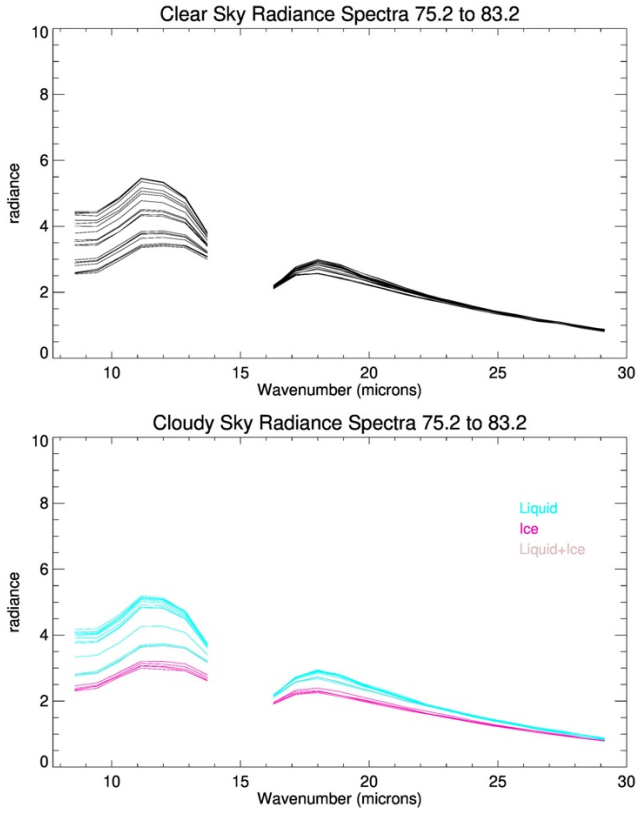


Figure 4: Top: clear-sky TIRS radiances. Bottom: cloudy-sky radiances with assigned phases based on Table 1. These profiles are from the clear and cloudy training data shown in Fig. 1.

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

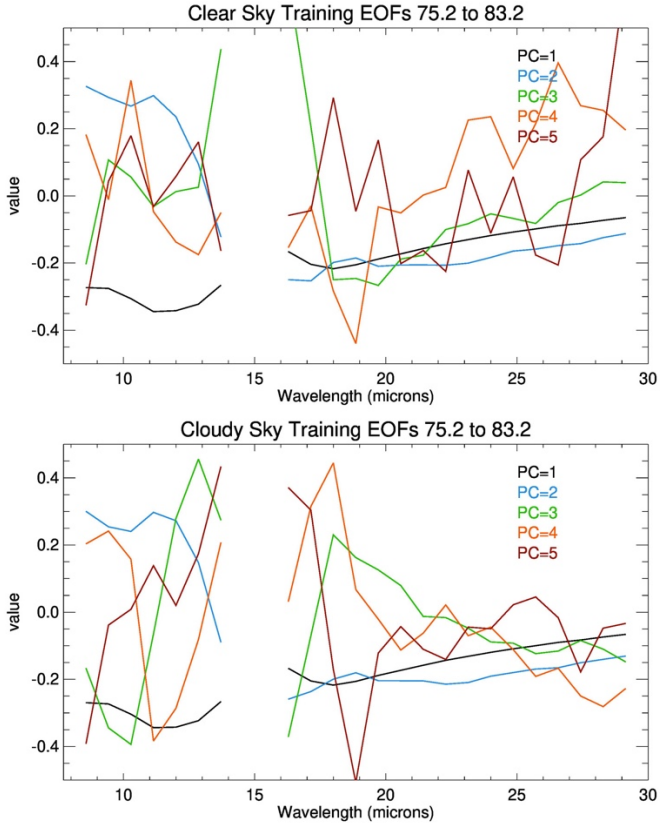
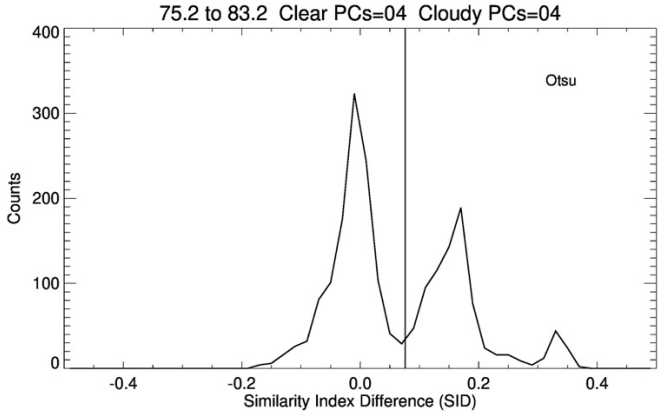


