Article

Study of the Effect of Temporal Sampling Frequency on DSCOVR Observations Using the GEOS-5 Nature Run Results (Part II): Cloud Coverage

Daniel Holdaway [1,2,*] and Yuekui Yang [2,3]

[1] Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[2] Goddard Earth Sciences Technology and Research, Universities Space Research Association, Columbia, MD 21046, USA; Yuekui.Yang-1@nasa.gov
[3] Climate and Radiation Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
* Correspondence: dan.holdaway@nasa.gov; Tel.: +1-301-614-6158

Abstract: This is the second part of a study on how temporal sampling frequency affects satellite retrievals in support of the Deep Space Climate Observatory (DSCOVR) mission. Continuing from Part 1, which looked at Earth's radiation budget, this paper presents the effect of sampling frequency on DSCOVR-derived cloud fraction. The output from NASA's Goddard Earth Observing System version 5 (GEOS-5) Nature Run is used as the "truth". The effect of temporal resolution on potential DSCOVR observations is assessed by subsampling the full Nature Run data. A set of metrics, including uncertainty and absolute error in the subsampled time series, correlation between the original and the subsamples, and Fourier analysis have been used for this study. Results show that, for a given sampling frequency, the uncertainties in the annual mean cloud fraction of the sunlit half of the Earth are larger over land than over ocean. Analysis of correlation coefficients between the subsamples and the original time series demonstrates that even though sampling at certain longer time intervals may not increase the uncertainty in the mean, the subsampled time series is further and further away from the "truth" as the sampling interval becomes larger and larger. Fourier analysis shows that the simulated DSCOVR cloud fraction has underlying periodical features at certain time intervals, such as 8, 12, and 24 h. If the data is subsampled at these frequencies, the uncertainties in the mean cloud fraction are higher. These results provide helpful insights for the DSCOVR temporal sampling strategy.

Keywords: cloud fraction; satellite sampling frequency; DSCOVR; EPIC; time series; GEOS-5; Nature Run; Fourier; spectral analysis

1. Introduction

Orbiting around the Sun–Earth L1 Lagrange point, the two instruments on board the Deep Space Climate Observatory (DSCOVR) spacecraft, the Earth Polychromatic Imaging Camera (EPIC) and the National Institute of Standards and Technology Advanced Radiometer (NISTAR), provide observations on the entire sunlit side of the Earth continuously [1]. Like any other satellite, DSCOVR can only discretely sample the Earth's climate system. Factors such as the dowlink bandwidth and receiver antenna availability also affect the number of observations that can be sent back. As has been shown in previous studies, this subsampling will affect the accuracy of satellite retrieved parameters [2,3]. The potential effects of temporal sampling frequency on the Earth's radiation budget derived from DSCOVR-like observations have been assessed using a range of different metrics in Part 1 of this study ([4]—hereafter, Part-1). Another important product of DSCOVR-EPIC is the cloud fraction.

This paper will address how temporal sampling frequency affects this quantity with a similar approach as Part-1.

EPIC observes the Earth with 10 spectral channels ranging from the ultraviolet to the near infrared with a 2048 by 2048 CCD array. At the vantage point of L1, EPIC provides snapshots of the cloud fields for the entire sunlit half of the Earth, which is valuable in interpreting the NISTAR observations and in climate applications. To study the effect of temporal sampling, a baseline truth dataset is needed. As in Part 1, the Goddard Earth Observing System version-5 (GEOS-5) Nature Run cloud fields are used as "truth". The Nature Run is well-suited for this purpose because of its high spatial and temporal resolution. It is run with a 7 km horizontal resolution, 72 model levels up to 0.01 hPa and with a timestep of a few minutes. The Nature Run provides a two year time series of an atmospheric state that is dynamically consistent. As discussed in Part 1 and the references therein (e.g., [5–7]), the GEOS-5 Nature Run and operational systems have been carefully validated against satellite data and shown to maintain a realistic atmospheric state throughout the length of integration.

A comparison between the current GEOS-5 operational system and the Nature Run mean total cloud fraction for September 2006 is shown in Figure 1. The maximum operational forecast length is 24 h. Although differences exist, as expected, the cloud fraction produced by the Nature Run has similar temporal variations to that seen in GEOS-5 operational products. For ocean cloud fraction, there is a slight positive bias for the Nature Run, but the variations and the land cloud fraction are very similar between the two models. Compared to the operational system, it is clear that the Nature Run field exhibits similar scales and structures. The Nature Run captures smaller-scale variations than the operational model due to the higher spatial and temporal resolution, which is beneficial in terms of testing the limits of temporal sampling frequency. Figure 2 shows a snapshot of the cloud fraction for the sunlit side of the Earth from the Nature Run at 1200 UTC on 25 September 2006. The image shows the field from the perspective of EPIC. For every time step of the Nature Run, a cloud field similar to Figure 2 can be generated. Analysis in this study is based on the mean cloud fraction of these simulated snapshots.

The paper is arranged as follows: Section 2 recaps the methodology outlined in Part-1 and introduces any new methodology used here. Section 3 shows all of the results for different sampling frequencies and discusses the implications for DSCOVR. Section 4 provides the concluding remarks.
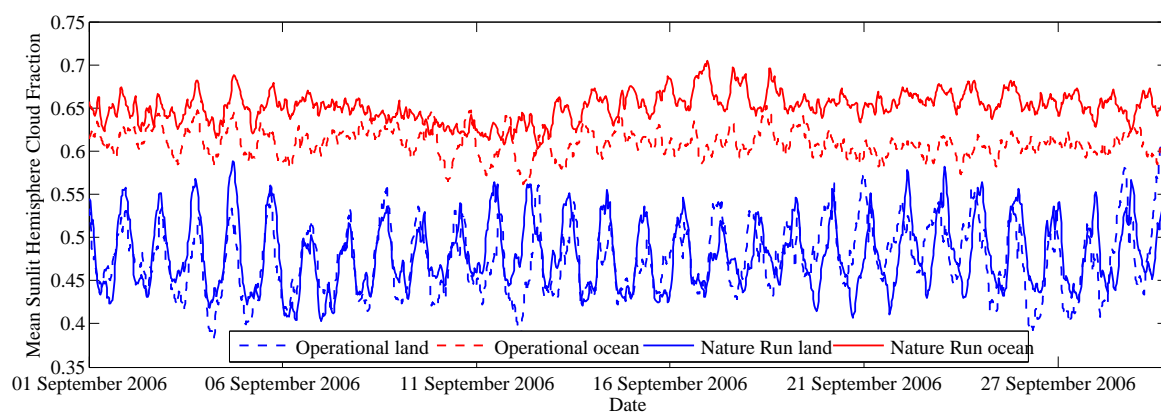

