Global Spectroscopic Survey of Cloud Thermodynamic Phase at High Spatial Resolution, 2005-2015: Response to Reviews

David R. Thompson[1], Brian H. Kahn[1], Robert O. Green[1], Steve A. Chien[1], Elizabeth M. Middleton[2], and Daniel Q. Tran[1]

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[2]Goddard Space Flight Center, Greenbelt, MD, USA

Correspondence to: David R. Thompson (david.r.thompson@jpl.nasa.gov)

1 Summary

This document combines four author responses that have already been uploaded to the public discussion under each reviewer's statement. We agree with nearly all the reviewer suggestions and have incorporated them into a new revision, appended at the end with changes tracked in red. A point-by-point response for each review follows below, with panel comments in blue, green, cyan, and magenta.

2 Response to RC1

This is an interesting use of Hyperion/EO-1 data to investigate cloud phase and, broadly speaking, spatial statistics. A significant part of the study is a juxtaposition of Hyperion and AIRS data. Cloud detection and classification methods for both instruments are compared, as are the resulting cloud climatologies for the lifetime of EO-1 (2005-2015). However, the new and timely result is to use the fine spatial resolution of Hyperion to build scale-by-scale statistics, basically, semi-variograms that show a robust scaling regime once the natural variability emerges from the instrumental noise. A 3-parameter nonlinear function is used to fit the data. This is a timely development because of the emergence of scale-aware parameterizations of subgrid cloud processes in GCMs.

This manuscript thus has the potential to become a significant contribution to the literature. However, it needs in my opinion a major revision to get there. What made me struggle with the narrative was the decision to use the classic "2. Method" (should be plural) then "3. Results" structure when there are actually three distinct exercises in data analysis: (1) cloud phase classification, (2) comparison with AIRS, and (3) spatial scale analysis. In their revision, I urge the authors to write three different sections on these topics, each with its subsections on "method" and "results". That way, we don't have to slog through a bunch of method

descriptions with applications postponed for several pages. I strongly suggest a "describe, use (show Figs., etc.), and move on" structure iterated three times. That would become a much more powerful build up to the truly new and interesting results on spatial scale analysis.

We reorganized the manuscript following this suggestion. There are now three independent sections on the cloud phase retrieval, AIRS comparisons, and spatial scale analysis. Each has its own method and results. The content does not change significantly, but we added introduction text to present the organization and improve narrative flow. Overall, we agree with the reviewer on this new organization and feel the manuscript is stronger for the change.

Questions to address on p. 11, eqs (14-16): Exponent in tropics is 2/3, the classic turbulence value. Nice that the N and S extra-tropical counterparts are very close (as are, to some extent the multiplicative and additive constants). But why the significant difference with classic 2/3? Same question about the multiplicative and additive constants? Significance of N-S difference in prefactor of scaling term?

On significance, there are two related questions: (1) is the difference statistically significant, and (2) is a statistically significant difference physically meaningful? We moved the model equations into a table of coefficients along with 95% confidence intervals to show the statistical significance and better isolate the two questions. We have also added confidence intervals to the curves in Figure 9. The scaling exponents are not significantly different between Northern and Southern hemispheres; their mutual alignment relative to the tropical case demonstrates that even the unevenly-sampled Hyperion dataset shows a significant difference in spatial scaling properties between distinct cloud populations. We added explanatory text, with two new references:

> We report scaling exponents in terms of variance; these could be translated to other conventions using structure functions or the power spectrum domain. Differences between the power law exponents for tropical and extra tropical clouds were statistically significant. Differences between the power law exponents for tropical and extra tropical clouds were statistically significant. The extratropical scaling exponents of 0.42 and 0.44 are similar to, but slightly in excess of the classic Kolmogorov scaling of 1/3 (-5/3 in the power spectral domain). The tropical scaling exponent of 0.62 is in excess of the classic Kolmogorov scaling of 1/3 but is consistent with tropical cloud reflectance variability reported in Barker et al. (2017), and mid-tropospheric water vapor mixing ratio in the tropics from the AIRS instrument, e.g. Kahn et al. (2011). At finer spatial resolutions there is also evidence of scale breaks dependent on 20 altitude. Consequently the scaling exponent is dependent on the length scale range calculated (Kahn et al., 2011).
>
> The additive offset c defined the variance at zero distance, i.e. the irreducible measurement error for each spectrum. This implied a noise equivalent change in liquid water fraction of approximately 7.5% for tropical and 10-11% for extra tropical clouds. The addend and prefactor coefficients differed significantly between extra tropical and tropical clouds, and to a much smaller degree between the two extra tropical populations. We note that all three populations were subject to the Hyperion datasets' biased sampling of ocean and land, and of instrument noise conditions dominated by solar zenith angles. Consequently we would not ascribe the differences between Northern

and Southern Hemispheres to cloud scaling properties, since these are small relative to the notable difference with the tropical population. In all cases, measurement error was a small contributor to observed variance — most of the variability arose from spatially-correlated structure. Outside the tropics, measurement error accounted for just 25% of variance observed at GCM grid scales of 100 km. The remaining 75% was therefore attributable to spatial structure at subgrid scales.

Out of curiosity I plotted the fits in Eqs. (14-16) in 2 ways: same axis limits as in Fig. 9, and lowering the $y_m in$ enough to see all. T curve looks the same. Something amiss with N and S.

This was due to a typo in our transcription of the offset coefficients, now remedied. The scaling exponents, scaling prefactors, chart, overall trends, and conclusions do not change. We thank the reviewer for catching this error.

2.1 Sequential Comments

- p. 3, Eq. (1): Use the conventional / once, rather than twice "-1". Fixed.
- p. 3, Eq. (1): Add the customary subscript $_0$ to θ for SZA. Fixed.
- p. 3, Eq. (2): Coefficient b usually precedes the variable λ in the writing of a 1st order polynomial. Fixed.
- p. 3, l. 10: Probably a good idea to list only water vapor, rather than generic gases, since it is the only one considered Then assign it explicitly to j=1, e.g., water vapor (j=1); then repeat, but incremented, for liquid and ice water. Then we know exactly what j is all about. Fixed.
- p. 3, Eq. (3), line 1: The first 3 terms are used to fit -log(a+bλ); only two parameters required, and your (m-n) is one of them. Please make this crystal clear in the math, not just with a vague justification in the following text. Fixed.
- p. 3, Eq. (3), line 2: - Fuzzy math: a formal vector of parameters can't be 0, but each of its elements can. Best to use words in this case. I understand that the fitting algorithm enforces positive values, hence the apparent need to use m and n even though one has to be 0 if the other is not. Better to say that both signs are tested for the ramp function. Fixed.
- p. 4, l. 15: What are the important ETW$_x$ properties? The outcome of (3) maybe? Please define quantity, not just acronym. We now define the quantity as the outcome of the prior expression, e.g. the optical absorption path length in millimeters.
- Table 1: I'm a bit confused by Land, Vegetation and Desert ... over ocean, but these are just results from the band math in (5-7). Right? Maybe clarify in caption. We removed these confusing labels.
- p. 5, l. 17: Unless ramp slope is redefined earlier, m already used. Changed to ℓ which is unused.
- p. 6,l. 1: n->m We have changed this symbol to ℓ.
- p. 6, l. 30: Do you mean LTF(x)? Rather that denote LTF by x. See next item. Yes, you are right - we fixed this notation.

- p. 7, Eq. (13): Isn't h actually d? Also, the squared difference is in *dependent* variables: LTF(x), I believe. Fixed; thank you for this catch.

- p. 7, l. 5: Marcotte (1993) -> (1996), also in References (p. 18, l. 27). Very interesting and, for this study, enabling paper, BTW, that I had to look into. That is how I uncovered the apparent error in year of publication. Fixed, thank you.

- p. 7, l. 12: What is a degenerate variogram? We clarified that we excluded scenes with variograms that evaluated to zero at short distances due to the lack of sufficient contiguous cloud pixels.

- p. 7, l. 14: This and the next paragraph should be a designated subsection on cloud phase discrimination, or something to that effect, having 3 figures to their credit. Agreed. we made this change.

- p. 8, Fig. 2: For completeness, please indicate date and location. And make it visible (different color or arrow?) on Fig. 1. Done.

- p. 10, l. 4: Profiles –> Distributions (since not along vertical here). Fixed.

- p. 11, l. 9: exponential scaling factor –> scaling exponent Fixed.

- p. 11, l. 10: exponential factor –> exponent Fixed.

- p. 13, Fig. 7: Distill caption down to "As in Fig. 6, but for ice phase." Fixed.

- p. 14, Fig. 8: One legend is enough, preferably located in the middle. Fixed.

- p. 15, Fig. 9: Why do I see the smallest (maybe Nyquist due to FFT-based estimation?) [wavenumber] of 90 m? I thought Hyperion pixels were much smaller. We assume the reviewer means "lag distance" rather than "wavenumber." We used selected lag distances to ensure a log-constant density. The first two bins corresponded to the identity bin which was not plotted, and 90m. Selected bins were dense near zero, and became widely separated at long ranges.

- p. 15, Fig. 9: Please distinguish or mark the N and S extra-tropical data and fits. Is N/S offset physically significant? The plot has been fixed. See the answer above with respect to significance. We emphasize -and have now written into the text- that each datapoint corresponds to a lag distance for the aggregate variogram, so separation between the curves does not imply separation between populations.

- p. 17, Fig. A2: Distill caption down to "As in Fig. A1, but for ice phase." Fixed.

3 Response to RC2

This manuscript develops a very useful approach to investigate cloud thermodynamic phase based on data from 2005-2015 obtained by the Hyperion imaging spectrometer on EO-1. The approach combines spectrum fitting and spatial scale analysis.

The validity is demonstrated with a comparison with AIRS. I recommend the work for publication in AMT after addressing/clarification of the comments listed below.

