JGR Atmospheres

RESEARCH ARTICLE
10.1029/2018JD030242

When Will Spaceborne Cloud Radar Detect Upward Shifts in Cloud Heights?

Hanii Takahashi[1,2], **Matthew D. Lebsock**[2], **Mark Richardson**[1,2], **Roger Marchand**[3], **and Jennifer E. Kay**[4,5]

[1]Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, CA, USA, [2]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, [3]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA, [4]Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder, Boulder, CO, USA, [5]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA

Key Points:

- In the absence of calibration error, pre-2050 detection occurs at some latitudes for all accumulated sensitivity bins below 0 dBZ
- The detection is most likely to occur sooner for the midlatitudes
- Reduced radar calibration error generally has a bigger effect on detection year than lowering the minimum detectable reflectivity

Supporting Information:

- Supporting Information S1

Correspondence to:
H. Takahashi,
Hanii.Takahashi@jpl.nasa.gov

Citation:
Takahashi, H., Lebsock, M. D., Richardson, M., Marchand, R., & Kay, J. E. (2019). When will spaceborne cloud radar detect upward shifts in cloud heights? *Journal of Geophysical Research: Atmospheres*, *124*, 7270–7285. https://doi.org/10.1029/2018JD030242

Received 31 DEC 2018
Accepted 28 MAY 2019
Accepted article online 18 JUN 2019
Published online 8 JUL 2019

Author Contributions:
Conceptualization: Hanii Takahashi, Matthew D. Lebsock
Data curation: Jennifer E. Kay
Formal analysis: Hanii Takahashi, Mark Richardson
Investigation: Hanii Takahashi, Matthew D. Lebsock, Mark Richardson
Methodology: Hanii Takahashi, Matthew D. Lebsock, Mark Richardson
Resources: Jennifer E. Kay
Supervision: Matthew D. Lebsock
Validation: Roger Marchand
Visualization: Hanii Takahashi, Mark Richardson
Writing - original draft: Hanii Takahashi
Writing – review & editing: Hanii Takahashi, Matthew D. Lebsock, Mark Richardson, Roger Marchand, Jennifer E. Kay

Abstract Cloud feedbacks remain the largest source of uncertainty in future climate predictions. Simulations robustly project an increase in cloud height, which is supported by some observational evidence. However, how much of this increasing trend is due to climate warming and how much is due to multiyear natural variability still remains unclear because of the brevity of existing observational records. Here we estimate when the signal will become detectable at 95% confidence by existing radar technology. We use output from a Representative Concentration Pathway 8.5 Community Earth System Model version 1 simulation in a Monte Carlo analysis to determine (1) what is the first year at which changes in the altitude of high cloud can be confidently estimated if we continue to fly W-band cloud radar, (2) what radar sensitivity is required to detect those changes, and (3) at what latitude will we first detect these changes? In Community Earth System Model version 1 a cloud radar record would be able to confidently detect upward shifts in cloud height over 20–60°N before 2030 for a radar with a sensitivity of −15 dBZ and stable calibration errors of ±0.25 dBZ. Furthermore, vertical resolution could be degraded to 1.6 km with little effect on detection year. Results are more sensitive to the magnitude of calibration errors than to the minimum detectable echo. Our earlier midlatitude detection contrasts with a previous lidar-based analysis, which may be due to radar detecting different parts of the clouds and our use of simulations that account for changing geographical patterns of forced warming through time.

1. Introduction

Clouds play an important role in mediating the climate sensitivity as they both cool the Earth by reflecting sunlight and warm the Earth by trapping heat as greenhouse gases do. In order to understand the future climate, it is critical to find clues about how clouds will change. However, the behavior of clouds in response to increasing greenhouse gases is unclear and cloud feedbacks remain the largest source of uncertainty in predictions of future climate (e.g., Soden & Held, 2006; Zelinka et al., 2012) in global climate models.

Many changes in cloud properties are expected with a warming climate including in cloud amount, thickness, height, and phase (e.g., Bony et al., 2009; Storelvmo et al., 2015; Zelinka & Hartmann, 2010). Many of these changes, particularly those associated with cloud amount, are highly variable across the models. A robust cloud response to climate warming predicted by climate simulations is an increase in cloud height following a deepening of the troposphere. Increases in the height of clouds in climate simulations were first noticed in the 1980s (e.g., Hansen et al., 1984; Wetherald & Manabe, 1988), and since then this expectation has been supported by simulations with global climate models and cloud-resolving models (e.g., Kuang & Hartmann, 2007), as well as theory (Zelinka & Hartmann, 2010). Because the rise in cloud top heights is a robust feature of climate simulations, we can make a reasonable hypothesis about when the signal will become detectable by existing observing technology.

Difficulty in understanding cloud feedbacks is, in part, due to a lack of long and stable observation records. A recent study (Norris et al., 2016) investigated evidence for changes using nearly 30 years of passive sensors from the International Satellite Cloud Climatology Project (ISCCP; Rossow & Schiffer, 1999) data set and the Extended Pathfinder Atmospheres (PATMOS-x; Heidinger et al., 2014) data set, which are the two longest satellite cloud records. ISCCP relies on geostationary radiances that are intercalibrated and gap filled by the Advanced Very High Resolution Radiometer series, while PATMOS-x is derived exclusively from Advanced Very High Resolution Radiometer data, so they are not truly independent. Norris et al. (2016) concluded that the altitude of the highest cloud tops increased between the 1980s and the 2000s, which they argued cannot be mainly explained by El Niño–Southern Oscillation (ENSO) variability. As such, while there is some observational support for the upward shift of clouds predicted by models, it remains unclear how much of the increasing trend in this record is due to climate warming and how much is due to multiyear natural variability. Furthermore, cloud heights based on passive sensors are indirect measurements with many uncertainties and are sensitive to instrument calibration drifts, which can introduce spurious trends in retrieved cloud top altitude (Chepfer et al., 2014). This is particularly problematic for the ISCCP and PATMOS-x data sets, which have been stitched together from a number of individual sensors that were not designed for climate monitoring. While the authors of studies such as Norris et al. (2016) worked extensively to remove spurious trends from their analysis, additional observational support would substantially bolster confidence in this result. Examination of stereo (geometrically derived) cloud top heights from the Multiangle Imaging Spectro Radiometer (launched 1998) should be insensitive to calibration and show large variations due to ENSO but no significant climate trend over 2000–2015 (Davies et al., 2017). However, the approach used in that analysis is weighted by cloud-free scenes and may conflate changes in cloud amount with cloud height (Davies & Molloy, 2012) and needed corrections for changes in Sun glint angle during the first 2 years of the mission as the Terra satellite's ascending equator crossing time was shifted from 10:45 to 10:30 a.m. Cloud heights can change with the diurnal cycle, so care must be taken when comparing absolute values from Multiangle Imaging Spectro Radiometer with those from instruments that sample at different times.

The most recent Multiangle Imaging Spectro Radiometer data also appear to show no trend in the altitude of tropical or midlatitude high cloud that is distinguishable from ENSO (Figure S1 in the supporting information). However, we stress that this data record spans less than 20 years and may simply be too short to detect any climatic trend.

Active spaceborne sensors (radar and lidar) provide range-resolved (equivalently altitude resolved) measurements and therefore a much more direct determination of cloud heights than can be obtained by passive sensors. CloudSat (Stephens et al., 2008) and the Cloud-Aerosol Lidar and Infrared Path Finder Satellite Observation (CALIPSO; Winker et al., 2010) operated for a decade as members of the A-Train constellation (Stephens et al., 2002) and currently operate several kilometers underneath the A-Train. CALIPSO and CloudSat carry the Cloud-Aerosol Lidar with Orthogonal Polarization and Cloud Profiling Radar (CPR), respectively. Cloud-Aerosol Lidar with Orthogonal Polarization is most sensitive to small cloud droplets and ice crystals, and CPR is sensitive to both cloud size and precipitation size particles. A previous study (Chepfer et al., 2014) used a climate model with a CALIPSO simulator to examine the capabilities of detecting changes in clouds due to climate warming. They found an unambiguous upward shift in lidar-detected cloud altitude in simulations when global sea surface temperatures are increased by +4 °C. By assuming that the change is linear in global temperature and that geographically uniform warming has the same effect as realistic warming patterns, they estimated an upward shift in cloud heights of ~20 m/year in the tropics for a strong CO_2 forcing. This suggests a lidar record of several decades, longer than the CALIPSO lidar record that currently exists, will be necessary to detect cloud changes.

Chepfer et al. (2014) encouraged us to investigate how the radar record begun by CloudSat could be used to detect upward cloud shifts, since lidar and radar have different sensitivities to different parts of clouds and can help increase understanding of cloud feedbacks from different perspectives. CALIPSO and CloudSat both stand on their last legs. A new satellite, EarthCARE (Illingworth et al., 2015) will have both lidar and W-band cloud radar sensors and continue the record begun with CALIPSO and CloudSat. EarthCARE's planned launch is in 2021, but with a short lifetime of 3 to 4 years due to fuel limitations. Beyond EarthCARE, the spaceborne lidar and cloud radar data record is uncertain. However, the recent Earth Science decadal survey (National Academies of Sciences, Engineering, and Medicine, 2018) recommended targeted observables of Aerosols and Clouds, Convection and Precipitation with candidate instrument approaches including both lidar and radar sensors that could potentially continue the active cloud sounding data record. If an Aerosols and Clouds, Convection and Precipitation mission came to fruition it could conceivably extend the cloud radar data record well into the 2030s. Is this sufficient to detect the expected upward shifts in cloud heights?

To this end, we investigate three questions in this study: (1) What is the first year at which changes in the altitude of high cloud can be confidently estimated if we continue to fly W-band cloud radar; (2) what radar sensitivity, range resolution, and calibration stability are required to detect those changes; and (3) at what latitude will we first detect these changes? To answer these questions, we use the Community Earth System Model version 1 (CESM1; Hurrell et al., 2013) coupled to the Cloud Feedback Model Intercomparison Project Observation Simulator Package (COSP; Bodas-Salcedo et al., 2011). A description of the COSP implementation in CESM1 and analysis of the cloud biases in CESM1 within a simulator framework can be found in Kay et al. (2012). The data sets from the CESM1 COSP model simulations and CloudSat observations are described in the next section. Section 3 explains our definition of high clouds, the Monte Carlo approach we use, and the methodologies for how to find a year where a significant trend can be detected with consideration of potential error sources. Analysis results and interpretations are presented in section 4. The paper concludes in section 5 with a discussion and summary of key findings.