Figure 1. Comparison of land and ocean cloud fraction for the Nature Run and operational versions of GEOS-5. The time series shows the month of September 2006.
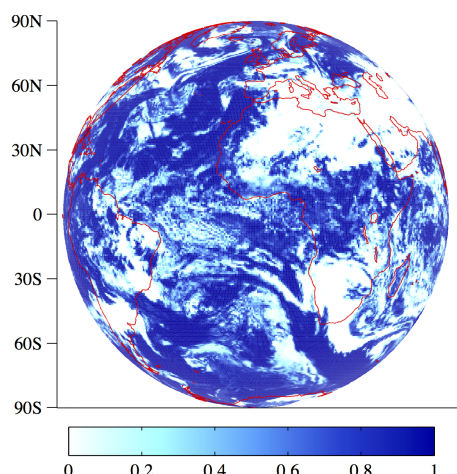
Figure 2. Simulated Earth Polychromatic Imaging Camera (EPIC) view from the cloud field produced by the Nature Run at 1200 UTC on 25 September 2006.

2. Methodology

The cloud fraction is output from the Nature Run every half hour, which is adopted as the "truth". Let t be the "truth" time series of the average cloud fraction for the sunlit hemisphere, then:

$$t = \{t_1, t_2, ..., t_N\} \tag{1}$$

where subscripts denote output times.

Sampling t with different frequencies mimics the observations that would be captured by EPIC. For a particular sampling frequency, a subsample can be constructed by choosing a specific starting point. Since the temporal resolution of the "truth" is half an hour, there are $2n$ potential starting points, where n is sampling frequency in hours; hence, $2n$ time series can be constructed for a sampling frequency of every n hours. Let $r_{n,j}$ be the subsample time series corresponding to the j-th starting point, where $j = 1, ..., 2n$, then:

$$r_{n,j} = \{r_{n,j,1}, r_{n,j,2}, ...\} \tag{2}$$

Some of the metrics used to analyze the subsamples r are similar to those used in Part-1, but are briefly reintroduced below for clarity.

2.1. Uncertainties in the Mean

The first metric is the uncertainty in the mean cloud fraction. To generalize for a specific time interval—be it a day, a month, a season, or a year—the uncertainty of the mean can be measured with the standard deviation of the differences between full and subsampled time series over all possible subsamples. Let $\overline{r_{n,j}}$ be the interval mean for the j-th subsample and $\bar{t}$ be the mean of the original time series (the "truth"). Note $\overline{r_{n,j}}$ and $\bar{t}$ are also time series, but with each point in the series being for each interval. The standard deviation of the difference between full and subsampled time series interval means can be calculated as

$$\sigma_{r,n} = \sqrt{\frac{1}{2n} \sum_{j=1}^{2n} \overline{r_{n,j}}^2 - \bar{t}^2} \tag{3}$$

The average standard deviation of the difference between full and subsampled time series over the entire Nature Run time period provides a measure of the overall uncertainty:

$$\overline{\sigma_{r,n}} = \frac{1}{m} \sum_{k=1}^{m} [\sigma_{r,n}]_k \tag{4}$$

where m is the number of possible intervals. For the Nature Run period, there are 720 days, 24 months, 8 seasons, or 2 years, so $m = 720$ if considering daily means.

2.2. Absolute Error in the Mean

Instead of computing the daily, monthly, seasonal, and annual means, the error compared to the true mean for that interval can be used. For a given sampling frequency of every n hours, the absolute error time series can be written as:

$$\delta_{r,n} = \frac{1}{2n} \sum_{j=1}^{2n} \left| \overline{r_{n,j}} - \bar{t} \right| \tag{5}$$

The average absolute error over the entire Nature Run for a given sampling frequency is,

$$\overline{\delta_{r,n}} = \frac{1}{m} \sum_{k=1}^{m} [\delta_{r,n}]_k \tag{6}$$

where m is the number of possible intervals.

2.3. Correlation

To complement the above metrics and asses the degree of aliasing that occurs for a certain sampling frequency, the correlation between the truth and the subsampled time series is computed. This is achieved by first linearly interpolating the subsample back to the original grid. The correlation coefficient:

$$R_{n,j} = \text{corr}\left(p_{n,j}, t\right) \tag{7}$$

where $p_{n,j}$ is the linearly-interpolated time series, is computed for each possible starting point. As above, a mean value is computed across all starting points within a particular frequency, denoted $\overline{R_n}$. In addition to that, the standard deviation of the correlation coefficients across all starting points is computed, $\sigma_{R,n}$.

2.4. Fourier Analysis

If strong underlying periodical signals are embedded within the time series, it could increase the uncertainty associated with sampling frequencies that are similar to that signal's frequency (wave number). To investigate whether this could be an issue for DSCOVR observations of cloud fraction, a Fourier analysis of the truth time series is conducted. The discrete fast Fourier transform (FFT) quickly converts the original signal t, which is a function of time, into one which is a function of wave number k. The new function, denoted T, is given by

$$T_k = \sum_{l=1}^{N} t_l \exp\left(-2\pi i (l-1)(k-1)/N\right), \text{ where } k = 1, 2, ..., N \tag{8}$$

Examining $|T|^2$ against wave number gives the power spectrum of the time series. Peaks in the spectrum correspond to where a signal with that underlying wave number is prevalent in the time series. Examining the original time series visually it can be difficult to determine whether underlying signals are present.

The time series of cloud fraction over the Nature Run period can contain large-scale trends, especially given the different regions of Earth viewed by DSCOVR throughout the year. These trends will predominantly affect the low frequency signals in the power spectrum, as discussed in [8,9], and should not alter the parts of the spectrum of interest here. However, it is useful to reduce this trend so the time series is periodic, alleviating the Gibbs phenomenon [10]. A discrete step at the end of the period being examined could arbitrarily project onto higher frequency oscillations. The trend is removed by subtraction of a four week running mean of cloud fraction as well as the remaining linear trend and a truncation of the time series to make the start and end points equal. The truncation

typically results in the loss of fewer than ten points. Experiments with different trend removal options showed little impact on the structure of the overall power spectrum.