Page 3, Eq. (3): The introduction/use of 'm' and 'n' is confusing. The development in equations (2-3) should be clarified. What about error due to this approximation?

Reviewer 1 made a similar suggestion. We changed this formula to use a single slope value, which makes the linear continuum more obvious. We clarified that we present both positive and negative slope (one of which must be zero) to the nonnegative least squares solver. Naturally, the notion of a continuum is a convenience, and the cloud particle scattering will invariably depart from a perfect linear relationship. However, linear continuua over the 1.4-1.8 micron interval were effective in prior modeling and validation experiments (e.g. Thompson et al., JGR 2016). A relatively simple 5 Degree of Freedom model fit all channels within our conservative noise estimate (see discussion below), suggesting that the model captured the major physical processes in play.

Page 4, line 8-9: 'These bulk absorption spectra were generally independent of particle scattering and did not relate directly to particle size'. Is there a reference or evidence for this statement?

This statement was simply a definition of the Kou et al. coefficients. In other words, the attenuation uses the same absorption coefficients that one would use for a thin sheet or volume of water or ice. We have added more text to clarify: "These bulk absorption spectra were molecular properties of H_2O, independent of particle size and scattering — a common practice for Shortwave Infrared observations of clouds (Kokhanovsky, 2004)."

Page 5, Eqs. (8-9): Are the "m" and "n" the same variables in Eq. (3)?

Reviewer 1 also noted this repetition; we've switched to a new variable here for clarity.

Page 5, Eq. (9): [Should] be divided by number of degree of freedom because your fit uses a reduced Chi-Square, which is defined as 'Chi-Square per degree of freedom'. So $\chi^2 = 1$ is a conservative estimate of measurement noise as shown in Fig.4. Please clarify.

We added text to clarify: "Specifically, [the reduced χ^2 score was the Chi-square score per degree of freedom, with χ^2=1 equivalent to estimated measurement noise. This was more appropriate than a classical Chi-square test for ur spectroscopic observations where errors could be correlated across adjacent wavelengths."

Page 11, Eqs. (15-16): I believe that the authors made a typo for the offsets. It should be [0.012 and 0.010]. Please check.

Reviewer 1 also noted this. We have corrected the typo in our revision.

4 Response to RC3

Recommendation: this paper has the potential to be an interesting contribution to knowledge, but requires major revision before being accepted as a publication.

The comparison with AIRS is flawed. It appears to be a purely statistical comparison, involving mainly land scenes on the part of Hyperion, and global scenes on the part of AIRS. Given the variability of clouds and the sharp differences between

maritime and continental clouds, the AIRS data should have been subsetted to match the Hyperion locations. Mention should be made of the differences (or similarity) between the Hyperion and AIRS sampling: I think that they sample at completely different times of the day? So even if both instruments were retrieving cloud phase perfectly, the comparison would be flawed by the different sampling strategies. The results shown in Figs. 6, 7, A1 and A2 consequently are troublesome to interpret.

We agree the two datasets are different. We would suggest that this is not necessarily a flaw, but rather conditions one must consider in interpreting the comparison. We intend it mainly as a "sanity check" of broad latitudinal distributions, and feel it is independently interesting for the fact that the instruments use very different measurement strategies. However, we agree that the differences are important and our revision calls them out from the start:

> [This] was a dramatically different measurement obtained from thermal infrared spectra with a coarse 13.5 km footprint rather than reflected solar energy at fine spatial resolution. Kahn et al. (2014) detail the algorithm, and Jin and Nasiri (2014) validate it using pixel-scale comparisons with CALIPSO data [....] We anticipated several differences in the result. First, AIRS sampled uniformly over the Earth's surface while Hyperion imaged only during the day and favored land areas. We also expected differences in sensitivity; AIRS was far more sensitive to thin clouds, while the Hyperion analysis intentionally excluded them with a strict cloud mask [...] Additionally, the AIRS algorithm classified ambiguous clouds as "unknown." This population likely contained mixed phase clouds but also a large fraction of supercooled liquid clouds due to the current AIRS phase algorithm [...] While we expected some discrepancies due to differences in instruments and sampling, the comparison provided a useful check between two very different measurement techniques.

Regarding the comments on comparing at the pixel scale, we agree with the reviewer that this is the most robust method that minimizes temporal and spatial mismatching. However, since the two instruments are in different orbits, and the Hyperion record is limited to targeted acquisitions, this would leave insufficient coincidences to provide robust statistics. A strict spatial distance and time difference criterion would have to be used in order to account for diurnal variability. These issues would persist even if we subset AIRS to the spatial points of the EO-1 observations, without requiring exact temporal coincidence.

Naturally, it is fairly common practice to compare data sets that retrieve the same geophysical variable and make statistical comparisons from completely different satellite platforms and spatial/temporal sampling. This is done with cloud properties including cloud microphysics, IWP, and LWP. One example is Stubenrauch et al., ASSESSMENT OF GLOBAL CLOUD DATASETS FROM SATELLITES: Project and Database Initiated by the GEWEX Radiation Panel, Bull Amer Met Soc, 2013. These types of comparisons still yield scientifically important insights.

Finally, we would emphaisze that the main Hyperion result stands apart as an independent contribution, and that the latitudinal distributions are consistent those reported for more similar instruments such as MODIS, as summarized in Hirakata et al. (2014). The main contributions of the manuscript are to demonstrate the first global scale cloud phase measurement from reflectance spectroscopy, to provide the first global study imaging cloud phase at 30 m spatial sampling, and to assess spatial scaling properties.

They should have addressed the sampling errors of each instrument, not simply a vague error bar for Hyperion and nothing for AIRS. The comparison is also weakened by the empirical correction factors to AIRS data discussed in section 2.5.

Our revision clearly describes our methodology for calculating 95% confidence intervals: "[We calculated] confidence intervals with nonparametric bootstrap variance estimation (Wasserman, 2006) that resampled the dataset 10,000 times with replacement." We now add that "The corresponding AIRS error bars would be far smaller due to the large number of samples, so we omit them for clarity."

The definition of LTF is flawed. The signal measured is based on the absorption of solar radiation integrated over the entire photon pathlength, yet eq. 2 refers only to the thickness of the cloud.

In fact, the reviewer has interpreted our equation exactly as we had intended: the LTF refers to the absorption along the photon optical path, with no implication for the physical vertical dimensions of the cloud. We call it a "thickness" for consistency with prior literature, such as Gao and Goetz (1995). This is also consistent with our own previous usage (e.g. Thompson et al. 2015, 2016).

By their nature, clouds are heterogeneous, so that horizontal variability dominates the radiative transfer process. [This also means that the retrieval technique is at a coarser scale than the postulated 30 m due to the effects of radiative smoothing, and is likely closer to 100 m.] I think this is correctly acknowledged in p.4 line 13 ff. However, it is not really clear whether the LTF is being interpreted correctly. I take it to be the fraction of average photon path that is liquid. Not the fraction of the cloud that is liquid, which would require all paths to extend to the cloud base. An opaque cloud has little transmission, so that most of the reflected paths relate to the top of the cloud. This probably doesn't matter much for the Hyperion retrievals standing alone, but becomes troublesome when compared to other techniques that sample cloud tops differently. It would be good to see a clearer discussion of what is meant by the "effective proxy for thermodynamic phase."

We modified the text to clarify our definition:

> In summary, following Equation 3 we modeled the entire interval from $1.4 - 1.8$ μm with five free parameters: a continuum offset l; a slope, represented by a single degree of freedom in the variables m and n; and the vapor, ice and liquid thicknesses u_j. These thicknesses represented the length of the optical path through an equivalent homogeneous volume, as in the Equivalent Water Thickness (Gao and Goetz, 1995). As in previous work, we wrote the absorption path length u_2 as the Equivalent Water Thickness due to Liquid in millimeters, EWT_{liquid}. Similarly, u_3 was the Equivalent Water Thickness due to ice, EWT_{ice}. We then defined the Liquid Thickness Fraction (LTF) as:
>
> $$\text{LTF} = \frac{\text{EWT}_{liquid}}{\text{EWT}_{liquid} + \text{EWT}_{ice}} \quad (1)$$
>
> Prior in situ validation had demonstrated a robust relationship between the LTF and thermodynamic phase (Thompson et al., 2016). We emphasize that "thickness" referred to the absorption along the optical path; clouds were heterogeneous, so the LTF was not necessarily related to their vertical dimension. In opaque clouds the measurement would be most sensitive to the upper layers.

Section 2.3 is a strange, stand-alone paragraph that seems incomplete. How is 'dominant' defined? Greater that 50%? What comparisons were made with historical datasets? This section should be rewritten to provide better context, or incorporated elsewhere.

Following a comment by another reviewer, we have restructured the manuscript to follow a more thematic organization which places this paragraph in better context. We have modified the text to indicate that we used a 50% cutoff threshold.

Section 2.4 presumably refers to the uncertainty in determining the LTF of a single scene, but this is not clear. It also stops abruptly with no relation to the esults. This needs to be rewritten for clarity and context.

We agree; our manuscript restructuring clarifies the implication of these χ^2 values. The fits demonstrate that the model explains the variability observed in the spectra to within our noise estimate, showing that the retrieval method of Thompson et al. (2016) also applies to Hyperion. We have added text to this effect. We also state that noise is calculated on a per-line basis (it is dominated by constant factors like the solar zenith). However, we calculate χ^2 statistics independently for each spectrum, since we fit a model independently for each 30 m $\times$ 30 m spatial location.

Section 2.5 should provide a reference to how AIRS obtains cloud phase and whether this has ever been validated.

We now state: "This was a dramatically different measurement obtained from thermal infrared spectra with a coarse 13.5 km footprint rather than reflected solar energy at fine spatial resolution. Kahn et al. (2014) detail the algorithm and validation."