2. Data

In general, it is difficult to compare the results based on climate model simulations and satellite observations because of the difference between scales (model grids are ~100 km and CloudSat footprints are ~1 km) as well as inherent differences between model variables and observed quantities. To address this issue, satellite simulators that convert model variables into pseudosatellite observations have been developed in an attempt to provide an apples-to-apples comparison between model simulations and satellite observations. Here we use the COSP Quickbeam radar simulator (Haynes, 2007). Our observables are summarized for each month and location in Contoured Frequency by Altitude Diagrams (CFADs). These are 2-D histograms showing the fraction of radar footprints that would have sufficient signal to trigger identification of a cloud, split into bins depending on altitude and reflectivity.

The CESM1 CFADs are from Quickbeam, which simulates W-band (94 GHz) reflectivities equivalent to those observed by CloudSat. It is important to emphasize that there are uncertainties associated with COSP, most importantly with regard to the issue of spatial scales. To address scale mismatch, the COSP package makes assumptions regarding the sub–grid-scale variability of model quantities to generate subcolumns, which are then used as input to the satellite simulators (Hillman et al., 2018). This is an imperfect approach, but as Hillman et al. (2018) show the primary impact of the subgrid scheme is to broaden the reflectivity distribution and this has little impact on the vertical distribution. The reflectivity simulations also depend on the model microphysical representation. It is possible, for example, that stronger updrafts in a future climate will alter particle sizes in a way that will affect reflectivity and is not well represented by models. This is certainly of considerable concern as regard the amount of anvil (Hartmann, 2016), but we expect these changes will be of minor importance with regard to cloud top metrics.

2.1. CESM1 CFAD

In order to simulate a warmer future world, we use the Representative Concentration Pathway (RCP) 8.5, a scenario of comparatively high greenhouse gas emissions in which projected radiative forcing reaches 8.5 W/m^2 at the end of the century. In this study, 90 years (2006 to 2095) of COSP radar monthly gridded CFADs in CESM1 RCP8.5 is used to evaluate trends in the altitude of high clouds as a function of latitude and radar sensitivity.

COSP radar monthly gridded CFAD has a vertical grid with 480-m spacing, a horizontal 2° × 2° latitude-longitude grid, and reflectivity bins that are 5 dBZ wide in the range of −50 to 25 dBZ. The CFAD is sampled at all times of day at all grid points. Examples of CESM1 COSP CFADs over 0–10°N for 2086–2095, 2006–2015, and their difference are shown in Figure 1. Our illustrative bin selection is arbitrary, but we chose a tropical bin because Chepfer et al. (2014) suggested the strongest cloud height response in the tropics. Our main results are shown for all latitude bins. The solid black lines and the dotted black lines represent the altitudes of the mean tropopause layer (TL; averaged monthly tropopauses pressure, CESM1 output TROP_P, from 2006–2095 over 0–10°N) and the height that is 50% below TL (TL50), respectively.

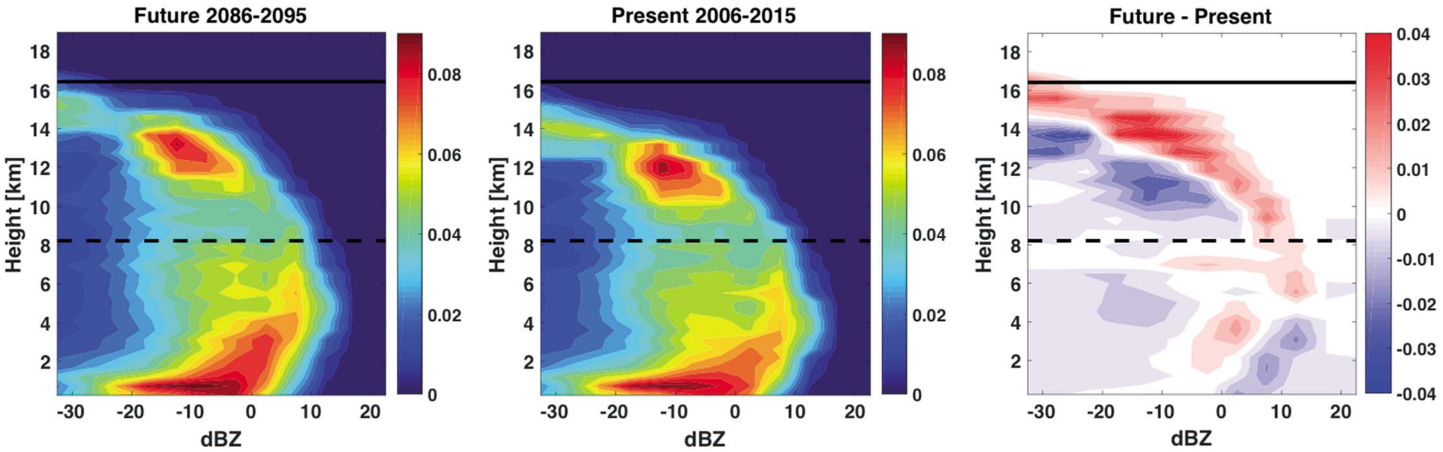

Figure 1. Community Earth System Model version 1 Contoured Frequency by Altitude Diagrams for the future climate (2088–2095 mean), for the present climate (2006–2015 mean), and the difference between them over 0–10°N. Black solid lines and dotted lines are the altitude of the tropopause layer and the height 50% below the tropopause, respectively.

Henceforth, we define high cloud as those above TL50, which is the height that is halfway between the surface and the tropopause. Figure 1 demonstrates that high clouds are higher in the future climate than the present-day climate. The upward shift in reflectivity is clearer for reflectivities <0 than >0 dBZ because higher reflectivities (>0 dBZ) are typically associated with precipitation which complicates the interpretation due to a variety of effects including non-Rayleigh scattering and attenuation of the radar by rising liquid layers.

2.2. CloudSat CFAD

Unlike the model CFAD, the CloudSat data are associated with specific spatiotemporal sampling characteristics. CloudSat measures at nadir in a Sun-synchronous orbit with an equator crossing time near 1:30 a.m./p.m. As is common with satellite data, there are uneven sampling artifacts in the CloudSat data. Most significantly, in April 2011, CloudSat experienced a battery anomaly and had to exit the A-Train constellation. CloudSat spent roughly 6 months drifting away from the A-Train while a new operations plan was being implemented. In November 2011, CloudSat switched to “Daylight Only Operations” mode to keep collecting cloud radar data during sunlit portions of the orbit (Nayak, 2012), and CloudSat successfully returned to the A-Train constellation in May 2012. Therefore, for approximately a 1-year period from April 2011 to May 2012 CloudSat was either out of operation or sampling an irregular part of the diurnal cycle.

We use CloudSat 2B-GEOPROF (Marchand et al., 2008) version R05 from 2006 to 2015 to generate monthly CFADs on a 2° × 2° latitude-longitude grid with a 480-m vertical resolution and 5-dBZ reflectivity bins to be consistent with the COSP CFADs. CloudSat 2B-GEOPROF provides a cloud mask and radar reflectivity, and we choose cloud mask ≥ 20 and radar reflectivity ≥ -30 dBZ as criteria to detect hydrometers. We exclude April 2011 to April 2012 when CloudSat was either suffering a battery anomaly or out of the A-Train and since the remaining data of 2011 (January to March) and 2012 (May to December) are roughly a year's data, we combine these into a single 2011–2012 entry. Since nighttime (a.m.) data are only available until 2010, we only use CloudSat daytime (p.m.) CFADs unless otherwise mentioned.

3. Methodologies

3.1. Definition of High Clouds and WCH

We first identify high cloud fraction as all hydrometeors above TL50, that is, 50% of the 2006–2095 average tropopause height for each latitude. We then generate a weighted cloud height (WCH) for each reflectivity bin:

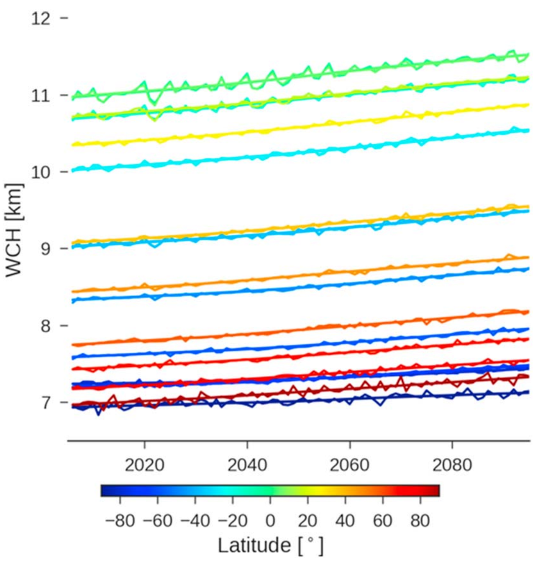

Figure 2. Representative Concentration Pathway 8.5 time series of WCH for the −20-dBZ accumulated bin with line colors representing each latitude. The solid lines through each is a 50-year locally weighted scatter plot smoothing fit that is used as the forced component, $\mathrm{WCH_F}$, for generating our Monte Carlo time series. WCH = weighted cloud height.

$$WCH(\mathrm{dBZ}) = \frac{\sum_{i,TL50}^{i,TOA} H_i \times CFAD(H_i, dBZ)}{\sum_{i,TL50}^{i,TOA} CFAD(H_i, dBZ)} \qquad (1)$$

Here H_i is the height of the ith vertical bin and $CFAD\ (H_i,$ dBZ$)$ refers to the CFAD fraction evaluated at H_i and the selected dBZ bin. The sums are over the bins from TL50 to the top of atmosphere (TOA), and we estimate WCH at each latitude and reflectivity bin.