3. Results

The Nature Run provides cloud fractions at different altitudes: high C_h, middle C_m, low C_l, and total C_t. For each altitude region, high, middle, and low, the maximum column cloud fraction between an upper and lower height is taken. The total cloud fraction is given using a random overlapping assumption $C_t = C_l + C_m + C_h - C_lC_m - C_lC_h - C_mC_h + C_mC_hC_l$. The mean total cloud fraction is computed for the entire sunlit half of the Earth—the same area that EPIC observes. To take into account the differences between clouds over ocean and over land, cloud fraction for the two different surface types are considered separately.

Figure 3a shows the daily mean $\bar{t}$ and mean plus and minus one standard deviation of the difference between full and subsampled time series $\bar{t} \pm \sigma_{r,n}$ for the 720 days of the Nature Run. The figure shows the results for the total cloud fraction C_t over land. The three panels show the results for the $n = 4$-, 8-, and 12-h sampling frequency. Figure 3b shows the monthly mean and standard deviation of the difference between full and subsampled time series; Figure 3c shows the seasonal values; and Figure 3d shows the annual values. Note that vertical scaling is kept fixed within each of the three sub-panels, but is not fixed for the entire figure.
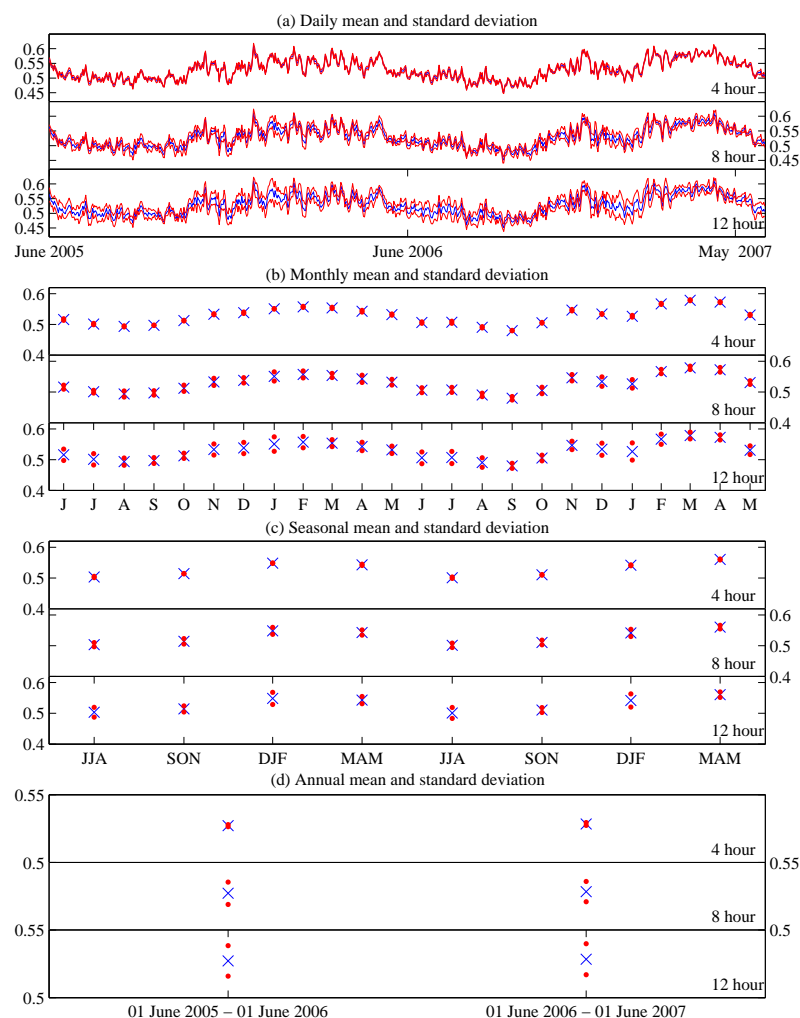

Figure 3. The blue curves/points show $\bar{t}$ for the sunlit land (**a**) daily; (**b**) monthly; (**c**) seasonal; and (**d**) annual intervals. The red curves/points show $\bar{t} \pm \sigma_{r,n}$. Within each panel, the three sub-panels show, from top to bottom, 4-, 8-, and 12-h sampling frequency. Vertical scale is fixed within each set of three panels.