The use of the word 'trends' p.5, p7, p.13. This is better reserved for long term climate change. Here we are looking at 'relative dependence on latitude' or similar.

We substitued the term "distributions."

Fig. 4 shows results for clear retrievals, yet the scene looks completely overcast. Are these all in error, despite the low values of χ^2 for many of these? Given the range of χ^2 shown, presumably the only results retained where when χ^2 was less than some threshold? This could be discussed better.

We understand how the clear sky case could be confusing. In fact, the scene is a fragment of a larger image that included clear sky areas, and these statistics come from the clear parts (which are outside the area shown in our figure). Ironically, the χ^2 values for the clear cases had been low despite the fact that the instrument saw the Earth's surface. The image was acquired over an ice shelf and ocean, and the former showed ice absorption while the latter had nearly zero reflectance in this range. Since we do not include clear sky cases in our global statistics (for obvious reasons) there is no reason to report these χ^2 values. For clarity, we have removed the case from the histogram and the text.

Fig.5 is too cryptic for the typical reader. If the vapor transmittance around 1.4 μm is zero, how can there be any reflectance to work with? Does the theory include the vapor paths both above and within the cloud? Probably need to explain what is meant by transmittance in this context.

Thank you - this was a good catch. It was an artifact of our figure, which had used the first order Taylor approximation of the transmittance. The approximation was not accurate near saturation. We replotted the figure using the true transmittances, which were naturally greater than zero. The product of plotted transmittances (and the continuum, not shown) now reproduces the observed TOA reflectance. The changes are very subtle outside the 1.4 micron vapor feature, and the general shapes and relative depths are not significantly altered.

Normalization of occurrence: p.6, l.11 is -60 to +60°, Fig. 6 is 0 to 60°. Which is it? Is the normalization done separately for each cloud phase? Are the AIRS data similarly normalized?

This was a typo in the figure caption; the normalization uses the whole -60 to +60° interval. We updated the caption, and changed the text to emphasize that both AIRS and Hyperion are normalized.

Fig. 8 is flawed by the nonuniform sampling with latitude. Perhaps an indication of the relative number of samples per histogram would help.

We have added a note to the caption and text reminding the reader that sampling is nonuniform, and that the results should be evaluated in light of the bootstrap uncertainty analysis. This figure is presented in the context of Figures 6-7, where bootstrap variance estimation shows the uncertainty due to sample size in each histogram bin.

p.10 line 9. Appendix 4? Appendix A.

Fixed, thank you.

Fig. 9 shows NH and SH curves for extra tropical clouds, but which is which? Eq. 14-16 don't seem to match the values on the figure.

We have remedied a typo in the offset values for these equations - now they match the figure. We have also changed the figure, labeling the two extra tropical curves.

p.13, line 4. This caveat comes far too late in my opinion as it dominates the comparison throughout. Note that CALIOP also offers high-resolution phase information that also has fewer sampling limitations.

The new version emphasizes nonuniformity throughout, in the following locations:

- Section 2.1: "[Hyperion] performed targeted acquisitions for specific regions of interest, with occasional pointing off nadir. Most targets were on land, with a high concentration in the mid-latitude northern hemisphere. There was sparser coverage of extreme latitudes and oceans, but several island targets offered a view into cloud systems over ocean (Figure 1). Many targets of interest were revisited multiple times during the mission."
- Figure 1: portrays the Hyperion image locations
- Section 2.2: "Note that Hyperion sampling is nonuniform across 20 histogram bins, and Section 1 quantifies uncertainty for different latitudes."
- Section 3.1: "AIRS sampled uniformly over the Earth's surface while the Hyperion dataset favored land areas."
- Section 5 (conclusions): "The Hyperion datasets were spatially biased and strongly favored land mass over ocean. Insofar as the latitudinal trends show asymmetries across northern and southern hemispheres, this may be related 15 to the spatial distribution of land mass in the southern hemisphere midlatitude areas. Southern hemisphere observations were often acquired over islands, which would exhibit a more oceanic influence on cloud cover."

We believe that the closing discussion is an appropriate place to contextualize these results and draw implications. Obviously the Hyperion mission was not designed for cloud observations. However, the ability to form cloud phase maps at 30m resolution, is a unique new capability that makes the investigation meritorious. Additionally, results generally agree with existing

cloud phase records from AIRS (and multi-instrument comparisons such as Hirakata et al., 2014). Because the measurement technique is distinct, is is a useful complement to other instruments using polarization (CALIOP) or thermal emission (AIRS). We have modified the conclusion to emphasize this. Finally, the Hyperion datasets provide a "first of a kind" observational record at sub-kilometer scales. Spatial granularity reaches a factor of three below CALIOP observations, though one should also note the reviewer's caveat about within-cloud scattering placing a lower limit on achievable resolution.

5 Response to RC4

Multilayered cloud systems: I found no description on how multilayered cloud systems are detected and handled in this study. In my view, "ice" cloud region shown in Fig. 2 looks like a multilayered cloud system with an optically thin, high cloud above an optically thick, low cloud deck. I am not sure on this because I am not an expert of this kind of imagery, but I was wondering why "ice" cloud region is more reflective than "mixed phase" cloud region.

In fact, the reflectance of the mixed phase is slightly higher than for the ice phase. This is captured in our new plot. It is indeed possible that the ice cloud hides a liquid cloud below (see below).

Satellite measurements show that multilayered cloud systems are quite common in the tropics and mid-latitude storm track regions. Thermal infrared measurements by AIRS are sensitive to the upper cloud, but the SWIR reflectance from Hyperion should be more sensitive to the lower cloud, depending on the optical thickness of upper cloud. If so, there should be more liquid cloud occurrence in Hyperion's results than in AIRS, in specific latitude zones. Is this a possible reason for statistically significant Hyperion–AIRS differences in the tropics and mid-latitude storm track regions, as in Figs 7 and 8? The authors just mentioned that distributions from the two instruments generally agreed and the differences were ascribed to sampling error and spectroscopic sensitivity difference. In my opinion, if there is a statistically significant difference, that difference is valuable to be discussed and should be clarified in the manuscript. In that way, this comparison is not just a "sanity check" but more valuable.

This is an excellent point - to the degree that there is very optically thin ice cloud above a liquid cloud, AIRS and Hyperion might give two different answers. We agree that this is one potential explanation for differences in AIRS and Hyperion, in addition to spatiotemporal sampling. More generally, AIRS and Hyperion will be sensitive to different altitudes within a large cloud. We have added a note to this effect in the AIRS section, and with due deference to the reviewer, will incorporate this phraseology directly: "Another potential contributor to the discrepancy is sensitivity to different altitudes in large or multilayer cloud systems. Multilayered clouds are abundant in tropical regions and and mid-latitude storm tracks. In cases where, for example, a translucent ice cloud overlays an optically thin liquid cloud, the two instruments would measure different thermodynamic phase."

On the comparison with AIRS: Oceanic and continental averages of cloud phase fraction can be derived from AIRS data. How can the difference between them explain the difference between results from the Hyperion and AIRS? It would be more insightful to compare the Hyperion's results with AIRS oceanic and continental averages.

We absolutely agree that the next natural step in a comparison would be a closer comparison of specific spatiotemporal subsets. It was not obvious how to do that in a paper of this scope, since there is insufficient direct coincidence to provide strong statistics, and there are other differences beyond the continent/ocean biases - for example, the fact that Hyperion observed only during the day, with observations concentrated in areas with human populations. Our response to reviewer 3 describes some of the other differences. We felt that the current evaluation was a simple story, and that a partial remedy of sampling differences might mislead the readership into expectations of precise alignment.

Page 6, Line 30, "The mixed phase clouds were ... nearly absent from the tropics": It seems to be not nearly absent.

Agreed; we changed "nearly absent" to "less abundant."

Page 9, line 15, "thin cloud": Is this an optically thin, high (or low) cloud?

We have modified the sentence for clarity: "We also expected differences in sensitivity; AIRS was far more sensitive to optically thin clouds, while the Hyperion analysis intentionally excluded thin clouds with a strict detection threshold." Hyperion thresholds applied to both high and low altitude clouds.

Global Spectroscopic Survey of Cloud Thermodynamic Phase at High Spatial Resolution, 2005-2015

David R. Thompson[1], Brian H. Kahn[1], Robert O. Green[1], Steve A. Chien[1], Elizabeth M. Middleton[2], and Daniel Q. Tran[1]

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[2]Goddard Space Flight Center, Greenbelt, MD, USA

Correspondence to: David R. Thompson (david.r.thompson@jpl.nasa.gov)

Abstract. The distribution of ice, liquid, and mixed phase clouds is important for Earth's planetary radiation budget, impacting cloud optical properties, evolution, and solar reflectivity. Most remote orbital thermodynamic phase measurements observe kilometer scales and are insensitive to mixed phases. This under-constrains important processes having outsize radiative forcing impact, such as spatial partitioning in mixed phase clouds. To date, the fine spatial structure of cloud phase has not been measured at global scales. Imaging spectroscopy of reflected solar energy from 1.4 - 1.8 μm can address this gap: it directly measures ice and water absorption, a robust indicator of cloud top thermodynamic phase, with spatial resolution of tens to hundreds of meters. We report the first such global high spatial resolution survey based on data from 2005-2015 acquired by the Hyperion imaging spectrometer onboard NASA's Earth Observer 1 (EO-1) spacecraft. Seasonal and latitudinal distributions corroborate observations by the Atmospheric Infrared Sounder (AIRS). For extra tropical cloud systems, just 25% of variance observed at GCM grid scales of 100 km was related to irreducible measurement error, while 75% was explained by spatial correlations possible at finer resolutions.