Instead of using the original COSP reflectivity bins, we aggregate all reflectivities above a given dBZ. For example, hereafter the −35-dBZ bin means the sum of all CFADs from −35 to 25 dBZ. These accumulated reflectivity bins are less sensitive to calibration uncertainty than individual bins because, for example, a ±0.5-dBZ calibration shift is equivalent to 10% of a 5-dBZ bin but is just 1.1 % of the example −20-dBZ accumulated bin we use for illustration, which spans −20 to 25 dBZ. These accumulated bins can be constructed for any instrument with a minimum detectable signal greater than the labeled bin. It is equivalent to a total detected hydrometeor fraction with reflectivity above the bin reflectivity value, and through equation (1), we expect it to be related to the weighted center of mass for high level clouds. Also note that using the average WCH (a continuous variable) allows us to detect changes smaller than the intrinsic resolution of CloudSat's CPR (480 m).

3.2. Monte Carlo Detection of Trends With Different Calibration Error and Internal Variability

3.2.1. Generation of Monte Carlo WCH Series

Only one CESM1 simulation is available for each scenario, and detection year statistics may depend greatly on the realization of the internal variability in this run. In order to more robustly assess the detection of cloud height trends, we use a Monte Carlo approach and generate sets of 1,000 pseudorandom series of WCH. We consider that the radar-derived cloud height within each latitude, ϕ, and dBZ bin consists of the sum of three independent terms:

$$WCH(t,\phi,dBZ) = WCH_F(t,\phi,dBZ) + \delta WCH_{\mathrm{cal}}(t,\phi,dBZ) + \delta WCH_{\mathrm{var}}(t,\phi,dBZ) \qquad (2)$$

where WCH_F is the forced climate component, $\delta WCH_{\mathrm{cal}}$ is variation introduced by changes in radar calibration, and $\delta WCH_{\mathrm{var}}$ is internal climate variability that is generated from a white noise process whose standard deviation is derived from the piControl simulation for the given dBZ latitude bin. We wish to address the question "when will spaceborne cloud radar detect upward shifts in cloud height?," and to build our pseudorandom series for the Monte Carlo analysis, we require estimates of the statistics required to build each of the terms in equation (2), that is, forced, calibration and internal variability. We use CESM1 to provide WCH_F and statistics of $\delta WCH_{\mathrm{var}}$ and the calibration variations described in section 3.2.4 for $\delta WCH_{\mathrm{cal}}$.

3.2.2. Forced Component of WCH, WCH_F

For every latitude-dBZ bin we take the RCP8.5 time series of WCH and use locally weighted scatter plot smoothing (LOWESS) with a 50-year window, local linear regression, and standard tricube weighting (Cleveland & Devlin, 1988). This accounts for nonlinearities in the WCH series, and the LOWESS output is taken as the WCH_F component. These nonlinearities can be seen in Figure 2, which displays the RCP8.5 WCH time series along with the 50-year LOWESS smooth split by latitude for a single dBZ bin.

3.2.3. Internal Variability Component of WCH, WCH_{var}

For every latitude-dBZ bin we take the piControl WCH time series and calculate its standard deviation, then use this to generate internal WCH_{var} from a Gaussian white noise process (Kolmogorov-Smirnov test does not reject normality at $p < 0.05$; see Figure S2). We also consider the effect of possible autocorrelation in the time series and find no evidence that it substantively changes our results, so our main results use

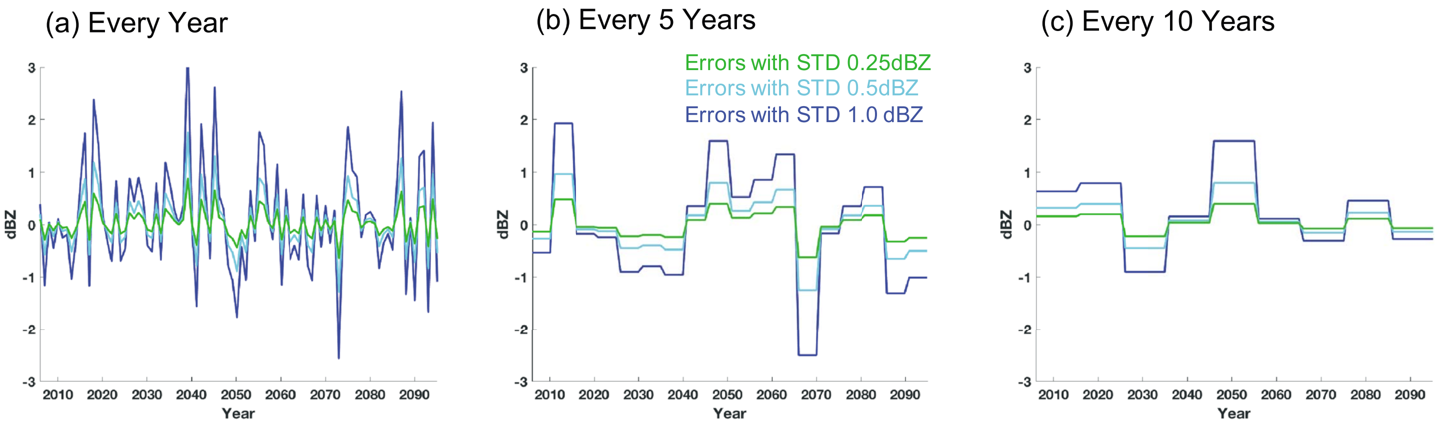

Figure 3. One realization of time series of random errors generated by (a) every year, (b) every 5 years, and (c) every 10 years with STD of 0.25 (green lines), 0.5 (cyan lines), and 1.0 dBZ (blue lines). STD = standard deviation.

Gaussian white noise throughout. A fuller description with detailed illustrations is in Figures S3–S7 and the associated supplementary text.

3.2.4. Calibration Component of WCH, WCH_{cal}

The quality of the reflectivity time series is contingent on the accuracy of calibration. In general, spaceborne radar is relatively well calibrated using the well-characterized reflection properties of the ocean in clear-sky scenes (Li et al., 2005; Tanelli et al., 2008). However, the calibration will vary in time due to drifts in radar output power and thermal effects. Additionally, we can consider shifts in the data record related to changing sensor characteristics over time as calibration uncertainty with a multiyear timescale. As a specific example, EarthCARE will have a slightly smaller footprint than CloudSat. Because the radar reflectivity is a nonlinear function of cloud properties, the sample distribution of reflectivity from the two sensors will be slightly different.

To investigate the effect of calibration errors on trend detection, we generate 1,000 realizations of random normal calibration errors (E) for nine different combinations of calibration error magnitude (standard deviation 0.25, 0.5, and 1.0 dBZ) and three values for the timescale of the error (1, 5, and 10 years). Figure 3 shows an example realization for each of the nine scenarios. In an approximate sense, the yearly change case represents intramission calibration errors, and the 5 and 10 years are examples of potential effects of intermission calibration errors. A calibration shift will cause the “hydrometeor” occurrence fractions in the CFAD to shift right or left, so for each year we estimate the calibration-adjusted CFAD by shifting the entire CFADs to the right when the calibration error is positive and to the left when the error is negative and then estimate from this the change in fractional occurrence assuming a uniform distribution of reflectivity values in each bin (see Figure S8 for a schematic and equations).

After shifting the 90 years of CFADs for each calibration error, we generate 1,000 $WCH_{\mathrm{cal,RCP}}(t)$ series for each latitude and radar sensitivity bin then obtain a calibration only component of WCH via

$$\delta WCH_{\mathrm{cal}}(t,\phi,dBZ) = WCH_{\mathrm{cal,RCP}}(t,\phi,dBZ) - WCH_{\mathrm{RCP}}(t,\phi,dBZ) \quad (3)$$

This results in 1,000 sets of $\delta WCH_{\mathrm{cal}}$ for each of the nine combinations of calibration error magnitude (0.25, 0.50, and 1.00 dBZ) and timescale (1, 5, and 10 years), plus a case of zero-calibration error (see Figure S9 for an example). These are then input into equation (2) as the final component of the Monte Carlo WCH series.

3.3. Trend Detection, Significance, and Stability

Detection of a significant trend at the 95% confidence level occurs when the WCH trend magnitude is larger than approximately twice the standard error of the regression estimate (i.e., trend > 2σ). For our uncorrelated random (“white”) noise the standard error is (e.g., Santer et al., 2000)

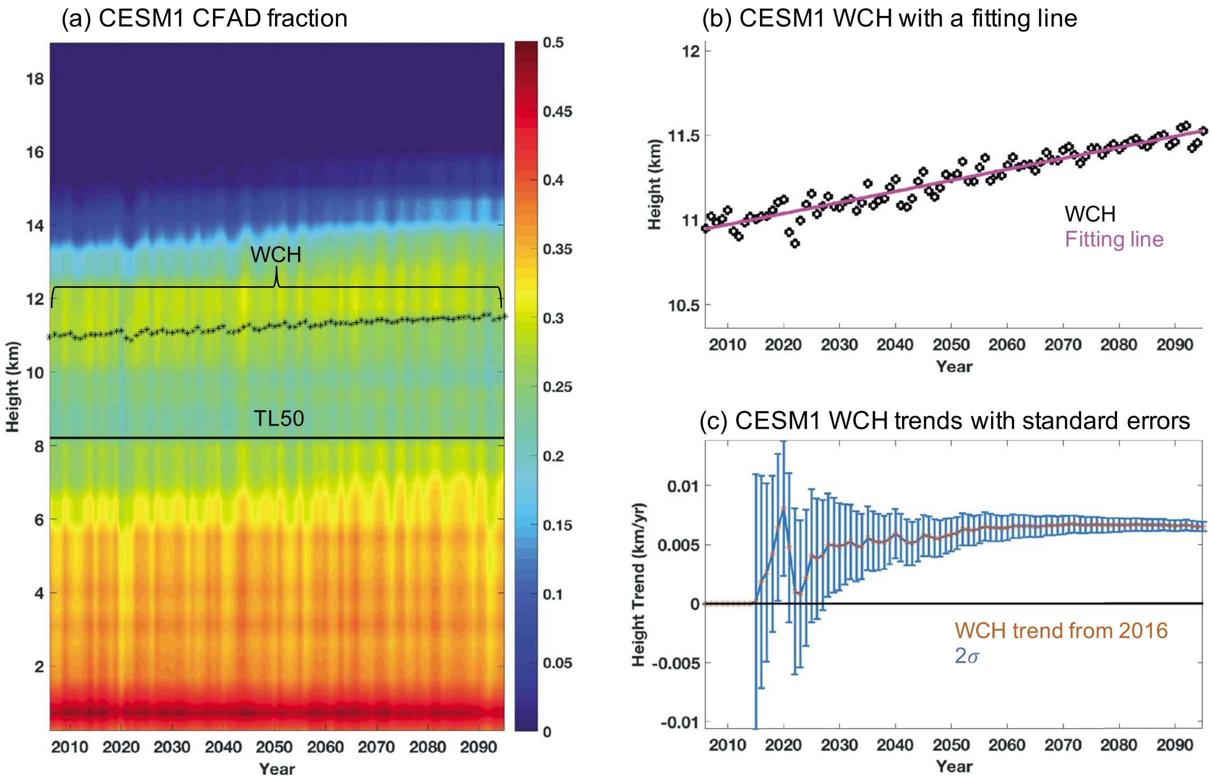

Figure 4. A time series of CESM1 (a) CFAD fraction, together with WCHs and TL50, (b) WCH with a fitting line, and (c) WCH trends with their standard errors (2σ). Data correspond to −20-dBZ accumulated reflectivity bin over 0–10°N. CESM1 = Community Earth System Model version 1; CFAD = Contoured Frequency by Altitude Diagram; WCH = weighted cloud height; TL50 = height 50% below the tropopause.