Figure 5: The leading five clear-sky (top) and cloudy-sky (bottom) eigenvectors calculated for the scene depicted in Figs. 1 and 2.

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023



585

Figure 6: The similarity index difference (SID) calculated for the scene depicted in Figs. 1 and 2. The vertical line indicates the separation between clear (on the left) and cloud (on the right) using the method of Otsu (1979).

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

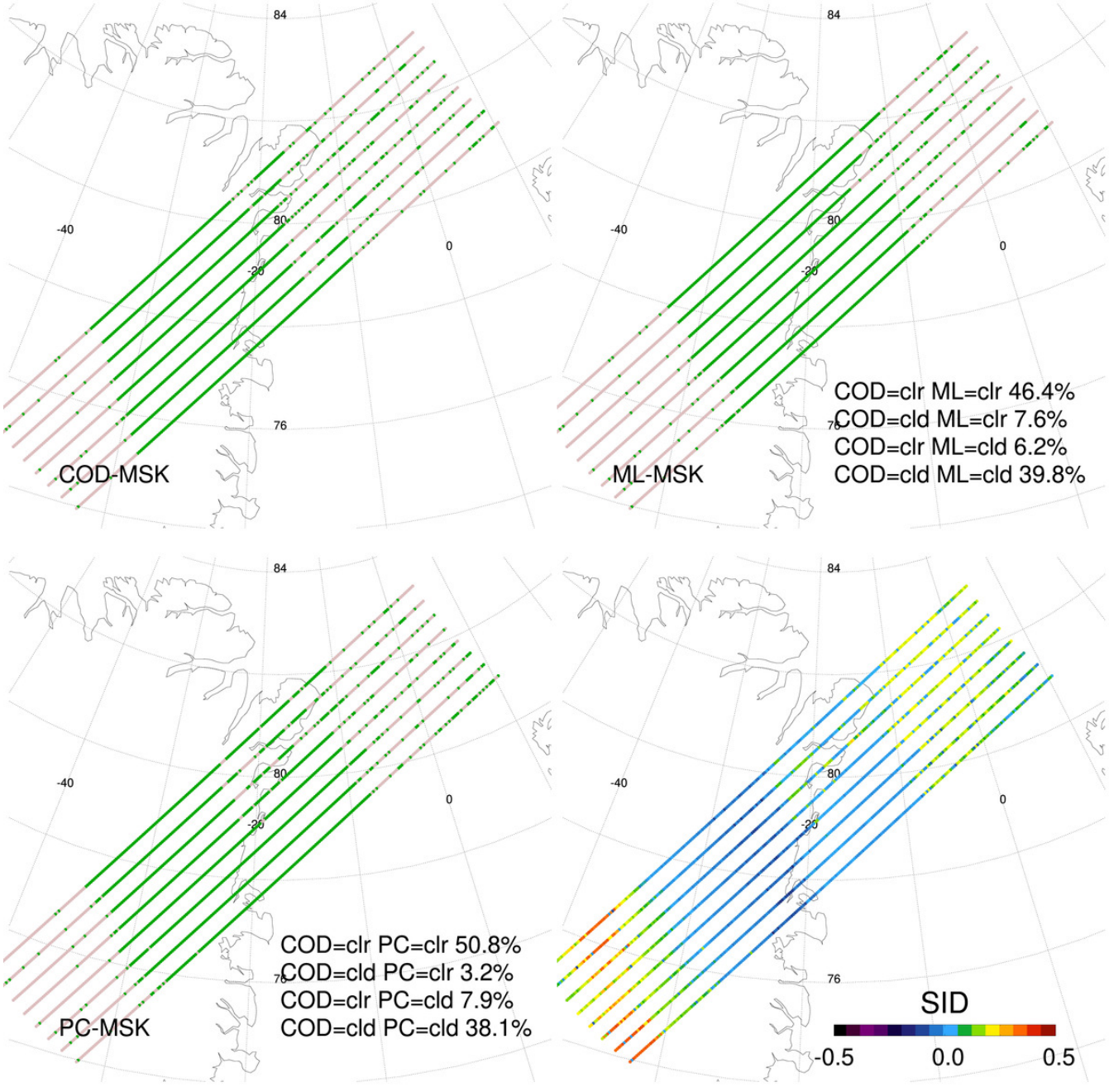


Figure 7: Upper left: the GEOS FP-IT "truth" COD-MSK. Lower left: the PC-MSK. Upper right: ML-MSK. Lower right: map of SID used to make PC-MSK. This is for the same scene depicted in Figs. 1 and 2.