1 Introduction

The distribution of ice, liquid, and mixed phase clouds is important for Earth's climate and planetary radiation budget (Chylek et al., 2006; Martins et al., 2011). Cloud thermodynamic phase affects radiative forcing by modulating absorption of incoming solar radiation, particle evolution, and lifetime (Ehrlich et al., 2008; Tan and Storelvmo, 2016). Previous satellite observational studies have shown that clouds are shifting poleward in the Northern and Southern Hemisphere extratropical storm tracks (Bender et al., 2012; Marvel et al., 2015; Norris et al., 2016). Within these shifting storm tracks, climate model experiments with forcing from increased CO_2 have shown losses of cloud ice phase and gains of cloud liquid phase (Ceppi and Hartmann, 2015; McCoy et al., 2015). This makes cloud phase an important property for accurate and continuous monitoring. Moreover, recent studies indicate that spatial partitioning of ice and liquid particles within clouds has outsized influence on climate, affecting Global Climate Model (GCM) predictions of future warming by over 1 degree Celsius (Tan et al., 2016). However,

current observing systems cannot reduce this uncertainty; they are unable to resolve differences at critical sub-100 m scales, or are insensitive to mixed phase clouds altogether.

Imaging reflectance spectroscopy from 1.4 - 1.8 μm could address this gap. Prior work used this spectral interval to measure cloud phase based on the optical absorption properties of liquid and ice (Pilewskie and Twomey, 1987). The fraction of the total path absorption due to liquid (the Liquid Thickness Fraction, or LTF) had high sensitivity to pure and mixed phases (Thompson et al., 2016). Liquid and ice show highly diagnostic absorption shapes which are robust to potential confounding effects such as surface reflectance, observation geometry, and mismatch in particle modeling assumptions. Remote imaging spectrometers measure these spectra from cloud tops at millions of spatial locations, resolving cloud phase at tens of meter scales (Thompson et al., 2016). This could constrain characteristic spatial lengths of these processes. However, global spectroscopic datasets have not yet been analyzed in this way.

Here we report a global spectroscopic survey of cloud phase from 2005-2015 based on data from the Hyperion imaging spectrometer instrument onboard NASA's Earth Observer 1 (EO-1) spacecraft. Hyperion also provides two novel contributions beyond prior records. First, its full spectrum fitting discriminates mixed phases by directly measuring the relative contributions of physical cloud top liquid and ice absorption. Second, Hyperion measures phase at horizontal scales of 30 m. These properties allow the first rigorous characterization of the spatial scales governing cloud top thermodynamic phase. The article first describes our estimation method, and reports seasonal and latitudinal changes of cloud thermodynamic phase. We then compare these distributions to measurements by NASA's Atmospheric Infrared Sounder, AIRS (Kahn et al., 2014). Finally, we show the spatial scaling properties of both extra tropical and tropical cloud populations. The study lays the groundwork for future orbital imaging spectrometers (Mouroulis et al., 2016), that can monitor cloud characteristics when cloud cover precludes their primary mission. Imaging spectroscopy investigations typically treat clouds as contamination, when in fact cloudy data can be exploited to dramatically increase these instruments' useful data yield.

2 Cloud Phase Estimation

2.1 Method

The Hyperion imaging spectrometer operated on the sun synchronous EO-1 spacecraft for over a decade prior to decommissioning in 2017. Hyperion measured reflected solar energy from approximately 400-2450 nm with approximately 10 nm spectral sampling. It performed targeted acquisitions for specific regions of interest, with occasional pointing off nadir. Most targets were on land, with a high concentration in the mid-latitude northern hemisphere. There was sparser coverage of extreme latitudes and oceans, but several island targets offered a view into cloud systems over ocean (Figure 1). Many targets of interest were revisited multiple times during the mission. Each targeted acquisition had a Ground Sampling Distance (GSD) of approximately 30 m over a cross-track width of approximately 7.5 km, and a typical along-track distance of approximately

120 km. The Jet Propulsion Laboratory (JPL) Hyperion archive included most normal science acquisitions from 2005 through 2015, with over 4.8×10^4 scenes and 3.7×10^{10} distinct spectra.

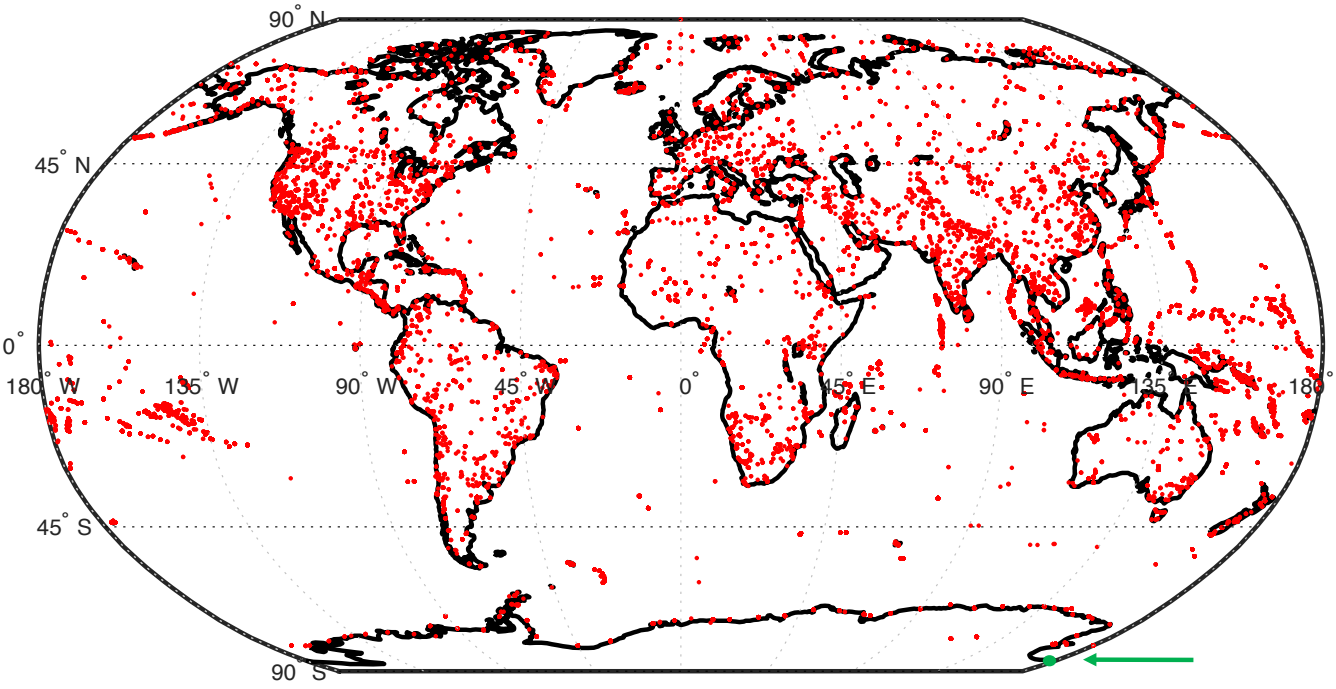

Figure 1. The JPL Hyperion archive comprised over 45,000 scenes, represented here by red dots. The majority were over land. A green arrow indicates the Ross Island acquisition of Figure 2.

The archive was stored as standard calibrated unorthorectified data. We first applied the cloud phase retrieval algorithm of Thompson et al. (2016), validated in the prior work by coincident remote and in-situ aircraft measurements. We defined the TOA reflectance, ρ:

$$\rho(\lambda) = \pi L(\lambda)/(F(\lambda)cos(\theta_0)) \tag{1}$$

where λ was the wavelength, L was the wavelength-dependent radiance at sensor, F was the extraterrestrial solar irradiance, and θ_0 was the solar zenith angle. Over short intervals we modeled ρ by a linear continuum with an offset a and slope b, attenuated by one or more Beer-Lambert absorbers j. Each absorber had a bulk absorption coefficient k_j and a nonnegative thickness u_j:

$$\rho(\lambda) = (a + b\lambda)\ \exp[-\sum_j k_j(\lambda)u_j] \quad \text{for} \quad u_j \geq 0 \tag{2}$$

We modeled three absorbers: atmospheric water vapor ($j = 1$); liquid water ($j = 2$); and ice ($j = 3$). As in Thompson et al. (2015, 2016) we applied a logarithic transformation resulting in a nonnegative least squares problem (Lawson and Hanson,

1974):

$$-log(\rho(\lambda)) \quad \approx \quad l + m\lambda + \sum_j k_j(\lambda) u_j$$
$$\text{for} \quad l \geq 0, \quad u_j \geq 0 \ \forall \ j \tag{3}$$

Here m permitted a log-linear continuum. We represented this to the solver as two nonnegative coefficients for upward and downward slopes, one of which would evaluate to zero. The nonnegative least squares solution provided stable global solutions without sensitivity to initialization.

We calculated the H_2O vapor absorption coefficients using the HITRAN 2012 line list (Rothman et al., 2013), via the Oxford Reference Forward Model (Dudhia, 2014). We initialized the vapor abundance using a band ratio retrieval as in Thompson et al. (2015), and used this to calculate an "effective" absorption coefficient of band-aggregated vapor lines for use in the least squares retrieval. Other atmospheric gases did not significantly impact the shape of vapor absorption features. We calculated liquid and ice absorption coefficients using the complex index of refraction measured by Kou et al. (1993). These bulk absorption spectra were molecular properties of H_2O, independent of particle size and scattering — a common practice for Shortwave Infrared observations of clouds (Kokhanovsky, 2004). Combined with a free continuum, they fit the observed spectra of opaque clouds over short spectral intervals without precise knowledge of particle properties (Thompson et al., 2016).