$$\sigma_{WN} = \sqrt{\frac{\sum(x(t)-x^*(t))^2}{(N-2)\left(\sum(t-\bar{t})^2\right)}} \tag{4}$$

where $x(t)$ is the WCH for year t (e.g., Figure 4a), $x^*(t)$ is the predicted WCH from the fit for year t (e.g., Figure 4b), N is the length of the time series, and $\bar{t}$ is the sample mean of the t values used in the regression.

In order to address when we could be confident in detecting a change in CESM1 WCH under RCP8.5, we calculate trends and their uncertainty beginning with 2006–2015 and then repeatedly lengthen the data series by 1 year until we reach 2006–2095. This result is a set of trend and σ values for each end year from 2015 through 2095, and our reported detection year is when the trend is greater than 2σ and all subsequent trends are also greater than 2σ. We refer to this as significant and stable, or "climatologically significant." This increases the confidence of having detected a climate trend, because even in the absence of a trend we would expect 5% of samples to show significant trends at 2σ. However, these would be temporary and linked to aspects of internal variability such as the ENSO that affect high clouds. This is illustrated in Figure 4c where CESM1 WCH trends $>2\sigma$ always occur after 2028 over 0–10°N with a minimum detectable signal of −20 dBZ. We therefore report 2028 as the detection year for a climatologically significant trend, even though significance at 2σ is first seen in 2019. This trend by 2019 is not maintained and we attribute it to internal variability based on Figure 5, where Figure 5a shows the area-weighted global mean temperature difference relative to 2006 (red line) and the WCH (blue line) time series, and Figure 5b shows the residual values of each after their first 50 years are detrended. The temperature residuals (red line) increase over 2006–2019, and the strong correlation with WCH residuals (blue line) means that natural variability around the longer-term trend inflates the WCH trend for this period. For this bin (i.e., −20-dBZ accumulated reflectivity bin over 0–10°N), the cloud height response derived from regressing the residual WCH against ΔT (400 ± 62 m/°K) is indistinguishable from the values calculated over the full piControl series (423 ± 64 m/°K). Since the piControl series only contains internal variability, this supports our argument

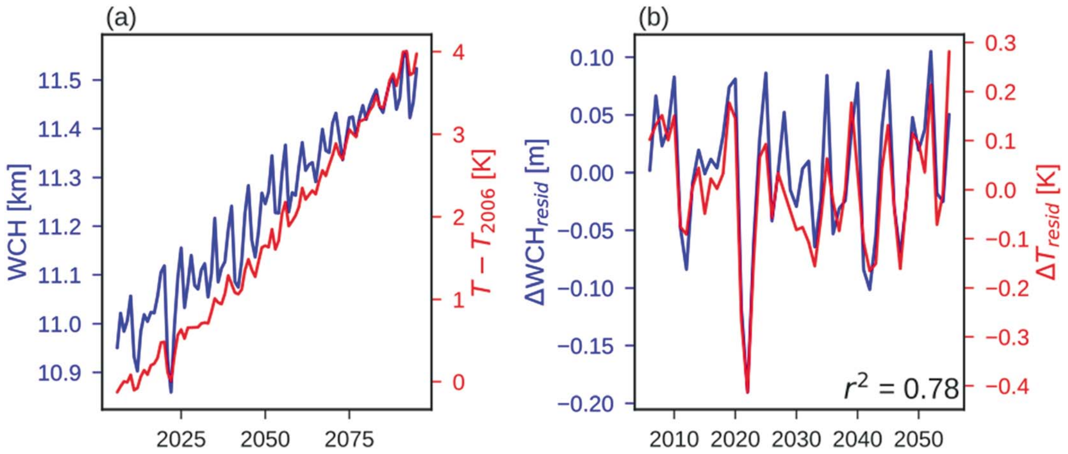

Figure 5. (a) Time series (2006 to 2095) of Community Earth System Model version 1 WCHs (blue, left axis) and the global mean temperature difference relative to 2006 ($T - T_{2006}$; red, right axis), and (b) time series (2006 to 2055) of residuals for the first 50 years following subtraction of a linear trend fitted to WCH (blue, left axis) and $T - T_{2006}$ (red, right axis) of the first 50-year series. The reported r^2 is between the two series. Data correspond to −20-dBZ accumulated reflectivity bin over 0–10°N. WCH = weighted cloud height.

that the temporary short-term trend significance is partly due to internal variability (further evidence is in Figures S10–S12). Our "climatologically significant" requirement of maintained significance means that these likely El Niño-driven cases are not mistakenly identified as a climate trend detection.

We wish to address the question "given a year, would we be 95% confident of seeing a significant and stable trend in WCH?," rather than "given a significant trend at a particular year, will it be stable?" The answer to this question must address uncertainty sources including natural variability and instrumental uncertainty. Our Monte Carlo approach described in section 3.2 allows us to test many possible realizations of internal variability instead of limiting our results to the single CESM1 simulation.

3.4. Model Representativeness

Unlike in the model, real-world observations are limited in space and time, which may contribute to model-observation discrepancies in the mean state or variability. In addition, CESM1's mean state and variability may differ from reality. We evaluate this by comparing CESM1 and CloudSat over 2007–2015 in Figure 6. Figure 6a shows the 2007–2015 WCH time series in CESM1 and CloudSat for a particular latitude and reflectivity bin. Note that the model mean WCH is ~200 m higher than the CloudSat estimate, and the trend differs in sign between the two data sets. The difference in trend signs may be related to internal variability: the CESM1 run shows a negative 2006–2015 trend in the NINO 3.4 region compared with the positive real-world trend, indicating opposite effects on WCH from a major component of internal variability (see Figure S13 and related discussion). Regarding the mean state discrepancy, we expect that most of this is a fundamental model bias but some may be because we use monthly averaged model output and compare with near-1:30 p.m. CloudSat sampling. Our only diurnal information is the additional 1:30 a.m. data from CloudSat for 2007–2010, so we separately calculate WCH including am data for comparison. Figure 7 compares the time series of CloudSat WCH during daytime only (WCH$_{pm}$) and during day and night (WCH$_{all}$) for one latitude and radar sensitivity. We reiterate the importance of only using the p.m. (sunlit) data from the observational record because of the loss of a.m. data after 2011. The WCH$_{all}$ series uses combined day-night data only for 2007–2010 before switching to day-only data. First, 2007–2010 shows that WCH$_{pm}$ and WCH$_{all}$ vary in the same way. Thus, the CloudSat data shows that (on annual scales) the variation in WCH are due more to large-scale changes than to any changes in the diurnal cycle. If this remains true in the future (which seems to us likely), then trends between constant-time observations (meaning observation made at the same time of day) and time-integrated model output can be reliably compared without time sampling the model output in the same way for the purpose of trend analysis. The discontinuity introduced by the loss of nighttime data in 2011 shifts the data, resulting in an artificial trend change and illustrating why observations must either ensure consistent time of day sampling, or otherwise correct for any changes to diurnal sampling.

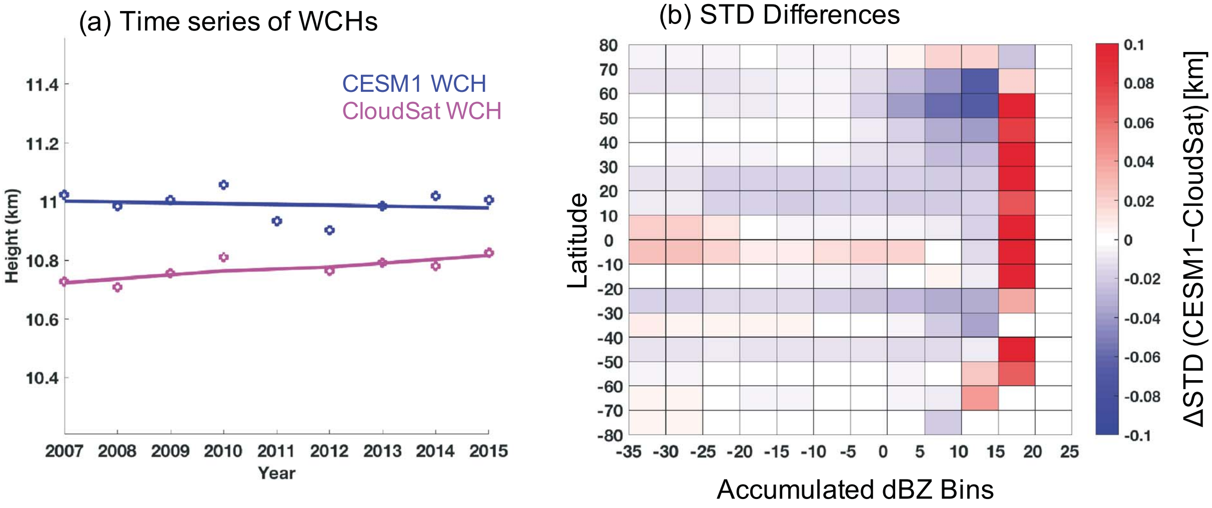

Figure 6. (a) Time series (2007 to 2015) of CESM1 WCHs (blue) and CloudSat WCHs (magenta) corresponds to the bin labeled −20 dBZ, that is, accumulated CFAD for all reflectivities greater than −20 dBZ, and for the latitude bin 0–10°N, and (b) the difference in modeled and observed STD of WCHs (CESM1 − CloudSat) over different latitudes and accumulated dBZ bin. CESM1 = Community Earth System Model version 1; WCH = weighted cloud height; STD = standard deviation.