Figure 3 shows that variations in the daily mean during the northern hemisphere winter season are greater than that during the summer months. At this time of year, a greater area of ocean is visible from the L1 Lagrange point as the large land masses of the northern hemisphere are orientated away from the sun. Due to the constantly changing character of cloud fields, the smaller the geographical area being included in the calculation, the larger the variation that is expected. Figure 3 also shows the general trend that the uncertainty increases as the sampling frequency decreases. For the 12-h sampling frequency, the one standard deviation marks are further from the mean. This is confirmed by examining $\delta_{r,n}$ over land, shown in Figure 4. Here it is clear that the absolute errors in the means generally increase as the sampling frequency decreases. It is also clear that the uncertainty increases in the northern hemisphere winter months, where the means were seen increasing. Note that, compared to coarser temporal resolutions, when a sampling frequency of 4-h is used the seasonal cycle is significantly less, demonstrating that using this level of sampling rate helps to ameliorate the uncertainty when a smaller area of land is in view.
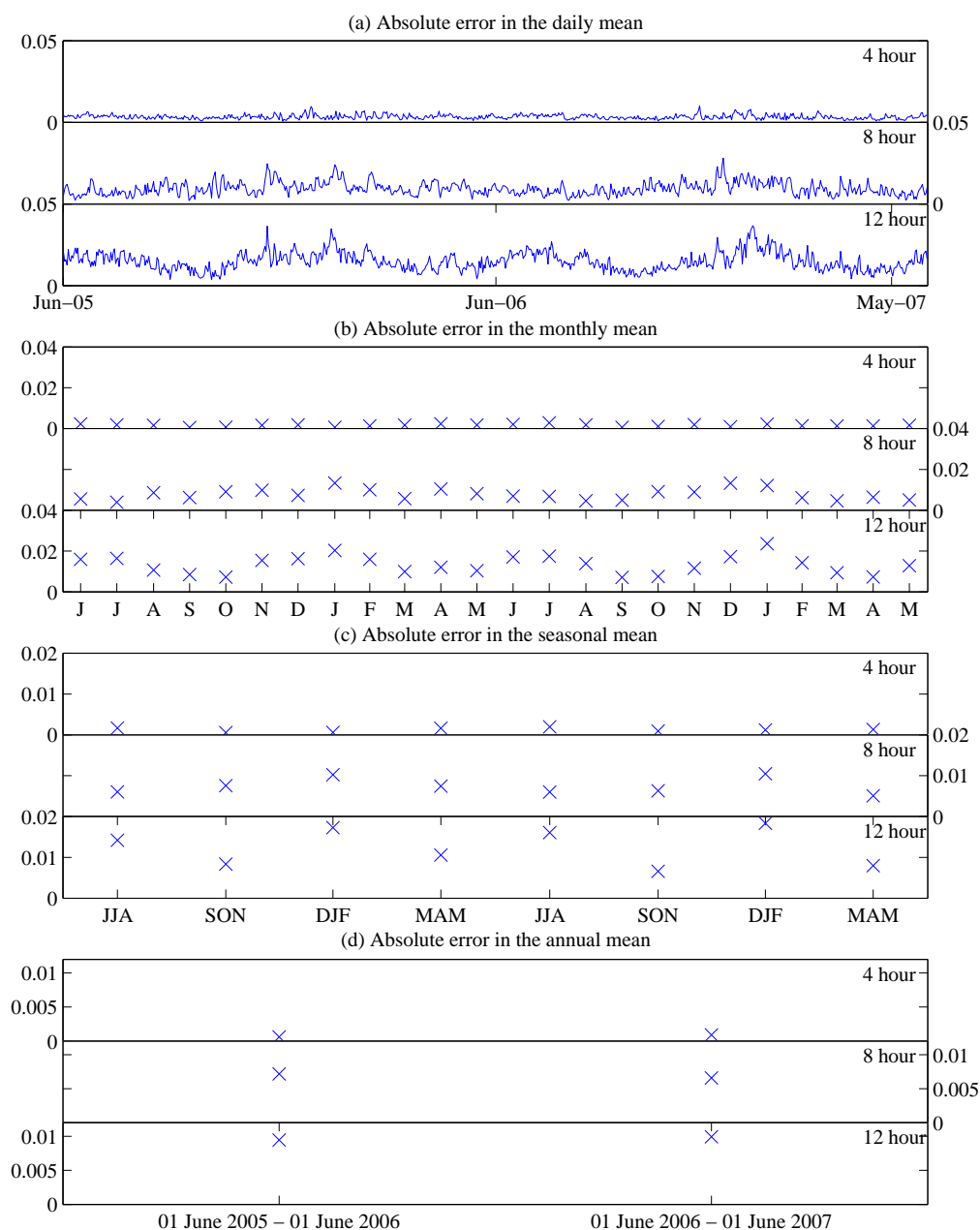

Figure 4. All panels (**a**–**d**) as for Figure 3 but showing just $\delta_{r,n}$.

Generally the behavior seen for cloud fraction in Figures 3 and 4 is similar to that seen for radiation in Part-1. A sampling frequency of 4-h or finer results in a smaller uncertainty in the calculation of the sunlit spatial mean. For 4-h sampling, the uncertainty is fairly consistent throughout the year for all periods—daily, monthly, seasonal, and annual.

The Nature Run mean total cloud fraction over land varies between 50% and 60% for the sunlit half of the Earth (Figure 3). Note that from Figure 4 the uncertainty in the daily mean is about 1% or smaller for 4-h sampling frequency, 2% for 8-h, and 4% for 12-h sampling. For annual mean, the uncertainty associated with 4-h sampling is 0.1%; it is slightly less than 1% for 8-h and 1% for 12-h sampling. These uncertainty levels are significant for climate studies, where cloud remains a major source of uncertainty in climate forcing [11]. Studies ([11] and the rerferences therein) have shown that the net global and annual radiative effect of clouds is about -20 Wm^{-2}. Assuming a global cloud coverage of 60%, one percent of change in cloud fraction is equivalent to about 0.33 Wm^{-2} in radiative forcing, which is comparable to the global aerosol direct radiative forcing (-0.35 Wm^{-2}) [12]. Since long term cloud coverage change can either magnify or reduce the CO$_2$-of great interest. Studies have found that the trend signal is very small. Over the tropical ocean it, is about 1.4% change per decade and over other areas, the signal is even smaller or nonexistent [13]. The uncertainties shown in the subsampling with temporal resolution coarser than 4-h would have significant impact on the cloud coverage trend study. We do note that in this work, uncertainties from other sources such as those in cloud detection (e.g., [14,15]) which can have larger effects are not considered.

Figure 5 is equivalent to Figure 3, but shows the mean and standard deviation for the sunlit ocean. The mean cloud fraction over the ocean is higher than over land, varying between around 0.65 and 0.7. The difference between the means and the standard deviations appears to be smaller for ocean than land, and this is confirmed by examining the errors in the mean in Figure 6. For all sampling rates and intervals, the uncertainty in measuring the mean cloud fraction over ocean is smaller. Similar as for land the uncertainty is more consistent throughout the year for the 4-h sampling than seen for the 8-h and 12-h sampling. The Earth's surface is around two-thirds ocean, so DSCOVR's view will always contain a significant amount of ocean. This is not the case for land, when at certain times of year only small geographical areas of land are in view. For these smaller regions, variations can be larger due to the large variations of cloud at specific locations. Since small geographical areas of ocean are not encountered, larger associated uncertainty is also not encountered.

Another feature in the above results is that the errors are not always smaller for 12-h sampling frequency *versus* 8-h. For the annual means over ocean, the errors are largest for the 8-h sampling frequency. This is different to what has been seen for land cloud fraction and outgoing radiation, where increasing sampling frequency generally resulted in smaller uncertainty. These results are confirmed by examining $\overline{\sigma_{r,n}}$ for land, ocean, and global regions over all periods, shown in Figure 7. Each data point in Figure 7 corresponds to the mean of sets of standard deviations shown in Figure 3 and Figure 5, with additional sampling frequencies also shown. For total cloud fraction over land, up to 4-h sampling frequency the mean uncertainty increases quite gradually—especially for monthly, seasonal, and annual means. For sampling frequency of 4-h or greater the uncertainty tends to increase more rapidly. For the ocean region, the results are similar up to 8-h sampling frequency, in that the increase in uncertainty is more gradual from 2-h to 4-h than it is from 4-h to 8-h. The results are very different from 8-h to 12-h though, here the uncertainty actually decreases for all time intervals.