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

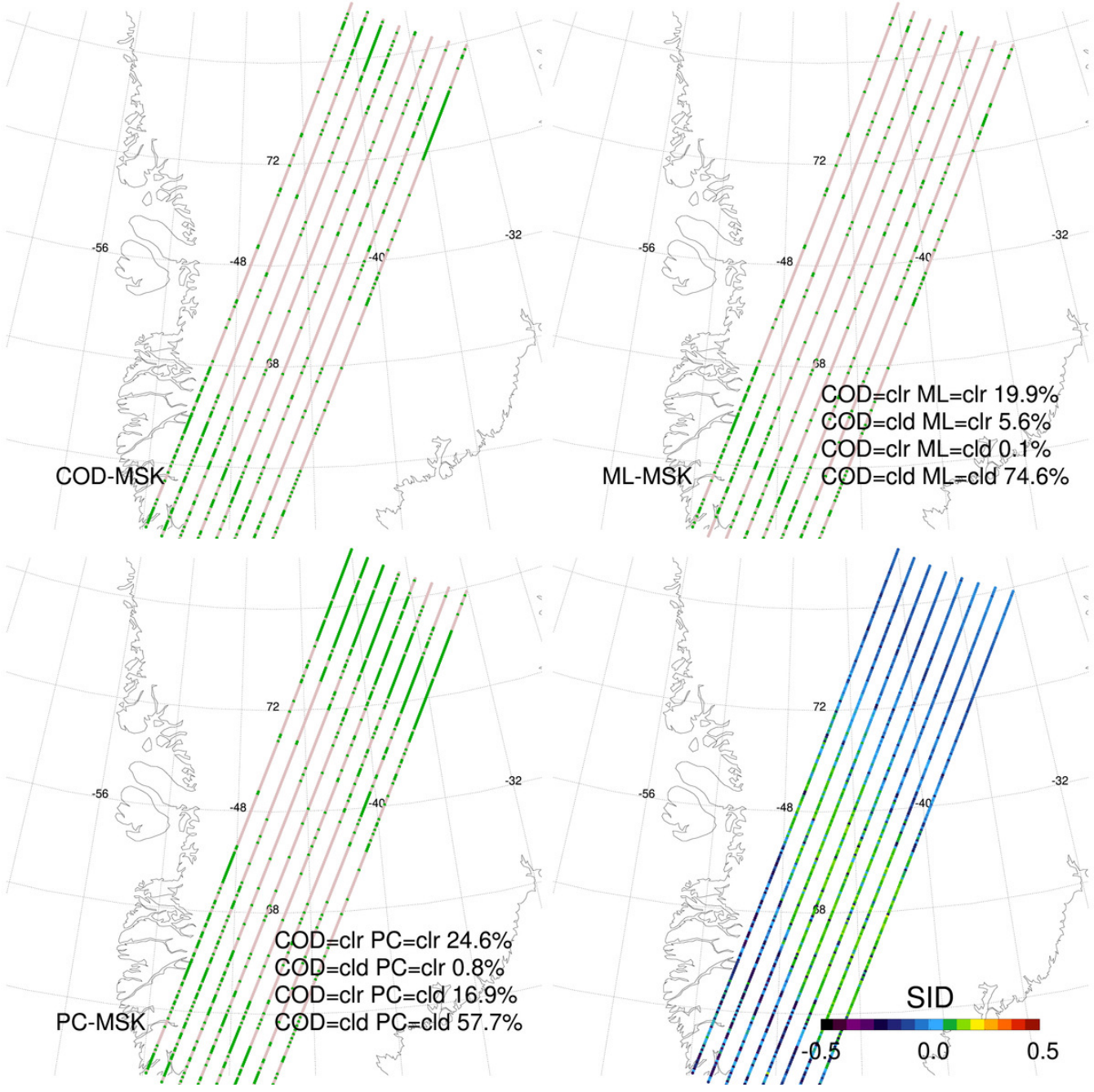


Figure 8: Upper left: the GEOS FP-IT "truth" COD-MSK. Lower left: the PC-MSK. Upper right: ML-MSK. Lower right: map of SID used to make PC-MSK. This is for the next descending orbital segment after the one depicted in Fig. 7.