Regardless of the time and space sampling and the offset in mean WCH, the standard deviation (STD) is similar: 0.047 km in CESM1 and 0.040 km in the CloudSat record. Figure 6b shows that the difference in modeled and observed STDs of WCHs is small in magnitude at all latitudes and radar sensitivities. To test for statistical significance, we bootstrapped the piControl run and found little evidence for detectable differences between the CloudSat and CESM1 9-year standard deviations (see Figure S14 and associated text). Our analysis proceeds using the CESM1 variability, with the caveat that the results are only valid for CESM1 and that conclusions could therefore be sensitive to biases in the simulated trends and internal variability.

We have already reported that within CESM1, WCH varies with global temperature, so we also convert our detection years into the amount of global warming that occurs by detection. We fit a 50-year LOWESS smooth to the global temperature series and subtract the 2006 fit temperature from the detection year temperature to give an amount of post-2006 warming associated with WCH trend detection. Table 1 shows forced warming levels at the beginning of each decade and these can be applied, for example, to Figure 4c where the detection year (2028) corresponds to 0.81 °C warming since 2006.

For context, the Berkeley Earth algorithm (Rohde et al., 2013) combined with the HadSST3 ocean record (Kennedy et al., 2011) is the method, which best approximates global temperature change (Richardson et al., 2018). This data record shows approximately 1.0 °C warming from 1850–2006. Given that we estimate 0.81 °C warming from 2006 is required to give a detection, our estimated detection corresponds to 1.8 °C warming from 1850, which is a commonly used baseline for estimates of total global warming.

We intend this to allow approximate inferences for detection times based on faster or slower estimates of global warming. We highlight that the response of WCH(T) requires future investigation and may affect the transferability of our results, and we note that the response of WCH for our example dBZ-latitude bin to forced temperature changes appears to be smaller than that for internal variability (Figure S11) and that while the forced WCH response corresponds to global T everywhere, for short-term variability only tropical WCH anomalies show strong correlation with global temperature (Figure S12).

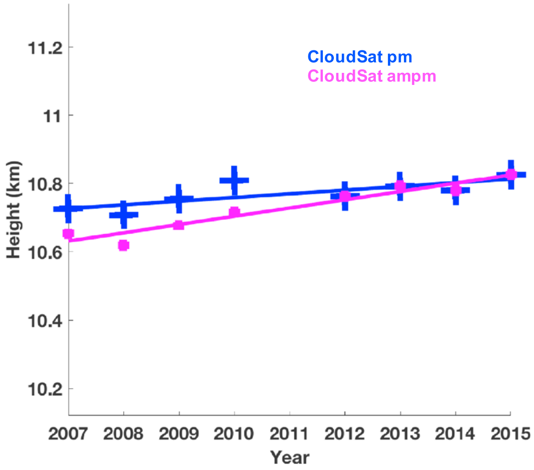

Figure 7. Time series (2007 to 2015) of CloudSat weighted cloud heights during daytime only (blue) and during day and night (magenta) corresponds to −20 accumulated dBZ bins over 0–10°N. Note that CloudSat only had night data until February 2011, so the final four magenta points are daytime only and are included to illustrate the effect of discontinuities on trends of this length.

Table 1
Forced Warming Estimates in This CESM1 Run Relative to 2006

Year	Forced warming since 2006 (K)
2006	0.00
2010	0.15
2020	0.51
2030	0.89
2040	1.30
2050	1.78

Note. Values are provided at each decade; annual locally weighted scatter plot smoothing fits are used in the main calculation.

3.5. Summary of Methodology

Given the complexity of our method, we provide a summary here:

1. Take the monthly CESM1 COSP output CFADs and for each dBZ bin *i*, replace its value with the sum of that bin and all higher-dBZ bins to produce accumulated-dBZ bin CFADs.
2. For each month and latitude calculate the WCH using equation (1) applied to the accumulated-dBZ CFADs, then convert these series to annual averages.
3. Estimate the effect of calibration uncertainty on WCH by shifting the CFADs randomly as described in section 3.2.4 (Figures S8 and S9) and generate WCH series including calibration error.
4. For each dBZ-latitude bin generate 1,000 Monte Carlo samples by summing forced, calibration and internal variability components as described in section 3.2.
5. For each of these 1,000 random samples, calculate the trend and its white noise uncertainty using equation (4) and report the detection year as the first year where the trend is $>2\sigma$ and all subsequent years also meet this criterion.
6. Rank the 1,000 samples and report the 950th value as the year by which we are 95% confident of detecting a climatological trend.
7. Convert this to a forced temperature change from 2006 by fitting a 50-year window LOWESS to the CESM1 area-weighted global temperature series and reporting the difference between the fit temperature in this year and 2006 (Figure S10).
8. Report these years and temperature changes for all dBZ-latitude bins.

We take the 0–10°N, −20 dBZ bin as an example to illustrate typical output of our Monte Carlo approach. The histogram of the resultant detection years for $\delta WCH_{\mathrm{cal}} = 0$ are shown in Figure 8a. The first bin in Figure 8a is 2016 and represents all the detections that would be achieved by looking at the data from 2006 to 2016. It thus includes some detections that would be found if we had used even a shorter period (e.g., just 9 years of data up to 2015). On the other hand, the other bins in Figure 8a are the number of additional cases found in each additional year. For example, 2017 means cases found when using data from 2006 to 2017 but *not* found when using data from 2006–2016. Notice the wide variation including almost 200 out of 1,000 detections by 2016 that maintain their statistical significance, implying that significant trends could already be detectable in the current CloudSat record.

In general, these early detections (i.e., 2016, so 10 or fewer years) represent Monte Carlo cases where internal variability over 2006–2016 showed a positive trend, strengthening the total WCH trend (i.e., internal + true $> 2\sigma$). While some of these early detections are cases in which the absence of a true trend would have generated a temporary or false detection (i.e., internal only $>2\sigma$), this will occur only 5% of the time. In fact, we find that about 4.5% of the 2016 cases show an early detection that is not stable (and thus

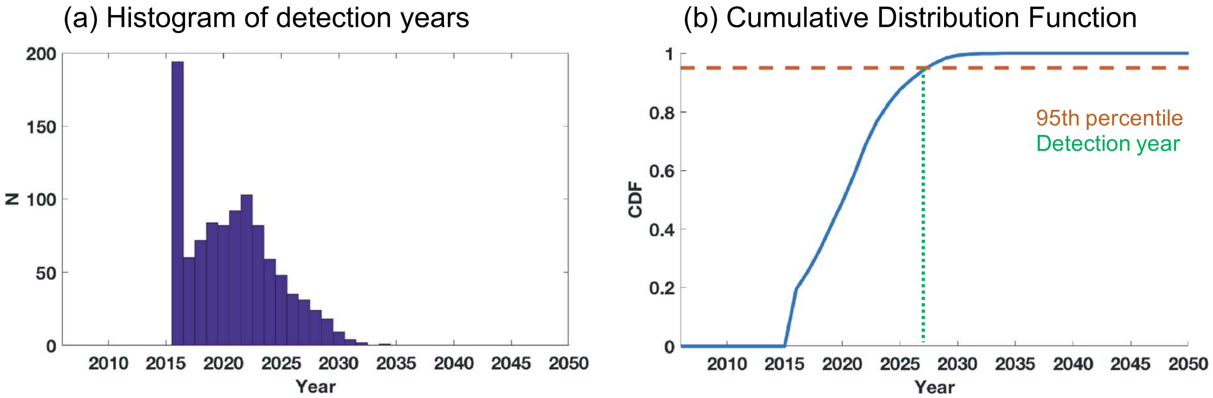

Figure 8. (a) A histogram of detection years based on 1, 000 randomly generated time series and (b) their cumulative distribution function, together with a dotted line of 95th percentile (red) and that of detection year (green), which corresponds to −20 accumulated dBZ bins over 0–10°N.

not identified as a climatologically robust trend in our statistics in Figure 8). In contrast, 20% of the 2016 cases are early detections that are found to be robust, which shows that the bulk of the early detections are cases where the true trend contributed significantly to the detection. Nonetheless, this also highlights the fundamental importance of continuing measurements after some initial detection to ensure the detection is climatologically robust, and more generally, the importance of collecting data that allow for detailed studies on the origins of any detected trend, to ensure we understand the physical changes leading to the observed trend and nominally to help quantify the component of the observed change that is due to internal variability. We report the 95th percentile of this (Figure 8b) and for the WCH_F and $\delta WCH_{\mathrm{var}}$ estimates here, there is a 95% chance of detecting a significant WCH trend in 2028 that will then be maintained. All results that follow will use these criteria to determine the detection year of *significant* and *stable* trends, which we will refer to as confident detection years.

4. Results

Current Observed Trends

Here we briefly show the trends and statistical significance of the current decade-long data record. The 2007–2015 trends, white noise standard errors, and significance levels from CloudSat and CESM1 are shown in Figure 9. Significant trends occur in both, but in different locations. Significant trends are seen in the tropics for CloudSat and the middle latitudes for the CESM1 simulation. However, due to a $p < 0.05$ requirement, on average, one of the 18 latitude bins would be significant at any point in time even if there is no true trend. Therefore, it is not surprising to see some trends based on such a short time period. In particular, Figure 5 shows a relationship between WCH and unforced temperature variability, and the positive real-world Nino3.4 trend over 2007–2015 favors a positive unforced real-world WCH trend (Figure S13). This may explain the spread of significant WCH trends across the tropics in the CloudSat record. Figure 9 indicates that there is spatial coherence between adjacent latitudes. We investigated the cross-correlation between latitudes (Figure S7) and found that it is only strong for the tropics.