Upon closer inspection, the increase in uncertainty for ocean cloud fraction between 8-h and 12-h sampling frequencies appears to result from an underlying signal within the time series. The issue is demonstrated by considering a sampling frequency of 24-h against other sampling frequencies, for which the problem is magnified. Figure 8a shows the time series of the mean sunlit ocean total cloud fraction for a two week period beginning 1 August 2006 at 0000 UTC. The figure shows the nature run time series *versus* the time series with a selection of sampling frequencies with the same starting point. Examining this small section of the time series it is clear that a highly diurnal cycle is present in the mean sunlit ocean cloud fraction time series. This is different to the normal atmospheric diurnal cycles in the sense that DSCOVR always views the sunlit hemisphere so it is not related to cloud fraction being

different during the day *versus* at night. Instead this is related to the side of the Earth that DSCOVR happens to be observing. When sampling frequency is below 24-h, all the curves capture the time series variations over this two week period, with varying degrees of aliasing. However, when the frequency is exactly 24-h, only the peaks of the time series are captured and the entire signal is effectively aliased. For this particular starting point, the mean for the two week period is captured very poorly. Indeed, the 26-h sampling frequency does better at capturing the range of possible values than the 24-h. This is because the sampling frequency is different to the period of the underlying signal. The annual means for all starting points and sampling frequencies for the period 1 Jun 2006 at 0000 UTC to 1 June 2007 at 0000 UTC are shown in Figure 8b. It is clear here that the uncertainty is much broader for the 24-h sampling frequency than any others, including higher frequencies. The figure also demonstrates that the uncertainty is not monotonic as the sampling interval increases. The largest uncertainty is seen for 8-, 12-, 16-, and 24-h sampling frequencies. It should be noted that even though sampling at certain longer time intervals may result in a more certain annual mean than sampling at shorter intervals, information is lost as the sampling interval becomes larger and larger, as will be demonstrated later in the discussion.
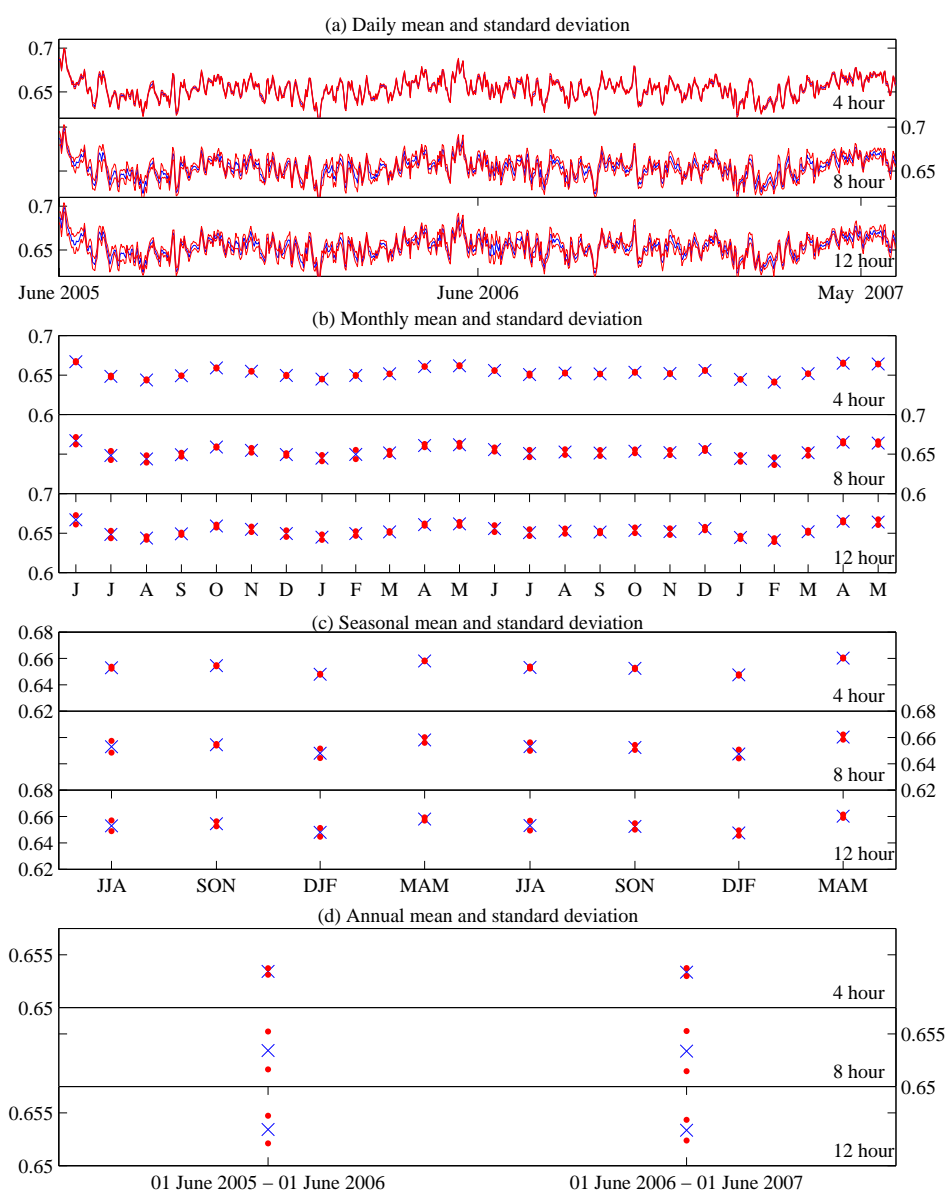