In summary, following Equation 3 we modeled the entire interval from $1.4 - 1.8$ μm with five free parameters: a continuum offset l; a slope, represented by a single degree of freedom in the variables m and n; and the vapor, ice and liquid thicknesses u_j. These thicknesses represented the length of the optical path through an equivalent homogeneous volume, as in the Equivalent Water Thickness (Gao and Goetz, 1995). As in previous work, we wrote the absorption path length u_2 as the Equivalent Water Thickness due to Liquid in millimeters, EWT_{liquid}. Similarly, u_3 was the Equivalent Water Thickness due to ice, EWT_{ice}. We then defined the Liquid Thickness Fraction as:

$$\text{LTF} = \frac{\text{EWT}_{liquid}}{\text{EWT}_{liquid} + \text{EWT}_{ice}} \tag{4}$$

Prior in situ validation had demonstrated a robust relationship between the LTF and thermodynamic phase (Thompson et al., 2016). We emphasize that "thickness" referred to the absorption along the optical path; clouds were heterogeneous, so the LTF was not necessarily related to their vertical dimension. In opaque clouds the measurement would be most sensitive to the upper layers.

We calculated the LTF of locations flagged by the Hyperion cloud detection algorithm of Griffin et al. (2003). In the Griffin et al. approach, a decision tree of threshold tests sorted spectra into different cloud types (including low and high clouds) as well as different land cover types (including snow, open terrain surfaces, and vegetation). This favored opaque clouds and generally ignored ambiguous translucent clouds that would be difficult to distinguish from land. The Griffin et al. algorithm defined tests using the top of atmosphere reflectance ρ and three intermediate quantities, the Normalized Difference Snow Index (NDSI),

	Criterion		
Test	**Over Ocean**	**Over Land**	**If True**
1	$\rho(1.38) > 0.1$	$\rho(1.38) > 0.1$	Cloud
2	$\rho(0.66) < 0.15$	$\rho(0.66) < 0.15$	Clear
3	VI < 0.6	VI < 0.7	Clear
4	DSI < 0.01	DSI < 0.05	Clear
5	-0.2 < NDSI < 0.2	0 < NDSI < 0.2	Cloud
6	0.6 < NDSI	0.6 < NDSI	Clear
7	$\rho(1.25) < 0.35$	$\rho(1.25) < 0.35$	Clear
8	$0.1 < \rho(1.38)$	$0.1 < \rho(1.38)$	Clear

Table 1. Hyperion Cloud Detection algorithm Griffin et al. (2003). We test each criterion in sequence, starting from the top, and ascribe a classification based on the first test that evaluates to true.

the Desert Sand Index (DSI), and the Vegetation Index (VI):

$$\text{NDSI} = [\rho(0.55) - \rho(1.65)] \,/\, [\rho(0.55) + \rho(1.65)] \tag{5}$$

$$\text{DSI} = [\rho(0.86) - \rho(1.65)] \,/\, [\rho(0.86) + \rho(1.65)] \tag{6}$$

$$\text{VI} = \rho(0.66)/\rho(0.86) \tag{7}$$

The method was originally formulated as a flowchart but was tantamount to an ordered sequence of tests (Table 1). It evaluated each spectrum independently, applying appropriate thresholds depending on whether the scene was over ocean or land. It ascribed a classification based on the first test that evaluated to true. Random manual validation of selected scenes demonstrated adequate performance, outside the brightest glacial scenes in Antarctica.

After calculating LTF maps over all cloud areas, we aggregated counts of liquid and ice clouds. LTF was a continuous-valued quantity so the maps revealed both pure and mixed phases. We binned LTF values in 10% graduations but also calculated binary classification using a 50% threshold. This facilitated comparisons with historical datasets of hard categorical classifications. We then analyzed zonal distributions, estimating confidence intervals with nonparametric bootstrap variance estimation (Wasserman, 2006) that resampled the dataset 10,000 times with replacement.

Evaluating the model fit to the spectrum demanded special care, since the global catalogue included many observing geometries, terrain surface types, and potential variability in instrument calibration over the decadal record. To account for this, we estimated the noise level independently for each integration timestep and each spectral channel. We used the common method of pairwise differences between spectra at neighboring locations (Boardman and Kruse, 2011). Since the spatial field was mostly uniform over small distances, these differences conservatively estimated the measurement noise σ in each channel. For n cross-track locations, we applied the nonparametric variance estimate of Von Neumann (1941), reprised in Brown and

Levine (2007):

$$\hat{\sigma}^2 = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (\rho_{i+1} - \rho_i)^2 \tag{8}$$

We then characterized the fit for each spectrum using the reduced χ^2 measure (Eldering et al., 2017), a statistical summary of the residual relative to the expected measurement errors. Specifically, it was the Chi-square score per degree of freedom, with χ^2=1 equivalent to estimated measurement noise. This was more appropriate than a classical Chi-square test for our spectroscopic observations where errors could be correlated across adjacent wavelengths. For ℓ spectral channels, with measured TOA reflectance spectrum ρ and the model estimate $\hat{\rho}$, the reduced χ^2 error was:

$$\chi^2 = \frac{\sum_\ell (\rho - \hat{\rho})^2}{\sum_\ell \hat{\sigma}^2} \tag{9}$$

The summations ran over all ℓ spectral channels.

2.2 Results

The retrievals clearly revealed distinct cloud phases. Figure 2 shows a typical example subimage. Figure 3 is the corresponding EWT$_{liquid}$ and EWT$_{ice}$ mapped to green and blue channels, respectively. LTF values in the mixed phase region range from 0.5-0.75, and values in ice areas are typically 0-0.2. Both regions were well-distinguished without significantly overlapping values, but also showed sub-kilometer interior spatial structure. Figure 4 shows normalized histograms of χ^2 scores for the entire scene, calculated independently for liquid and ice clouds. Fits were generally quite good, with χ^2 scores below the conservative noise estimate, $\chi^2 = 1$. In other words, the algorithm fit these cloud spectra to within the measurement accuracy showing applicability of the Thompson et al. (2016) three absorber model o Hyperion data.

Figure 5 shows typical spectrum fits from locations in this subimage, with mixed cloud and ice cloud. The top rows show the model fit to spectrum, confirming that the five-parameter model fit the 1.4 - 1.8 μm interval. The middle rows shows transmittance due to each absorbing component in the model. The continuum component is not shown. The three absorbers together were sufficient to explain observed spectrum shapes. Finally, the bottom rows show residual error, which was again below the estimated measurement noise level.

Figure 6 shows the entire distribution of mixed and pure phase clouds for all 10 years of observations and binned in 10 degree increments, ranging from pure ice ($LTF < 0.2$) to pure liquid ($LTF > 0.8$). Note that Hyperion sampling is nonuniform across histogram bins, and Section 1 quantifies uncertainty for different latitudes. The dataset clearly resolved key features apparent in records from other sensors like MODIS and CALIPSO (Hirakata et al., 2014; Hu et al., 2010). A band of ice clouds peaked in the Intertropical Convergence Zone (ITCZ) at approximately $5° - 10°$ latitude, and other seasonally-dependent maxima appeared at approximately 60° and -60°. The population associated with the LTF range from 0.6 to 0.8 was fairly large, and possibly included some pure water pixels that were misclassified due to estimation noise. The mixed phase clouds were most numerous in the middle and extreme high latitudes, and less abundant in the tropics. There was a general increase in the

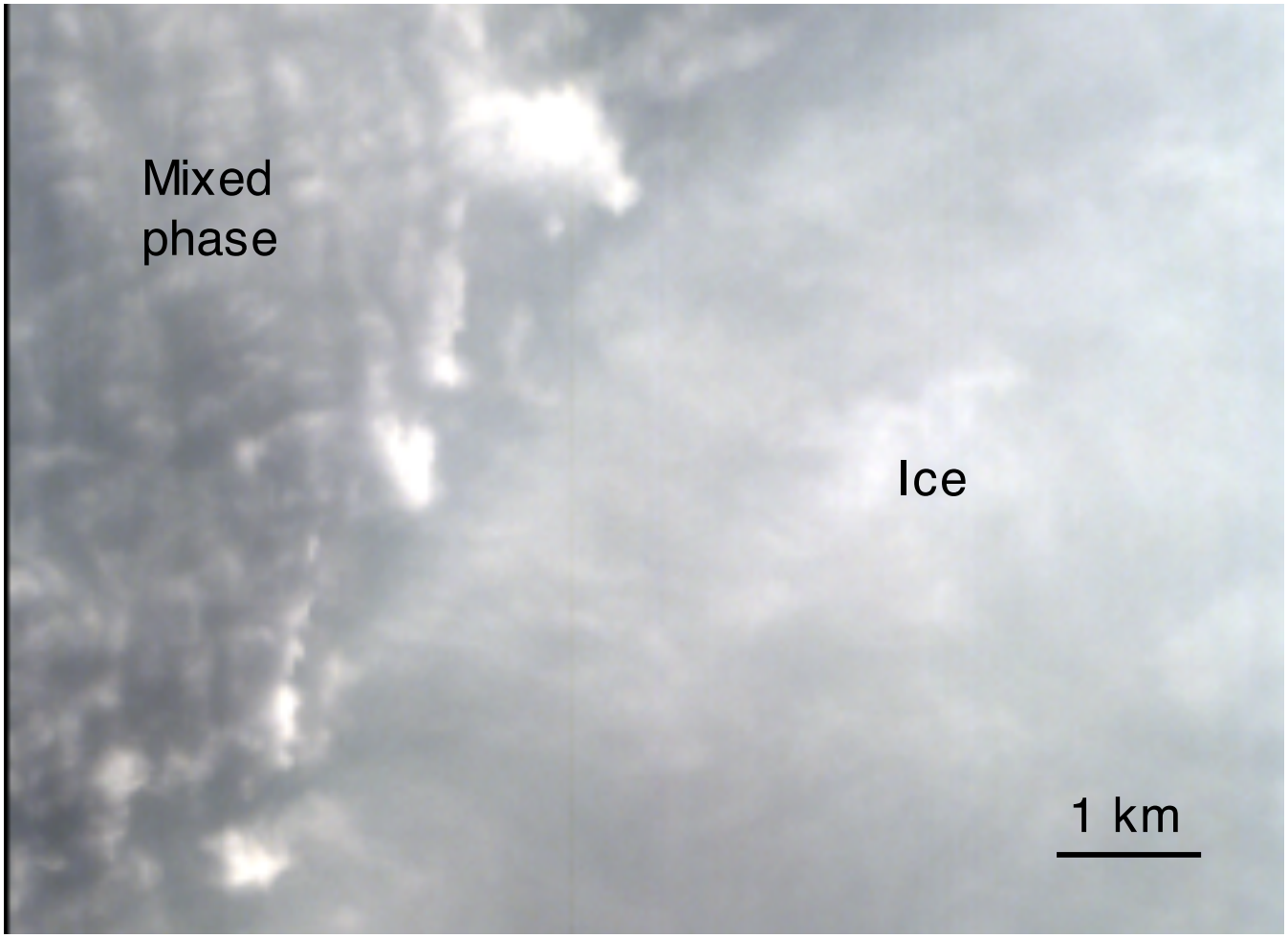

Figure 2. Typical fragment of a Top of Atmosphere Reflectance image, drawn from Hyperion product ID EO1H2221282005350110KF over Ross Island, Antarctica, and displayed in visible red, green, and blue channels.