Predicted Detection Years From CESM1

Figure 10 shows the expected detection years in the absence of calibration error (Figure 10a), together with the global forced temperature change relative to 2006 (Figure 10b) for each dBZ-latitude bin's detection year. For radar sensitivities between −35 and 5 dBZ, the year of significance tends to be in the mid-2020s or early 2030s, except over the tropics and polar regions. Over the tropics, radar sensitivity between −20 and −10 dBZ is needed for a confident detection in the 2030s.

When using CloudSat-based internal variability in place of CESM, tropical detection can advance to the late 2020s for some dBZ < −15 bins, and the southern midlatitude detection is delayed by up to 5 years. However, the finding of the earliest detection in the northern midlatitudes is robust (Figures S15 and S16).

Figure 11 shows that our radar CFAD-based approach indicates larger trends in the tropics under RCP8.5, but internal variability is also larger, which delays trend detection. The largest source of internal variability in global temperature is ENSO, and the strongest response of WCH to short-term global temperature variability is in the tropics, so we expect this tropical variation to be largely related to ENSO. The tropical trends also tend to be larger in the −15-dBZ accumulated bin than the lower dBZ bins. This implies that the vertical profile of cloud mass change is nonuniform. If this −15-dBZ bin represents vertical shifts in the within-cloud mass structure, cloud radar could be obtaining information which is independent of lidar, which tends to be attenuated below approximately three cloud optical depths.

Chepfer et al. (2014) reported that lidar would detect the greatest shifts in cloud altitude in the tropics for a fixed global warming of 4 °C. We suspect that our findings are not inconsistent with that result for several reasons, including how tropical lidar-detected cloud tops and radar-derived WCHs may differ in their forced response and variability. The Chepfer et al. method is also not directly comparable to ours in many ways. For example, their simulations use a spatial ΔT pattern based on quadrupling atmospheric CO_2 over 140 years. In most CMIP5 models, the warming pattern evolves (Andrews et al., 2015; Stephens et al., 2016**)** with tropical warming generally being delayed. It is likely that the 140-year-long pattern of warming in Chepfer et al. includes more tropical warming and therefore a stronger local thermodynamic effect on clouds than we would expect from the 50–60 years from the initialization of warming in the 1970s (Cahill et al., 2015) to detection years from 2025–2035.

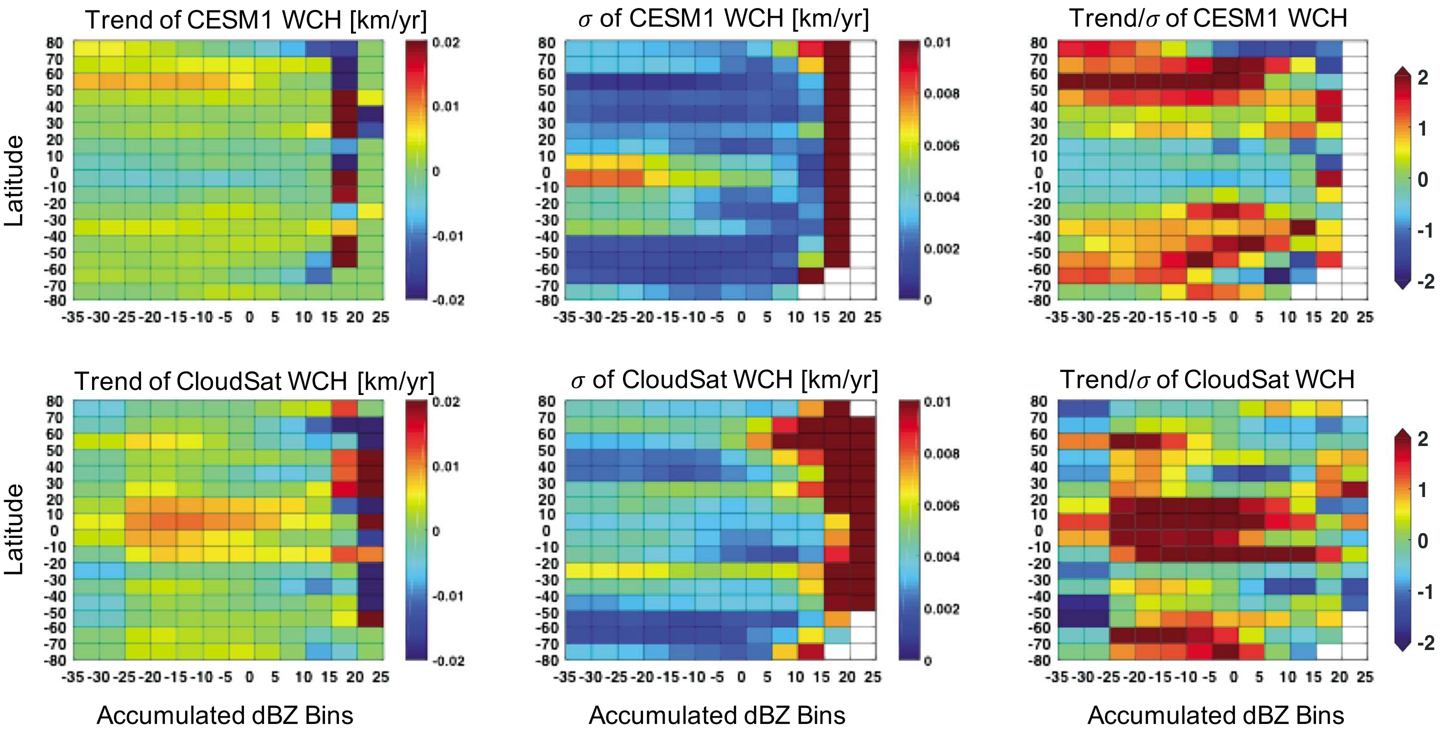

Figure 9. Trends of WCHs (left column), standard errors (σ) of WCHs (middle column), and the ratio between them (right column) for CESM1 (top row) and CloudSat (bottom row) over different dBZ bins (*x* axis) and latitudes (*y* axis) for 2007 to 2015. In the right-hand panels the magnitude of the shading corresponds to the statistical significance of the trend (1 ~ 66%, 2 ~ 95%). WCH = weighted cloud height; CESM1 = Community Earth System Model version 1.

Overall, these results indicate that high sensitivity (i.e., −30 dBZ or better) is not fundamentally needed to detect trends in cloud heights prior to 2030, given the trends and variability simulated by CESM1. Radars that are sensitive to nonprecipitating cloud hydrometeors (<0 dBZ), associated with convective outflow and dense cirrus clouds rather than precipitation hydrometeors (>0 dBZ) are best able to capture the expected trends. We also find that autocorrelation in the WCH_{var} time series has little effect on the detection year, with a 95% confidence interval of a 1-year earlier to 2-year later detection, justifying our use of a simplified white noise structure (Figure S17).

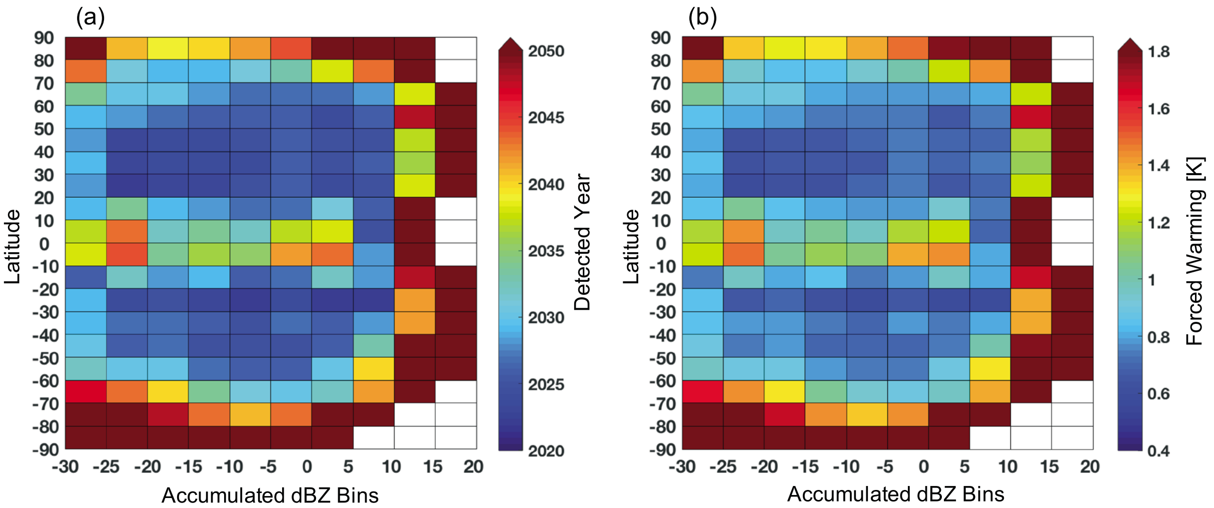

Figure 10. (a) Summary of years that significant trends can be detected and (b) forced warming (right) relative to 2006 over different dBZ bins (*x* axis) and latitudes (*y* axis) based on Community Earth System Model version 1 Contoured Frequency by Altitude Diagram.

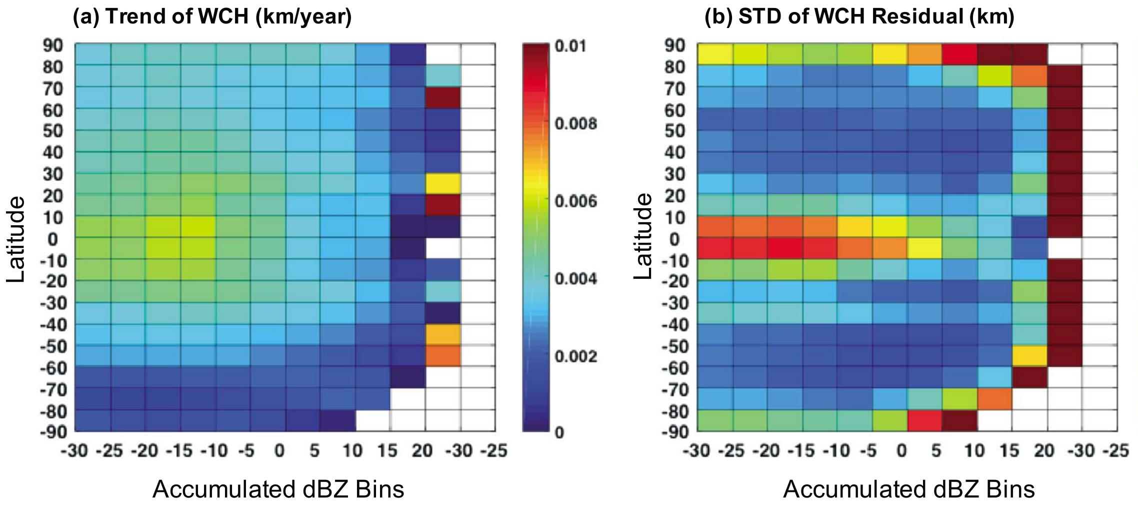

Figure 11. (a) Trend of WCH and (b) STD of WCH residual over different dBZ bins (x axis) and latitudes (y axis) for 2006 to 2050. WCH = weighted cloud height; STD = standard deviation.