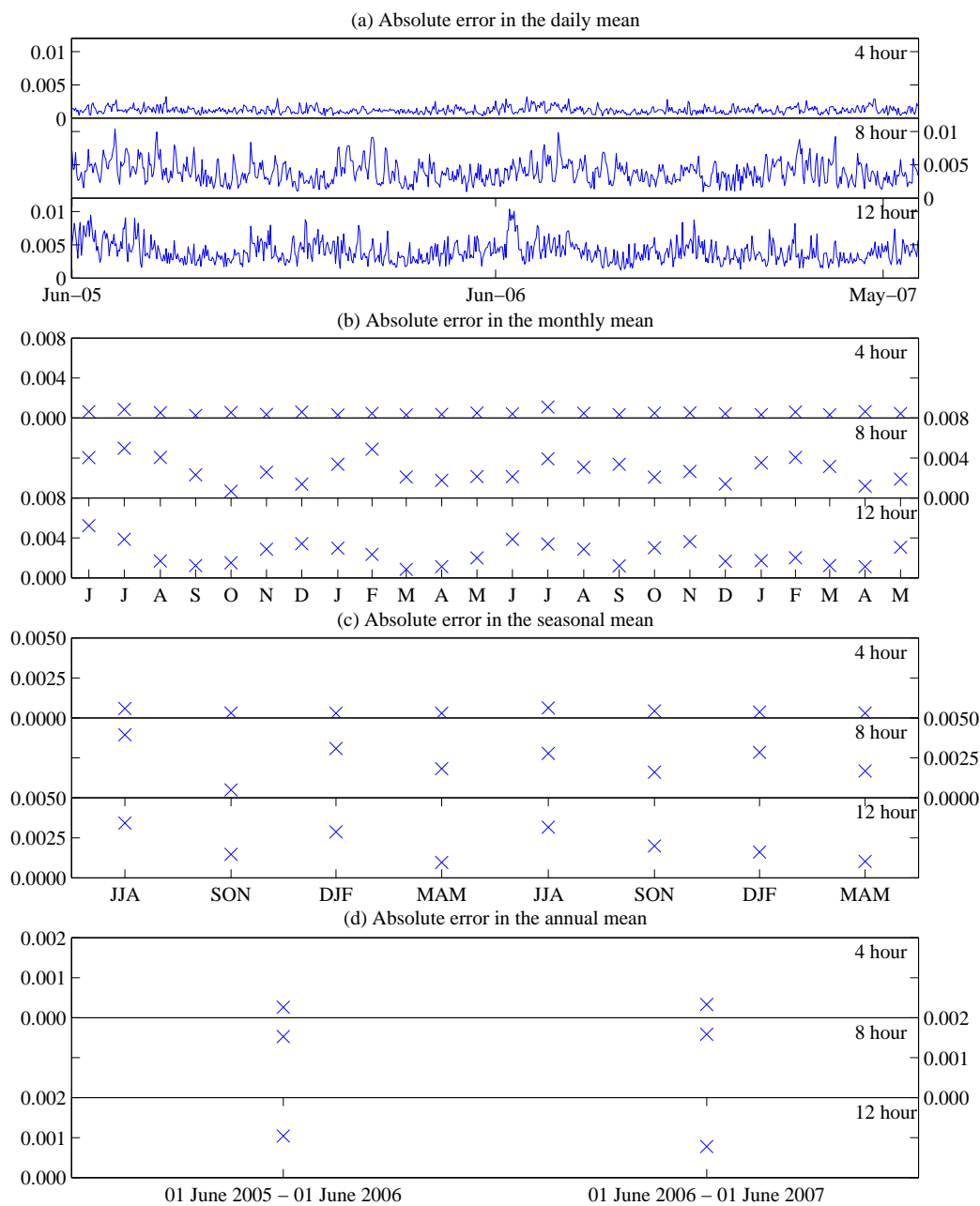**Figure 5.** All panels (**a**–**d**) as for Figure 3 but for the sunlit ocean.

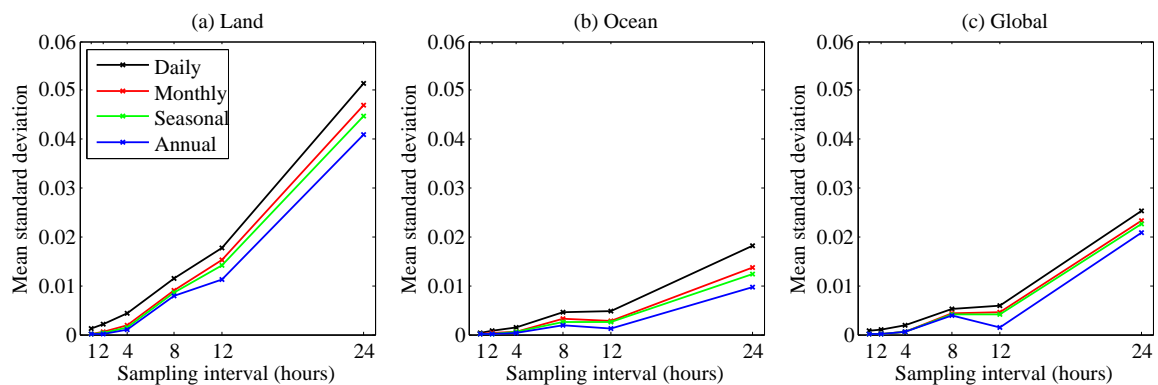**Figure 6.** All panels (**a**–**d**) as for Figure 4 but for the sunlit ocean.

Figure 7. The mean of the standard deviations $\overline{\sigma_{r,n}}$ for (**a**) land; (**b**) ocean; and (**c**) globally.
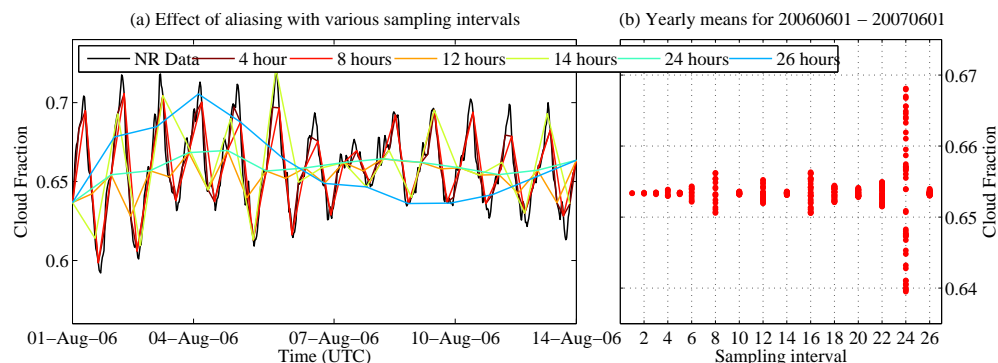
Figure 8. (**a**) The time series of sunlit ocean cloud fraction for the two week period beginning 1 August 2006 at 0000 UTC—curves show different sampling frequency but with the same starting point; (**b**) The mean for the year beginning 1 June 2006 at 0000 UTC for different sampling frequencies and all possible starting points.