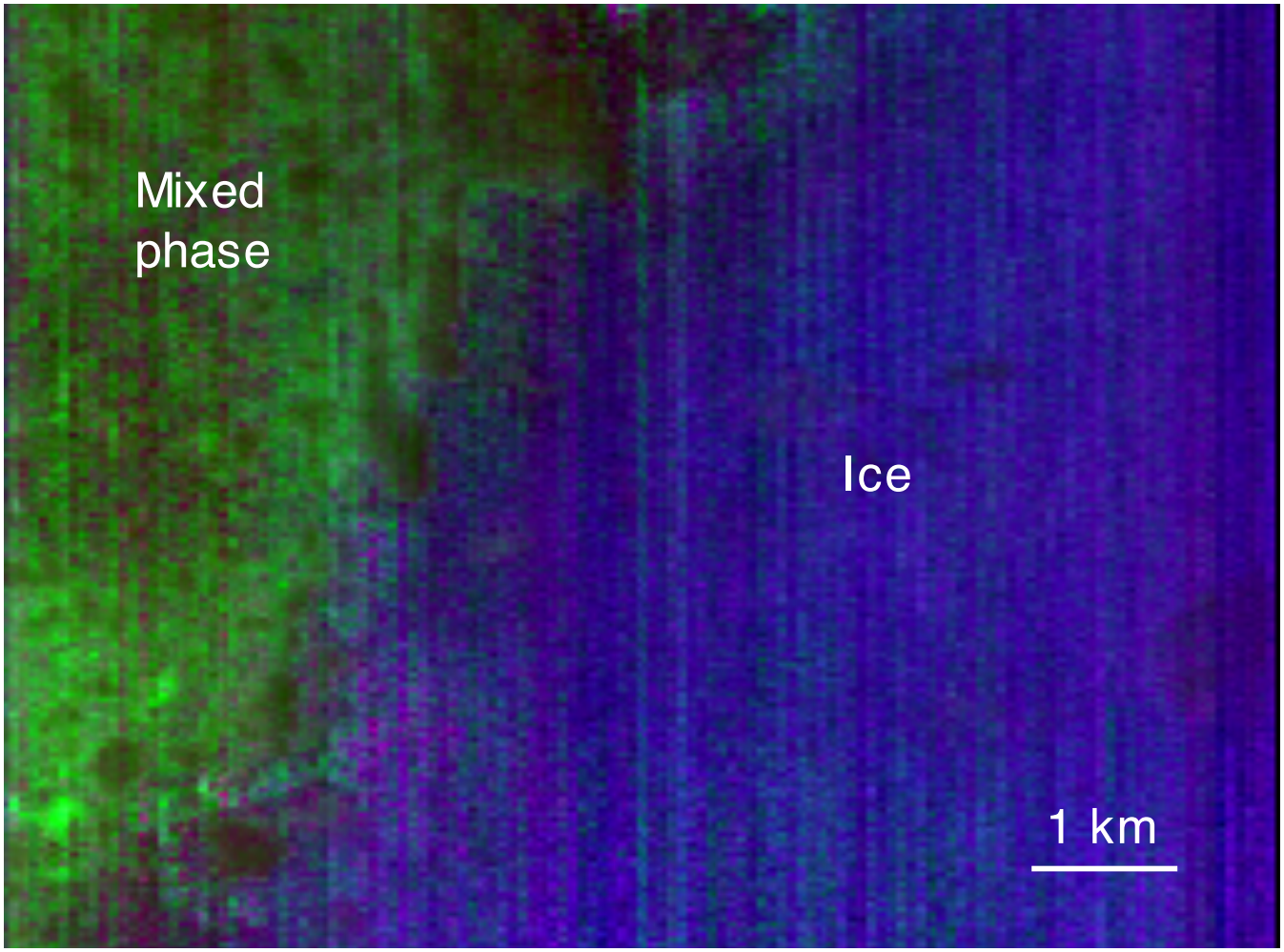

Figure 3. Thermodynamic phase map corresponding to Figure 2, with green indicating Equivalent Liquid Thickness, and blue indicating Equivalent Ice Thickness.

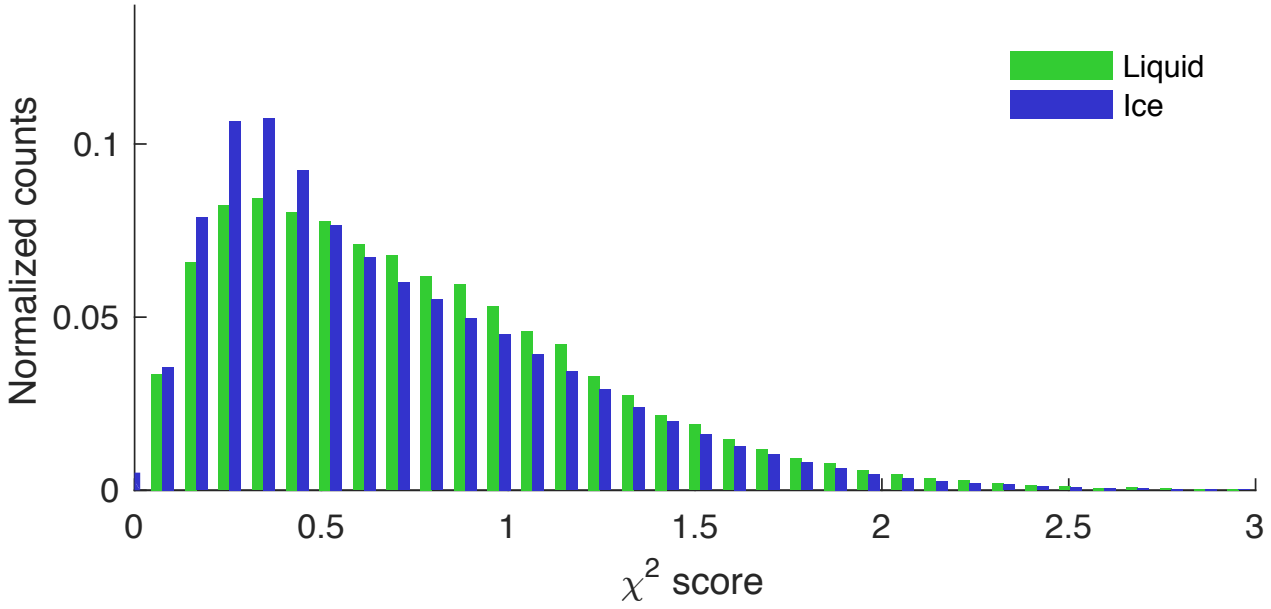

Figure 4. Normalized histogram of χ^2 spectrum fit scores for the entire scene of Figures 2 and 3. Low χ^2 values indicate good fits; $\chi^2 = 1$ is a conservative estimate of measurement noise.

occurrence of liquid clouds at middle and high latitudes, with a strong asymmetry near the solstice: a seasonal peak in liquid cloud coverage for the summer months was apparent in both hemispheres.

3 AIRS Comparison

3.1 Method

Next, we compared the resulting thermodynamic phase retrievals to a decadal dataset of cloud phase retrievals by the Atmospheric InfraRed Sounder (AIRS) instrument (Pagano et al., 2003). This was a dramatically different measurement obtained from thermal infrared spectra with a coarse 13.5 km footprint rather than reflected solar energy at fine spatial resolution. Kahn et al. (2014) detail the algorithm, and Jin and Nasiri (2014) validate it using pixel-scale comparisons with CALIPSO data. We filtered clouds using an AIRS sensitivity threshold (Effective Cloud Fraction, or ECF) of 0.1, and binned AIRS phases by latitude and season for direct comparison.