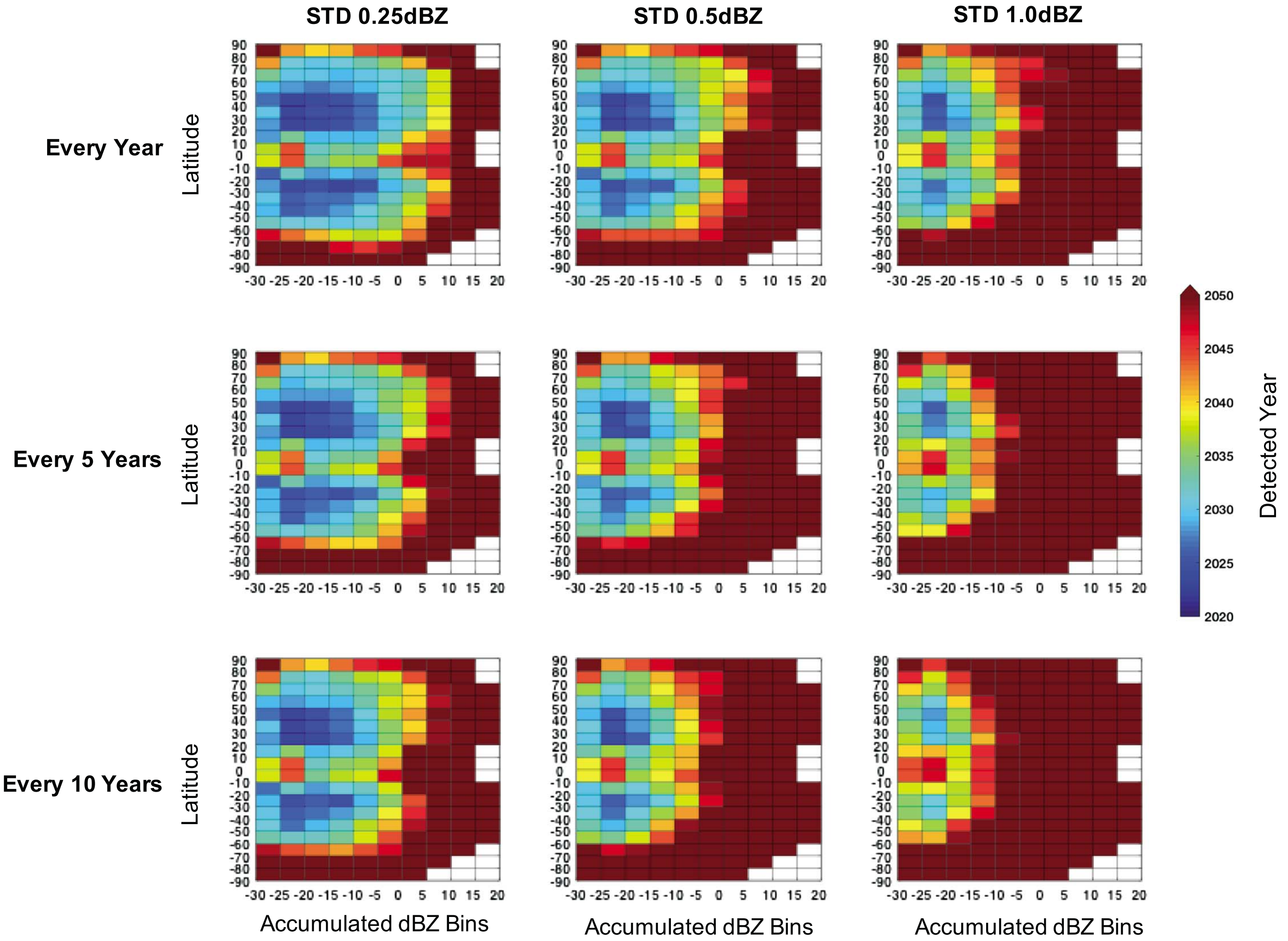

Figure 12. Similar to Figure 9 but with different calibration errors based on 1,000 different simulations. STD = standard deviation.

Influence of Instrument Characteristics and Calibration Uncertainty

CloudSat has a true vertical bin size of 480 m, and we find that our results are not sensitive to degradation in vertical resolution by a factor of 2 times or 3 times (Figure S18). This means that even if a measurement has only ~1.5-km vertical resolution, we can still detect the trend approximately as early as a measurement with ~0.5-km vertical resolution.

We now present the various hypothetical calibration errors described in section 3.2.4, based on the 95th percentile of the stable detection year in our Monte Carlo simulations. Figure 12 shows the resultant detection year estimates for a white noise $\delta WCH_{\mathrm{var}}$ (as in Figure 10a) with each combination of the calibration error magnitude (0.25, 0.50, and 1.00 dBZ) and timescale (1, 5, and 10 years) as illustrated in Figure 3. Overall, the trend detection year is more sensitive to the magnitude of the calibration accuracy than the timescale of calibration uncertainty: Longer timescales cause slightly later detection, while larger calibration uncertainties cause considerably later detection, especially for the larger dBZ bins (>-10 dBZ). For example, all radar sensitivity values between −30 and 0 dBZ are able to detect extratropical changes in WCH by the late 2030s with a calibration uncertainty of 0.25 dBZ, but a calibration uncertainty of 1-dBZ delays the detection by several years for smaller dBZs to more than a decade for larger dBZs.

Given that ±1 dBZ is a conservative estimate for calibration uncertainty, these results generally suggest that pushing the radar sensitivity is not the limiting factor in detecting a significant trend. Instead, the calibration stability is a more critical factor in determining the year of detection.

5. Summary and Discussion

This study investigates the first year at which changes in altitude of high cloud can be confidently detected if we continue to fly W-band cloud radar in space, based on analysis of an RCP8.5 CESM1 simulation. We also identify what radar sensitivity is required to detect such changes before 2050 and at what latitude we first expect such a detection. We evaluate trends in the altitude of the weighted height of high clouds in the CESM1 CFADs as a function of latitude and radar sensitivity.

The primary detection year results we present here are based on analysis of CESM1 simulations, and care should be taken when relating these to real-world applications. For example, we provide estimates based on available RCP8.5 output, a simulation of strong global warming. The largest differences relative to other scenarios occurs later in the 21st century (Rogelj et al., 2012), and we focus on pre-2050 detection of cloud height changes, but if the real-world experiences a different climate trend through the 2040s due to, for example, volcanism or differences in transient climate response relative to the model, then the expected detection year would shift. Also, note that this study has used a single model, which happens to have a comparatively high equilibrium climate sensitivity (Meehl et al., 2013). If we make the assumption that the magnitude of the cloud deepening scales with the climate sensitivity, then this would imply that our results might be skewed toward earlier detection than would be found with a lower sensitivity model. We also therefore provide estimates of the forced warming that have occurred by each decade in Table 1, allowing others to infer how detection year might change under faster or slower warming. For example, our bins showing detection by 2030 correspond to 0.9 °C global warming from 2006, although the exact relationship between year and global warming is dependent on the internal variability magnitude and on the trend length, since longer time periods reduce the trend uncertainty regardless of its magnitude.

Another limitation of the current study is that we do not have sufficient observational data to provide a reliable constraint on the magnitude of internal variability or identify which noise process is most representative. The short CloudSat time series is not detectably different from CESM1 output, but even quite substantial differences would likely not yet be detectable due to the small sample size. A bootstrap-based analysis showed few latitude-dBZ bins where the CloudSat variability magnitude was outside the 95% confidence interval from CESM1 (Figure S14), and a test using the CloudSat-based variability showed changes of up to ±5 years in individual bin detection years, but did not affect the conclusion of earliest detection in the northern mid-latitudes (Figures S15 and S16). Based on analysis of CESM1, we identified regions, particularly in the tropics, where standard white noise processes are likely not applicable. However, our Monte Carlo tests indicate that our calculated detection years on average do not change when considering an autocorrelated noise process. We therefore believe that our analysis is robust to the most probable range of internal variability.

The key findings are as follows:

1. In the absence of calibration error, pre-2050 detection occurs at some latitudes for all accumulated sensitivity bins below 0 dBZ, that is, for cloud, rather than precipitation radars. Detection is delayed for lower sensitivity (higher dBZ) and rarely occurs before 2050 for >0 dBZ.
2. If missions do not share CloudSat's orbit and overpass time, then careful accounting of differences in mean WCH due to the diurnal cycle will be needed to remove discontinuities between satellites. For 0–10°N, the mean offset between 1:30 p.m. and combined 1:30 a.m./p.m. WCH is ~83 m, equivalent to 13 years of the climate trend. Note that the EarthCARE mission will have a 30-min delayed equatorial crossing time compared to CloudSat and careful analysis will be necessary to combine these data records in a manner that accounts for diurnal sampling shifts.
3. Sampling with ~1.5-km vertical resolution can be as useful as sampling with ~0.5-km vertical resolution.
4. Although the largest shifts occur over the tropics, the greater magnitude of internal variability (both in CloudSat and CESM) there means that detection is most likely to occur sooner for the midlatitudes.
5. Reduced radar calibration error generally has a bigger effect on detection year than lowering the minimum detectable reflectivity.

Results suggest that a longer cloud radar record may be able to detect the expected upward shifts in cloud heights over 20–60°S/N before 2030 for a radar with a sensitivity of -15 dBZ and stable calibration errors of ± 0.25 dBZ, under CESM1-like global warming. For a calibration error of ± 1 dBZ, detection in these areas either requires a sensitivity closer to -25 dBZ or the extension of the record by approximately a decade and into the 2030s.