To further understand the varying of uncertainty with sampling frequency, the FFT of the "truth" time series is analyzed. Figure 9 shows the power spectrum $|T|^2$ against period (in hours) for the (a) land and (b) ocean "truth" time series. Only the portion of the spectrum for wavelengths between 1-h and 25-h is shown. The FFT is computed for a six month period from 1 Jun 2006 at 0000UTC to 1 December 2006 at 0000UTC. Peaks in the spectrum highlight where the original time series contains a strong underlying signal. The peak at 24-h in Figure 9b, for example, correlates with the highly diurnal cycle of ocean cloud fraction that was identified in Figure 8a. Examining Figure 9a, it is clear that this diurnal cycle is present in the land region as well. It is not surprising to see such a large signal with a 24-h period, since this matches the orbital period. For frequencies greater than 24 h, there are no additional peaks and the spectrum increases approximately linearly. This is omitted from Figure 9 to allow focus on the more interesting part of the spectrum. The linear increase in the spectrum throughout appears to result from a so-called red noise in the time series [8]. The spectrum for a time series of red noise with an autocorrelation coefficient of around 0.95 was examined and found to have a very similar rate of increase as well as structure in the low frequency oscillations.
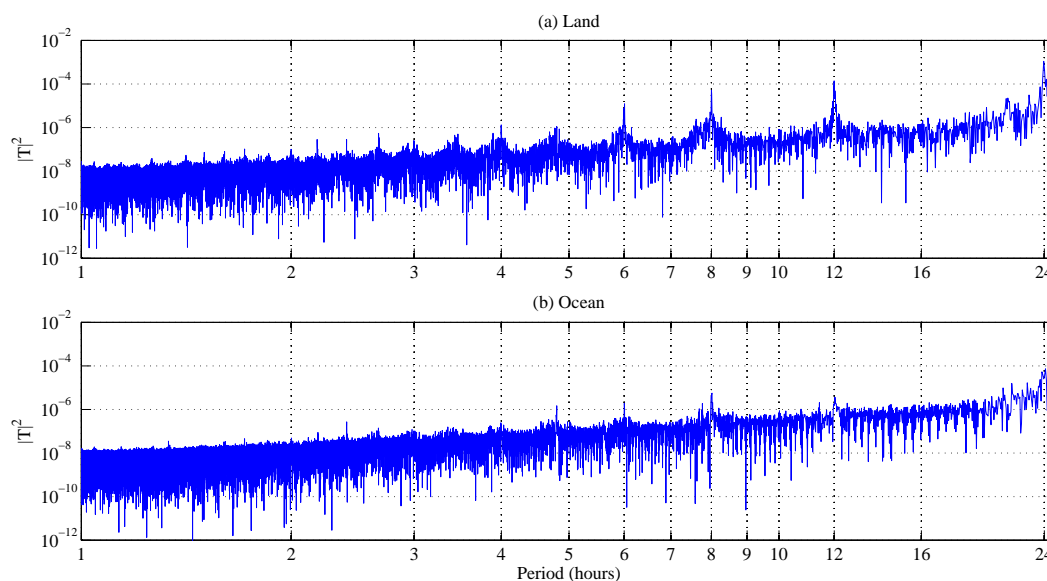

Figure 9. Power spectrum of the Nature Run time series of mean (**a**) land; and (**b**) ocean cloud fraction for 1 June 2006 at 0000 UTC to 1 December 2006 at 0000 UTC.

The results in Figure 7 show that mean uncertainty was higher over land than over ocean. The magnitude of the spectral coefficients for land are higher than for ocean, and the peaks have more contrast with other wave number signals. Peaks in the land region spectrum occur not just for 24-h but also for 4-, 6-, 8-, and 12-h, which are all shown in Figure 7a as well. For daily means, it remains true that the finer the sampling frequency, the smaller the mean uncertainty. The reason that the 12-h sampling frequency has smaller uncertainty than that of the 8-h over ocean (Figure 7b) is because the magnitude of $|T|$ is larger for 8-h than 12-h.

Given the above results, it is clear that a sampling frequency should be used that does not match peaks in the power spectrum. For computing monthly, seasonal, and annual means of cloud fraction, this means avoiding sampling frequencies of 4-h, 6-h, 12-h, 24-h, and multiples thereof. It has already been noted that a sampling frequency of not more than 4-h should be used for computing daily means.

The results above consider the uncertainty in the period means. Another consideration is the aliasing associated with a particular sampling frequency. This provides insight into the uncertainty in a particular instantaneous measurement. For example if the area mean cloud fraction is computed from DSCOVR, how uncertain is it at a particular moment, rather than over a particular period? Figure 10 shows the mean $\overline{R_n}$ (solid curves) and standard deviation $\sigma_{R,n}$ (dashed curves) of the correlation coefficients for land, ocean, and global regions between the subsamples and the "truth". Correlation coefficients are computed for daily, monthly, seasonal, and annual intervals then averaged over all of the occurrences of those periods.
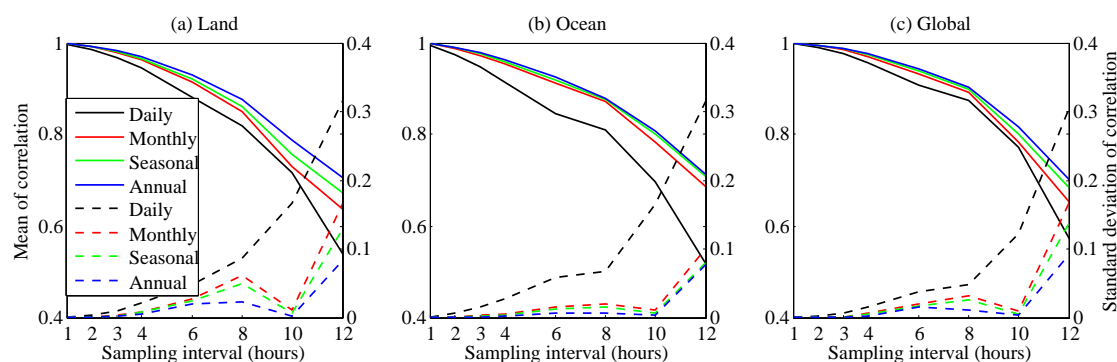

Figure 10. The mean of the correlation coefficients $\overline{R_n}$ (solid curves) and the standard deviation of the correlation coefficients $\sigma_{R,n}$ (dashed curves) for (**a**) land; (**b**) ocean; and (**c**) global regions. The axis on the left of the panels shows the mean values, and the right axis shows the standard deviation values.