We anticipated several differences in the result. First, AIRS sampled uniformly over the Earth's surface while Hyperion imaged only during the day and favored land areas. We also expected differences in sensitivity; AIRS was far more sensitive to optically thin clouds, while the Hyperion analysis intentionally excluded thin clouds with a strict detection threshold. To

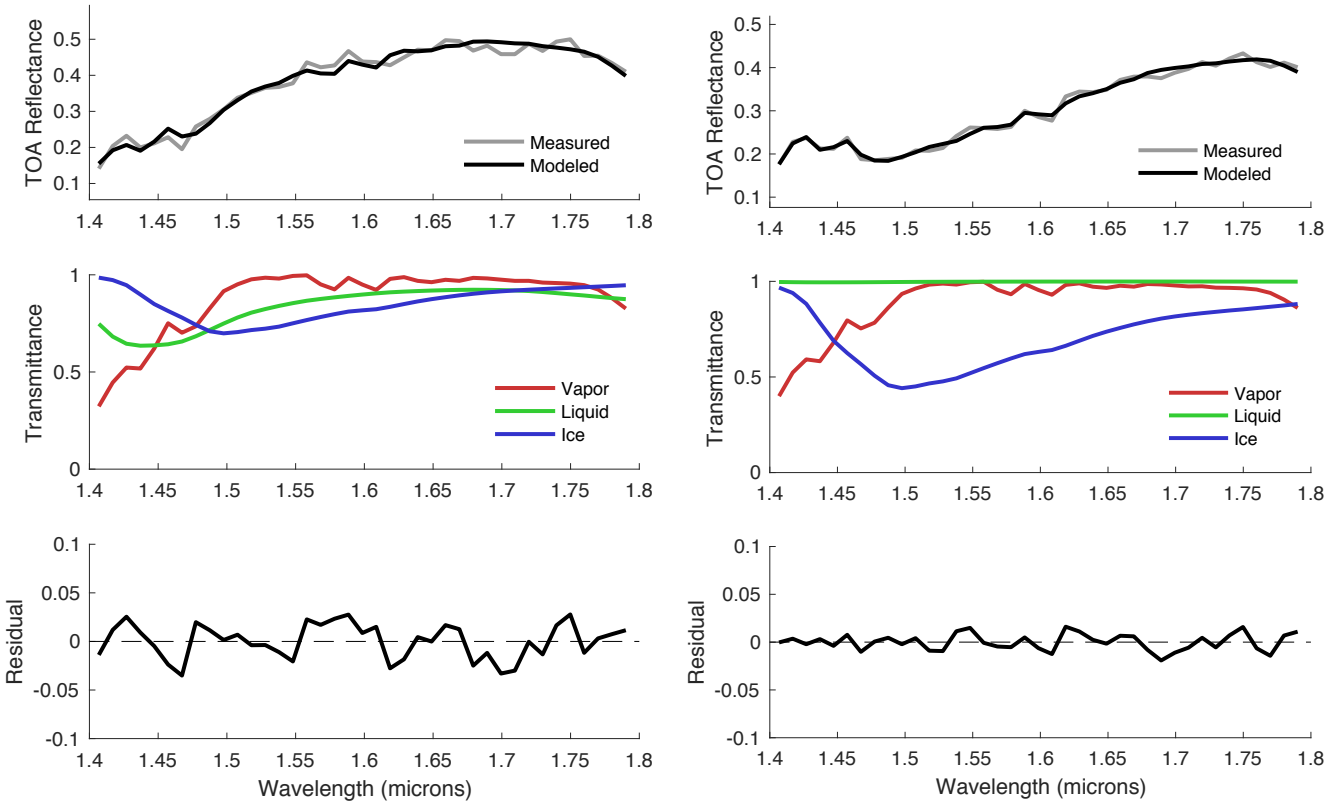

Figure 5. Spectrum fits from the region of mixed and ice cloud in Figure 2. Top row: model fit to spectrum. Middle row: Transmittances for each absorber. Bottom row: Residual error.

help account for these differences, we normalized the relative cloud abundances of both instruments. We designated a reference latitude band from -60 to 60 degrees (the area of densest Hyperion coverage) to have a mean occurrence of unity, and compared zonal changes relative to this standard.

Additionally, the AIRS algorithm classified ambiguous clouds as "unknown." This population likely contained mixed phase clouds but also a large fraction of supercooled liquid clouds due to the current AIRS phase algorithm. The liquid tests are based on warm liquid water indices of refraction rather than the supercooled liquid indices of refraction (Rowe et al., 2013). Following from results of Jin and Nasiri (2014), we reassigned some of the unknown clouds to form 10% of total ice and 60% of total liquid. For liquid clouds, we defined a corrected occurrence L' in each latitude bin and took just 40% of this to be from the original AIRS estimate L:

$$(1-0.6)\sum L' = \sum L \tag{10}$$

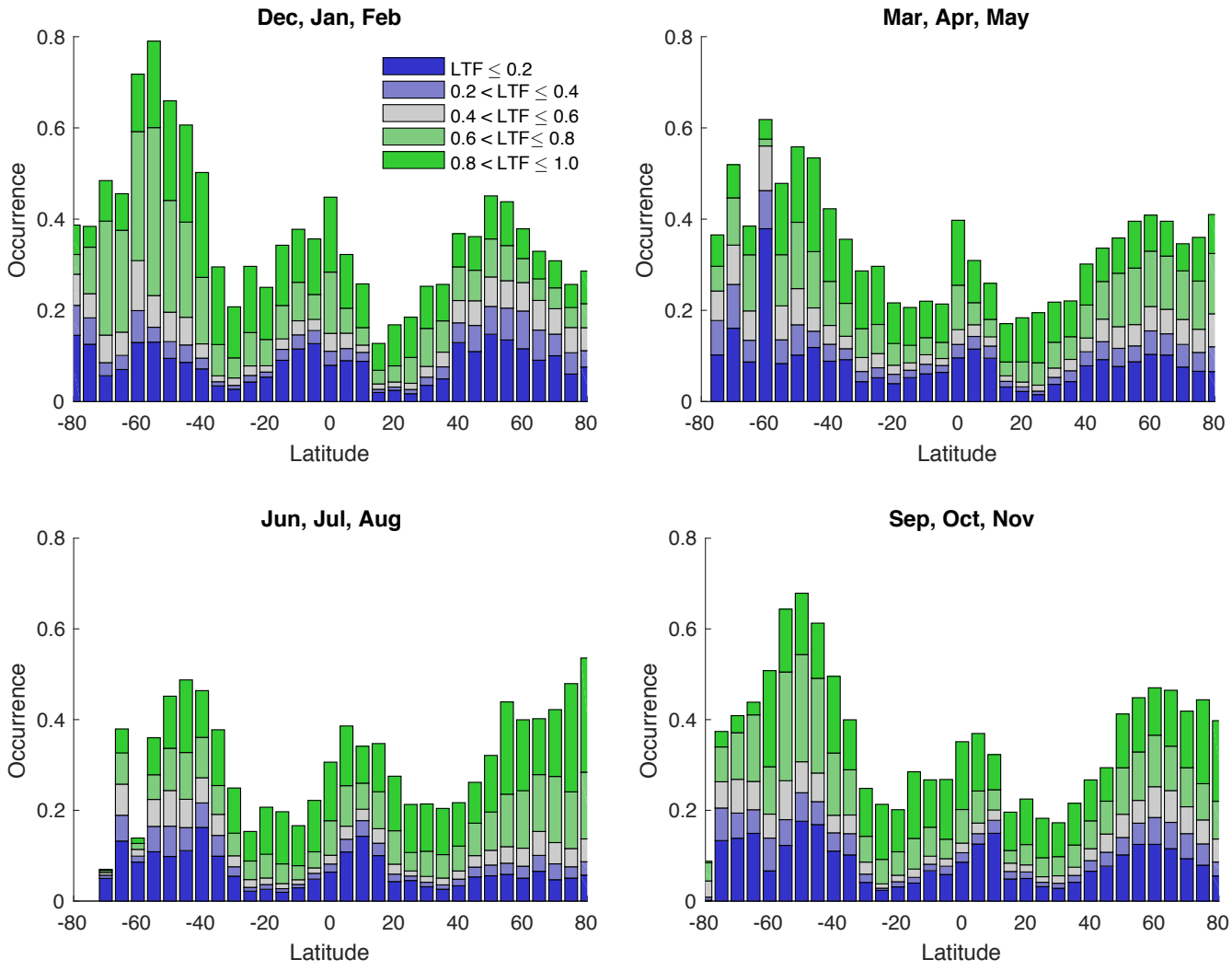

Figure 6. Zonal average cloud phase, partitioned by season, for all 10 years of observations. Sampling is nonuniform across latitudes; Section 1 quantifies uncertainty due to finite sample sizes.

where summations ran over all latitude bins. This allowed a unique correction for each latitude bin, given as the product of its unknown cloud fraction U and a global multiplicative coefficient α:

$$L' = L + \alpha U \tag{11}$$

Equations 10 and 11 gave:

$$\alpha = \frac{0.6 \sum L}{(1 - 0.6) \sum U} \tag{12}$$

We defined a similar relation for ice cloud with a missing fraction of 0.1 rather than 0.6. Together, the two adjustments (magnitude normalization and reassignment of unknown clouds) accounted for known biases which permitted a comparison of zonal gradients between the two instruments. While we expected some discrepancies due to differences in instruments and sampling, the comparison provided a useful check between two very different measurement techniques.

3.2 Results

Figures 7 and 8 show liquid and ice cloud phase spatial distributions across latitudes, partitioned by season, for all 10 years of observations binned in 10 degree increments. Error bars show 95% confidence intervals for the mean. Thin lines indicate decadal averages derived from the AIRS dataset. Distributions from the two instruments generally agreed - particularly for ice, to which AIRS was very sensitive. Tropical ice clouds showed the best agreement, as in prior studies comparing AIRS and CALIPSO (Jin and Nasiri, 2014). Differences may be related to the much stronger sensitivity to thin ice cloud, coarse spatial resolution and near global daily sampling, and ambiguity of the unknown category with respect to how many liquid, ice, and mixed phase clouds are contained within that categorization. Another potential contributor to the discrepancy is sensitivity to different altitudes in large or multilayer cloud systems. Multilayered clouds are abundant in tropical regions and and mid-latitude storm tracks. In cases where, for example, a translucent ice cloud overlays an optically thin liquid cloud, the two instruments would measure different thermodynamic phase.. For reference, we place a non-normalized comparison of the two instrument datasets in Appendix A.