595

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

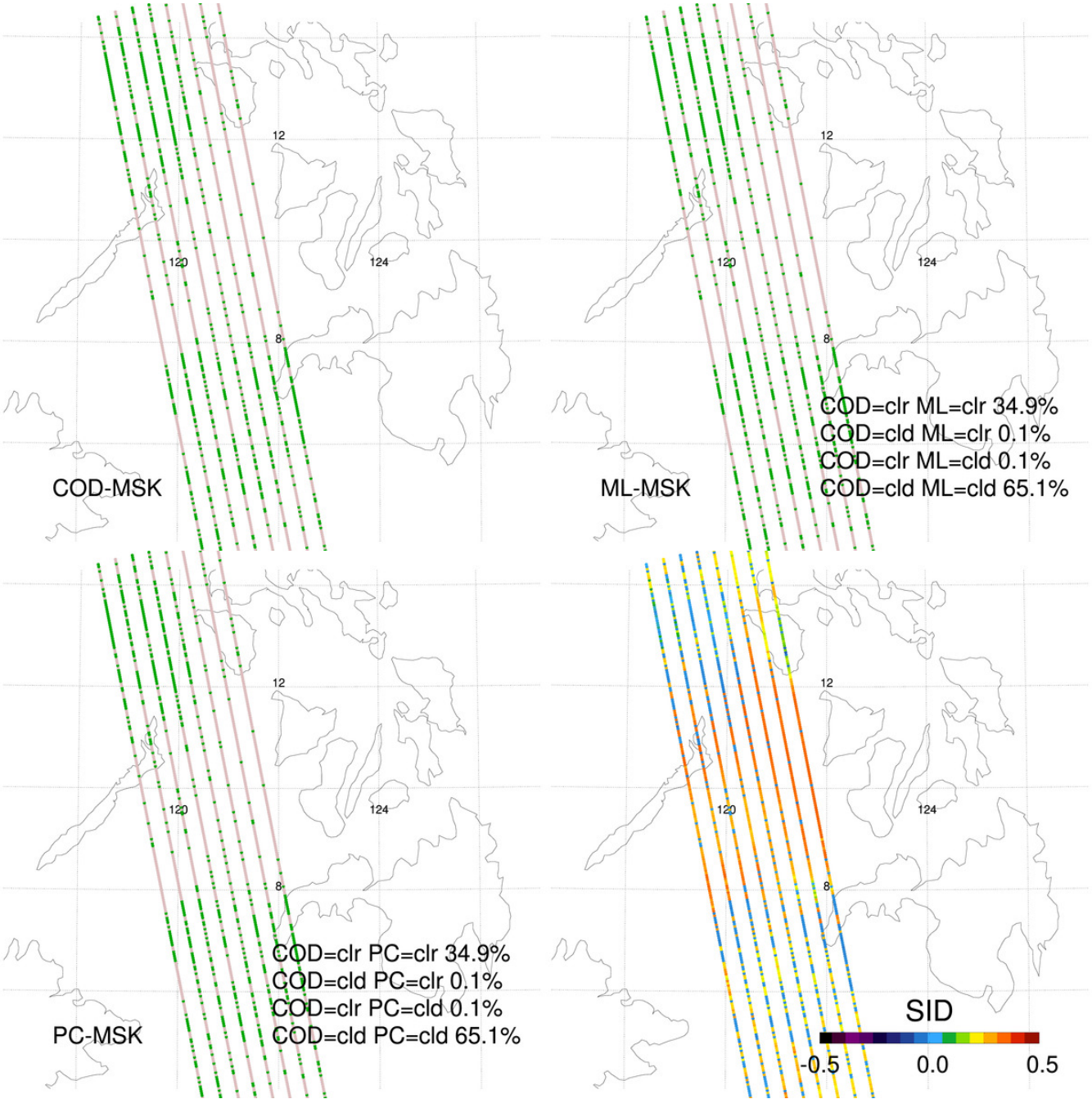
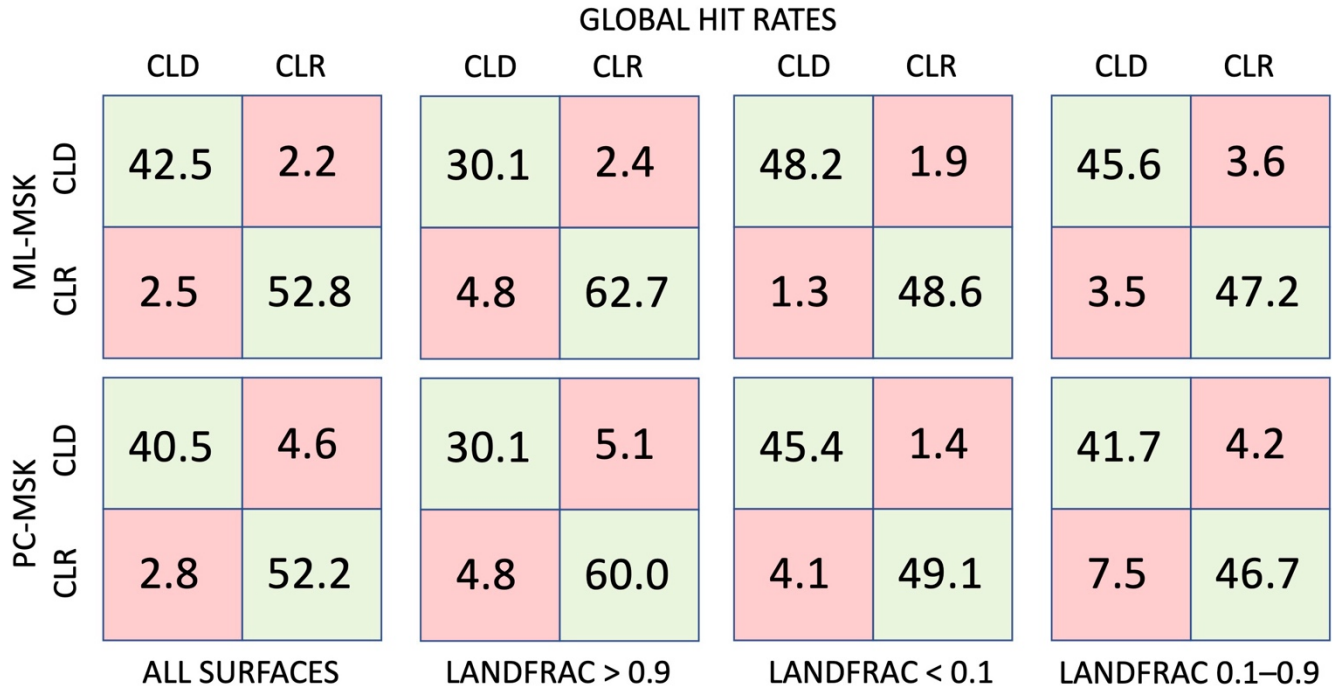