The correlation coefficients decrease as the sample frequency becomes coarser. This demonstrates that even though it may be possible to obtain a mean that is close to the "truth", subsampled time series are farther and farther away from the "truth" as the temporal resolution becomes coarser and coarser. There is little difference among land, ocean, and global regions, demonstrating that the aliasing is similar for each.

4. Conclusions

The time series of cloud fraction on the sunlit hemisphere of Earth as produced by the GEOS-5 high resolution Nature Run has been analyzed in support of the DSCOVR mission. DSCOVR orbits around the Sun–Earth L1 Lagrange point, providing an uninterrupted view of the entire sunlit hemisphere. The aim of this work is to simulate the observations of cloud fraction observed by DSCOVR and to understand how temporal sampling frequency affects the observations.

A set of metrics have been used to analyze the time series of the sunlit mean cloud fraction. These are uncertainty in the mean across a range of possible start times, absolute errors in the mean, and correlation between original and sampled time series. In addition to these metrics, a Fourier spectral analysis of the time series was also conducted.

Cloud fraction was examined over land and ocean separately. Generally it was found that using a sampling frequency finer than 4-h resulted in small uncertainties. Over ocean, the uncertainty in annual mean cloud fraction decreased when using 12-h sampling frequency *versus* 8-h, which was caused by an underlying signal in the time series with a period of 8-h. Fourier analysis revealed a number of strong underlying signals, notably with 6, 8, 12, and 24-h periods (all integer factors of 24). Comparing with other results, it is shown that uncertainty can be significantly higher when the sampling frequency matches these periods. To achieve accurate results on the means, DSCOVR needs to avoid sampling frequencies that correlate with peaks in the power spectrum.

Analysis of correlation coefficients between the subsamples and the "truth" shows that even though subsampling at certain longer time intervals may not increase the uncertainty in the mean significantly, the subsampled time series are further and further away from the "truth" as the sampling interval becomes larger and larger. These results provide helpful insights for DSCOVR's temporal sampling strategy.

Author Contributions: Y. Yang proposed undertaking this work, and D. Holdaway carried out the experiments and generation of the data from the GEOS-5 Nature Run and operational forecasts. Both authors chose the metrics used to analyze the data and designed the presentation of the results. D. Holdaway prepared a first draft of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yang, Y.; Marshak, A.; Mao, J.; Lyapustin, A.; Herman, J. A method of retrieving cloud top height and cloud geometrical thickness with oxygen A and B bands for the Deep Space Climate Observatory (DSCOVR) mission: Radiative transfer simulations. *J. Quant. Spectrosc. Radiat. Trans.* **2013**, *122*, 141–149.

2. Gebremichael, M.; Krajewski, W.F. Characterization of the temporal sampling error in space-time-averaged rainfall estimates from satellites. *J. Geophys. Res.* **2004**, *109*, doi:10.1029/2004JD004509.

3. Tiao, G.; Reinsel, G.; Xu, D.; Frederick, J.H.; Zhu, X.; Miller, A.J.; DeLuisi, J.J.; Mateer, C.L.; Wuebbles, D.J. Effects of auto-correlation and temporal sampling schemes on estimates of trends and spatial correlation. *J. Geophys. Res.* **1990**, *95*, 20507–20517.

4. Holdaway, D.; Yang, Y. Study of the effect of temporal sampling frequency on DSCOVR observations using the GEOS-5 nature run results (Part I): Earth's radiation budget. *Remote Sens.* **2016**, *8*, 98.

5. Gelaro, R.; Putman, W.M.; Pawson, S.; Draper, C.; Molod, A.; Norris, P.M.; Ott, L.; Prive, N.; Reale, O.; Achuthavarier, D.; *et al*. *Evaluation of the 7-km GEOS-5 Nature Run*; Technical Report Series on Global Modeling and Data Assimilation 36; NASA Global Modeling and Assimilation Office: Greenbelt, MD, USA, 2014.

6. Bloom, S.; da Silva, A.; Dee, D.; Bosilovich, M.; Chern, J.D.; Pawson, S.; Schubert, S.; Sienkiewicz, M.; Stajner, I.; Tan, W.W.; *et al*. *Documentation and Validation of the Goddard Earth Observing System (GEOS) Data Assimilation System—Version 4*; Technical Report Series on Global Modeling and Data Assimilation 26; NASA Global Modeling and Assimilation Office: Greenbelt, MD, USA, 2005.

7. Molod, A.; Takacs, L.; Suarez, M.; Bacmeister, J.; Song, I.S.; Eichmann, A. *The GEOS-5 Atmospheric General Circulation Model: Mean Climate and Development from MERRA to Fortuna*; Technical Report Series on Global Modeling and Data Assimilation 28; NASA Global Modeling and Assimilation Office: Greenbelt, MD, USA, 2012.

8. Mann, M.E.; Lees, J.M. Robust estimation of background noise and signal detection in climatic time series. *Clim. Chang.* **1996**, *33*, 409–445.

9. Thomson, D.J. The seasons, global temperature, and precession. *Science* **1995**, *268*, 59–68.

10. Foster, J.; Richards, F.B. The gibbs phenomenon for piecewise-linear approximation. *Am. Math. Mon.* **1991**, *98*, 47–49.

11. Boucher, O.; Randall, D.; Artaxo, P.; Bretherton, C.; Feingold, G.; Forster, P.; Kerminen, V.-M.; Kondo, Y.; Liao, H.; Lohmann, U.; *et al*. Clouds and aerosols. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2013.

12. Myhre, G.; Shindell, D.; Bréon, F.-M.; Collins, W.; Fuglestvedt, J.; Huang, J.; Koch, D.; Lamarque, J.-F.; Lee, D.; Mendoza, B.; *et al*. Anthropogenic and natural radiative forcing. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK; New York, NY, USA, 2013.

13. Wylie, D.; Jackson, D.L.; Menzel, W.P.; Bates, J.J. Trends in global cloud cover in two decades of HIRS observations. *J. Clim.* **2005**, *18*, 3021–3031.

14. Frey, R.A.; Ackerman, S.A.; Liu, Y.; Strabala; K.I.; Zhang, H.; Key, J.R.; Wang, X. Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5. *J. Atmos. Ocean. Technol.* **2008**, *25*, 1057–1072.

15. Yang, Y.; Di Girolamo, L. Impacts of 3-D radiative effects on satellite cloud detection and their consequences on cloud fraction and aerosol optical depth retrievals. *J. Geophys. Res. Atmos.* **2008**, *113*, doi:10.1029/2007JD009095.