The sensitivity and range resolution requirements we have identified could realistically be achieved by low mass, volume, and power solid-state radar on a small satellite bus at significantly lower cost than CloudSat or EarthCARE radars. An example of this approach is the recent successful demonstration of the RainCube, Radar in a Cubesat (Peral et al., 2015) Ka-band radar. Development of a similar miniaturized W-band radar is ongoing. The results presented here provide systems-level requirements on sensitivity, calibration, and range resolution that can help guide this development. These exciting developments in the radar and spacecraft technology open the door for low cost monitoring of upward shifts in cloudiness that are one of the more robust features of climate simulations.

Acknowledgments
The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (NASA). The data of A-Train observations can be found from the CloudSat Data Processing Center (www.cloudsat.cira.colostate.edu). We are also grateful to three anonymous reviewers for their valuable comments.

References

Andrews, T., Gregory, J. M., & Webb, M. J. (2015). The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models. *Journal of Climate*, *28*(4), 1630–1648.

Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J. L., Klein, S. A., et al. (2011). COSP: Satellite simulation software for model assessment. *Bulletin of the American Meteorological Society*, *92*(8), 1023–1043. https://doi.org/10.1175/2011BAMS2856.1

Bony, S., Webb, M., Stevens, B., Bretherton, C., Klein, S., & Tselioudis, G. (2009). The cloud feedback model intercomparison project: Summary of activities and recommendations for advancing assessments of cloud-climate feedbacks. CFMIP Doc.

Cahill, N., Rahmstorf, S., & Parnell, A. C. (2015). Change points of global temperature. *Environmental Research Letters*, *10*(8), 084002.

Chepfer, H., Noel, V., Winker, D., & Chiriaco, M. (2014). Where and when will we observe cloud changes due to climate warming? *Geophysical Research Letters*, *41*, 8387–8395. https://doi.org/10.1002/2014GL061792

Cleveland, W. S., & Devlin, S. J. (1988). Locally weighted regression: An approach to regression analysis by local fitting. *Journal of the American Statistical Association*, *83*(403), 596–610.

Davies, R., Jovanovic, V. M., & Moroney, C. M. (2017). Cloud heights measured by MISR from 2000 to 2015. *Journal of Geophysical Research: Atmospheres*, *122*, 3975–3986. https://doi.org/10.1002/2017JD026456

Davies, R., & Molloy, M. (2012). Global cloud height fluctuations measured by MISR on Terra from 2000 to 2010. *Geophysical Research Letters*, *39*, L03701. https://doi.org/10.1029/2011GL050506

Hansen, J., Lacis, A., Rind, D., Russell, G., Stone, P., Fung, I., et al. (1984). Climate sensitivity: Analysis of feedback mechanisms. *Climate Processes and Climate Sensitivity*, *29*, 130–163. https://doi.org/10.1029/GM029p0130

Hartmann, D. L. (2016). Tropical anvil clouds and climate sensitivity. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(32), 8897–8899.

Haynes, J. M. (2007). QuickBeam radar simulation software user's guide v 1.1 a. Department of Atmospheric Science, Colorado State University.

Heidinger, A. K., Foster, M. J., Walther, A., & Zhao, X. (2014). The pathfinder atmospheres–extended AVHRR climate dataset. *Bulletin of the American Meteorological Society*, *95*(6), 909–922.

Hillman, B. R., Marchand, R. T., & Ackerman, T. P. (2018). Sensitivities of simulated satellite views of clouds to subgrid-scale overlap and condensate heterogeneity. *Journal of Geophysical Research: Atmospheres*, *123*, 7506–7529. https://doi.org/10.1029/2017JD027680

Hurrell, J., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P., et al. (2013). The Community Earth System Model: A framework for collaborative research. *Bulletin of the American Meteorological Society*. https://doi.org/10.1175/BAMS-D-12-00121.1

Illingworth, A. J., Barker, H. W., Beljaars, A., Ceccaldi, M., Chepfer, H., Clerbaux, N., et al. (2015). The EarthCARE satellite: The next step forward in global measurements of clouds, aerosols, precipitation, and radiation. *Bulletin of the American Meteorological Society*, *96*(8), 1311–1332. https://doi.org/10.1175/BAMS-D-12-00227.1

Kay, J. E., Hillman, B. R., Klein, S. A., Zhang, Y., Medeiros, B., Pincus, R., et al. (2012). Exposing global cloud biases in the Community Atmosphere Model (CAM) using satellite observations and their corresponding instrument simulators. *Journal of Climate*, *25*(15), 5190–5207. https://doi.org/10.1175/JCLI-D-11-00469.1

Kennedy, J. J., Rayner, N. A., Smith, R. O., Parker, D. E., & Saunby, M. (2011). Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 1. Measurement and sampling uncertainties. *Journal of Geophysical Research*, *116*, D14103. https://doi.org/10.1029/2010JD015218

Kuang, Z., & Hartmann, D. L. (2007). Testing the fixed anvil temperature hypothesis in a cloud-resolving model. *Journal of Climate*, *20*(10), 2051–2057.

Li, L., Heymsfield, G. M., Tian, L., & Racette, P. E. (2005). Measurements of ocean surface backscattering using an airborne 94-GHz cloud radar—Implication for calibration of airborne and spaceborne W-band radars. *Journal of Atmospheric and Oceanic Technology*, *22*(7), 1033–1045. https://doi.org/10.1175/JTECH1722.1

Marchand, R., Mace, G. G., Ackerman, T., & Stephens, G. (2008). Hydrometeor detection using CloudSat—An Earth-orbiting 94-GHz cloud radar. *Journal of Atmospheric and Oceanic Technology*, *25*(4), 519–533.

Meehl, G. A., Washington, W. M., Arblaster, J. M., Hu, A., Teng, H., Kay, J. E., et al. (2013). Climate change projections in CESM1 (CAM5) compared to CCSM4. *Journal of Climate*, *26*(17), 6287–6308. https://doi.org/10.1175/JCLI-D-12-00572.1

Nayak, M. (2012). CloudSat anomaly recovery and operational lessons learned. In SpaceOps 2012 (p. 1295798).

National Academies of Sciences, Engineering, and Medicine (2018). *Thriving on Our Changing Planet: A Decadal Strategy for Earth Observation from Space*. Washington, DC: The National Academies Press. https://doi.org/10.17226/24938.

Norris, J. R., Allen, R. J., Evan, A. T., Zelinka, M. D., O'dell, C. W., & Klein, S. A. (2016). Evidence for climate change in the satellite cloud record. *Nature*, *536*(7614), 72–75.

Peral, E., Tanelli, S., Haddad, Z., Sy, O., Stephens, G., & Im, E. (2015). Raincube: A proposed constellation of precipitation profiling radars in CubeSat. In *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 1261–1264). IEEE. https://doi.org/10.1109/GARSS.2015.7326003

Richardson, M., Cowtan, K., & Millar, R. J. (2018). Global temperature definition affects achievement of long-term climate goals. *Environmental Research Letters*, *13*(5), 054004.

Rogelj, J., Meinshausen, M., & Knutti, R. (2012). Global warming under old and new scenarios using IPCC climate sensitivity range estimates. *Nature Climate Change*, *2*(4), 248.

Rohde, R., Muller, R., Jacobsen, R., Perlmutter, S., Rosenfeld, A., Wurtele, J., et al. (2013). Berkeley Earth temperature averaging process. *Geoinfor Geostat: An Overview*, *13*, 20–100.

Rossow, W. B., & Schiffer, R. A. (1999). Advances in understanding clouds from ISCCP. *Bulletin of the American Meteorological Society*, *80*(11), 2261–2288.

Santer, B. D., Wigley, T. M. L., Boyle, J. S., Gaffen, D. J., Hnilo, J. J., Nychka, D., et al. (2000). Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. *Journal of Geophysical Research*, *105*(D6), 7337–7356. https://doi.org/10.1029/1999JD901105

Soden, B. J., & Held, I. M. (2006). An assessment of climate feedbacks in coupled ocean–atmosphere models. *Journal of Climate*, *19*(14), 3354–3360.

Stephens, G. L., Kahn, B. H., & Richardson, M. (2016). The super greenhouse effect in a changing climate. *Journal of Climate*, *29*(15), 5469–5482.

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., et al., & the CloudSat Science Team (2002). The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation. *Bulletin of the American Meteorological Society*, *83*(12), 1771–1790. https://doi.org/10.1175/BAMS-83-12-1771

Stephens, G. L., Vane, D. G., Tanelli, S., Im, E., Durden, S., Rokey, M., et al. (2008). CloudSat mission: Performance and early science after the first year of operation. *Journal of Geophysical Research*, *113*, D00A18. https://doi.org/10.1029/2008JD009982

Storelvmo, T., Tan, I., & Korolev, A. V. (2015). Cloud phase changes induced by CO_2 warming—A powerful yet poorly constrained cloud-climate feedback. *Current Climate Change Reports*, *1*(4), 288–296.

Tanelli, S., Durden, S. L., Im, E., Pak, K. S., Reinke, D. G., Partain, P., et al. (2008). CloudSat's cloud profiling radar after two years in orbit: Performance, calibration, and processing. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(11), 3560–3573. https://doi.org/10.1109/TGRS.2008.2002030

Wetherald, R. T., & Manabe, S. (1988). Cloud feedback processes in a general circulation model. *Journal of the Atmospheric Sciences*, *45*(8), 1397–1416.

Winker, D. M., Pelon, J., Coakley, J. A. Jr., Ackerman, S. A., Charlson, R. J., Colarco, P. R., et al. (2010). The CALIPSO mission: A global 3D view of aerosols and clouds. *Bulletin of the American Meteorological Society*, *91*(9), 1211–1230. https://doi.org/10.1175/2010BAMS3009.1

Zelinka, M. D., & Hartmann, D. L. (2010). Why is longwave cloud feedback positive? *Journal of Geophysical Research*, *115*, D16117. https://doi.org/10.1029/2010JD013817

Zelinka, M. D., Klein, S. A., & Hartmann, D. L. (2012). Computing and partitioning cloud feedbacks using cloud property histograms. Part II: Attribution to changes in cloud amount, altitude, and optical depth. *Journal of Climate*, *25*(11), 3736–3754.