Figure 9: Upper left: the GEOS FP-IT "truth" COD-MSK. Lower left: the PC-MSK. Upper right: ML-MSK. Lower right: map of SID used to make PC-MSK. This is for the first orbital segment in the simulated GEOS FP-IT data set on 01 January 2021.

600

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023



605 **Figure 10: Global hit rates of ML-MSK versus COD-MSK (upper row) and PC-MSK versus COD-MSK (lower row) for all surface types, land surfaces (land fraction > 0.9), ocean surfaces (land fraction < 0.1), and coastlines (land fraction between 0.1 and 0.9). The threshold of clear and cloud is set to COD=0.25 following B23.**

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023

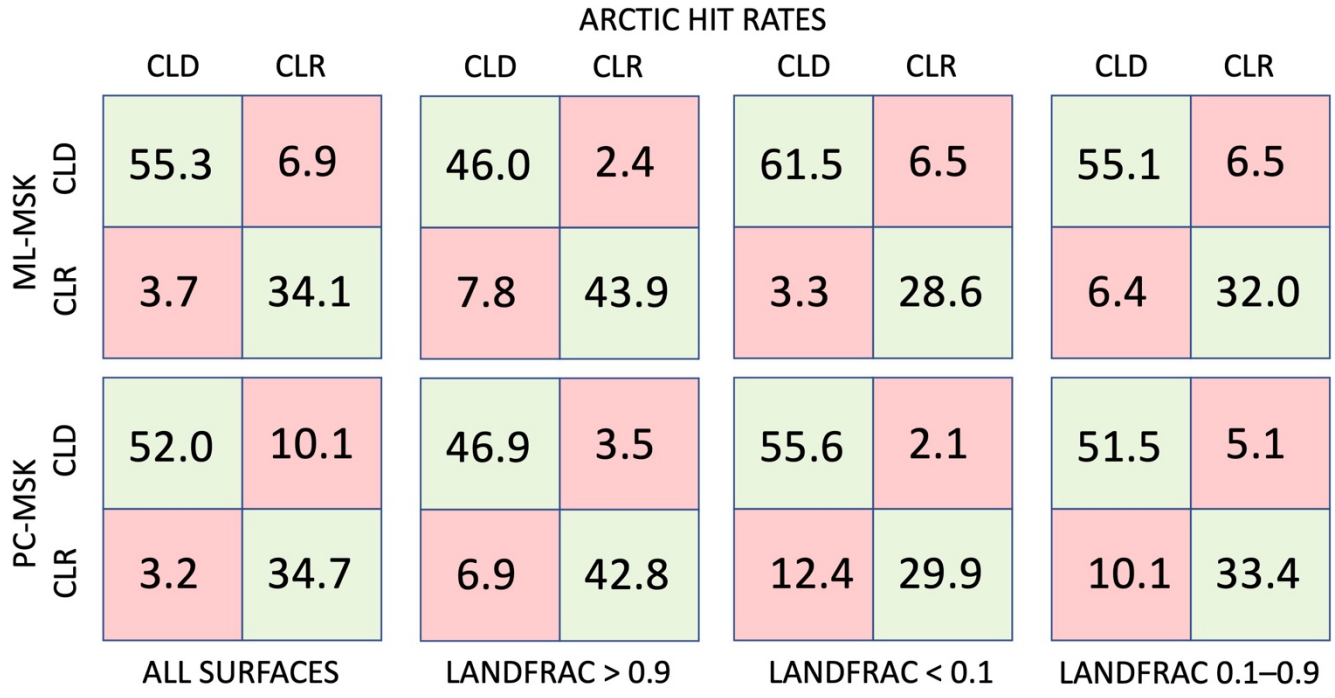
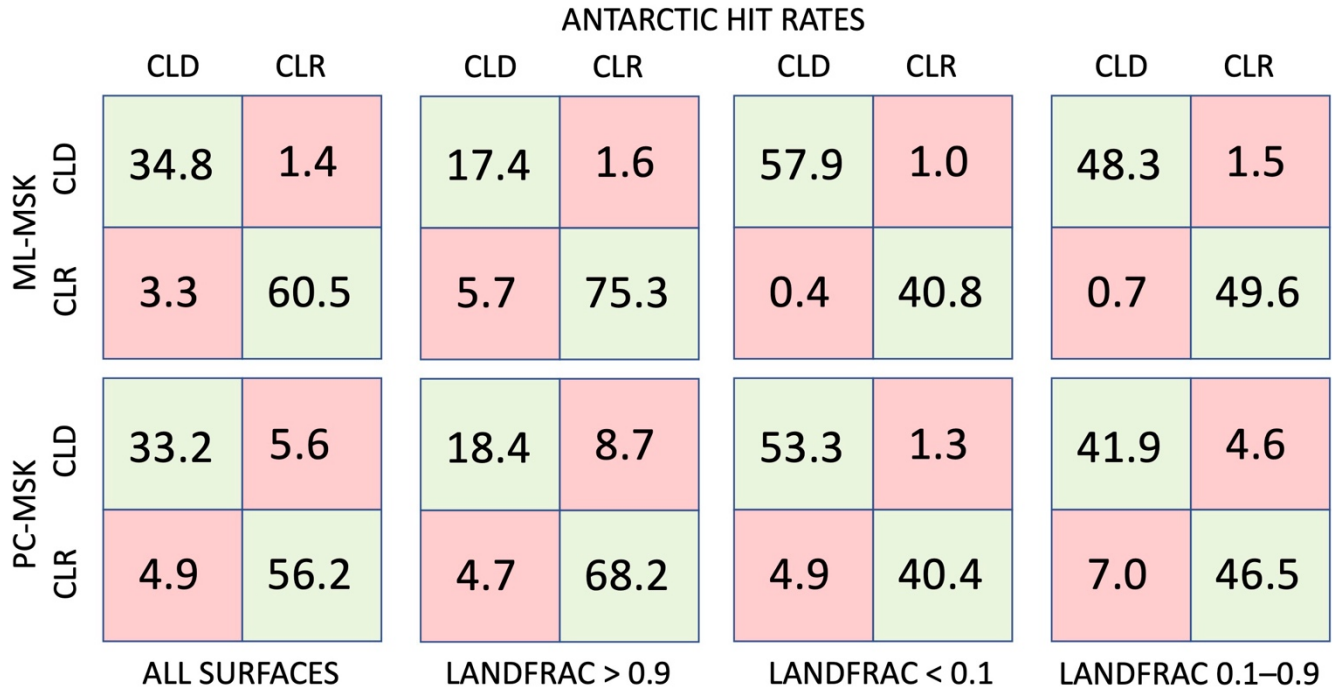


Figure 11: Same as Fig. 10 except for the Arctic region (60°N–90°N).

https://doi.org/10.5194/egusphere-2023-2463
Preprint. Discussion started: 8 November 2023



620

Figure 12: Same as Fig. 10 except for the Antarctic region (60°S–90°S).