4 Spatial Scale Analysis

4.1 Method

We next characterized spatial scaling properties of the thermodynamic phase maps. A variogram (Garrigues et al., 2006) estimated the expected squared differences between LTF at any two locations x and x' in the map as a function $v(d)$ of the *lag distance* d between the paired points. We used the classical estimator based on the sample variance at each lag distance:

$$v(d) = \frac{1}{2N(d)} \sum_{\|x - x'\| = d} (LTF(x) - LTF(x'))^2 \tag{13}$$

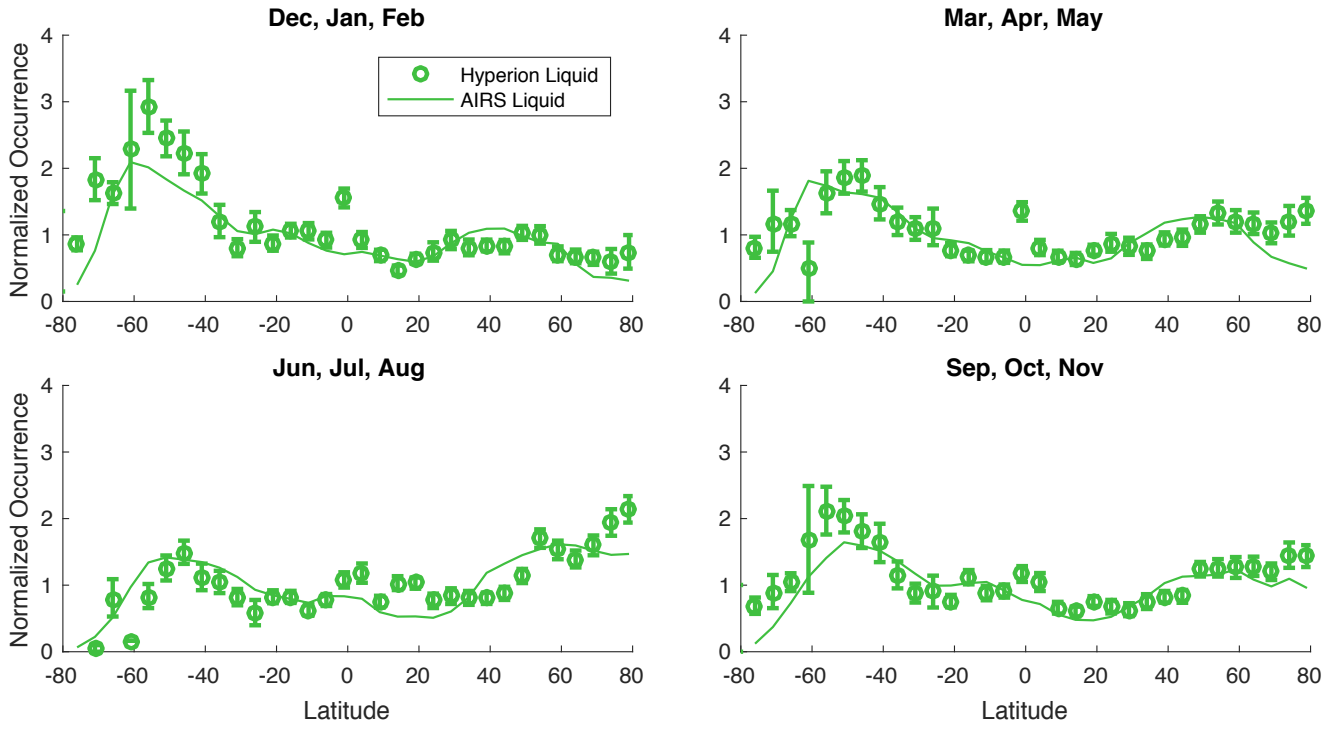

Figure 7. Comparison of Hyperion and AIRS cloud statistics, both normalized by the -60 to 60 degree latitude range to account for differences in cloud mask cutoff thresholds. The 10-year seasonal observations are shown separately for the liquid phase. Here Hyperion uses a hard classification, i.e., liquid thickness fractions less than 0.5 are considered ice clouds. Error bars show 95% confidence intervals calculated via nonparametric bootstrap estimation. The corresponding AIRS error bars would be far smaller due to the large number of samples, so we omit them for clarity.

where $N(d)$ was the number of such points in the sample. The näive formulation involving calculations of all point pairs would have required over 10^{10} squared differences per scene, an intractable number. Instead, we calculated the same quantity efficiently in the Fourier domain using the method of Marcotte (1996), with a spatial mask to ensure that only cloud pixels influenced the calculation.

We fit the resulting variogram with a power law of the form $v(d) = ad^b + c$, subsampling the data to achieve log-constant point density and optimizing free parameters with the Levenberg-Marquardt method. We considered the hypothesis that tropical clouds would have different spatial scaling from extra tropical clouds because of the dominant influence of convective versus baroclinic systems, respectively. To test this, we analyzed the scenes in three segments: a tropical band within 20 degrees of the equator, and extra tropical scenes poleward of 30 degrees North and South latitude. We used cloud fields larger than a 3 km cutoff, excluding variograms that evaluated to zero at shorter lag distances due to the lack of sufficient cloud pixels in the scene

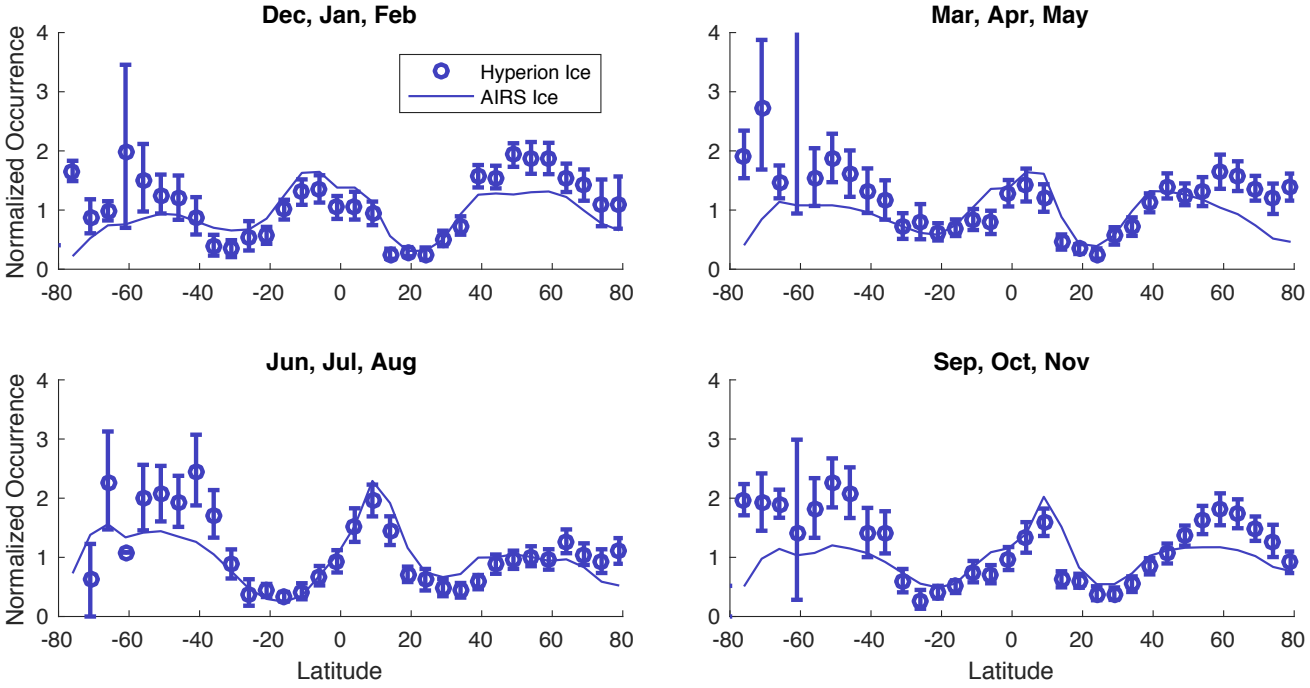

Figure 8. As in Figure 7, for the ice phase.

Subset	**Latitudes**	**a** ($\pm 95\%$)	**b** ($\pm 95\%$)	**c** ($\pm 95\%$)	$\mathbf{R}^2$
Northern Hemisphere	30 to 90	0.0058 (0.0052, 0.0064)	0.419 (0.399, 0.440)	0.0118 (0.0111, 0.0125)	0.995
Southern Hemisphere	-90 to -30	0.0046 (0.0040, 0.0051)	0.441 (0.414, 0.467)	0.0098 (0.0091, 0.0106)	0.995
Tropics	-20 to 20	0.0026 (0.0024, 0.0028)	0.619 (0.599, 0.639)	0.0056 (0.0052, 0.0060)	0.996

Table 2. Model parameters of tropical and extra tropical clouds, for $v = ad^b + c$, with confidence bounds and R squared coefficients.

4.2 Results

Finally, Figure 9 shows variograms over all years. Each datapoint corresponds to a lag distance for an aggregate variogram, so separation between the curves does not imply complete separation between populations. However, statistical analysis does reveal some significant differences in scaling properties. Table 2 shows the best-fitting coefficients for variance v in the liquid thickness fraction of tropical clouds, Northern hemisphere extra tropical clouds, and Southern hemisphere extra tropical clouds, as a function $v(d) = ad^b + c$ of distance d in kilometers. R^2 values indicate the power law is an excellent least-squares fit. Parenthesis indicate 95% confidence intervals on the parameters.

We report scaling exponents in terms of variance; these could be translated to other conventions using structure functions or the power spectrum domain. Differences between the power law exponents for tropical and extra tropical clouds were statistically significant. The extratropical scaling exponents of 0.42 and 0.44 are similar to, but slightly in excess of the classic Kolmogorov scaling of $1/3$ ($-5/3$ in the power spectral domain). The tropical scaling exponent of 0.62 is in excess of the classic Kolmogorov scaling of $1/3$ but is consistent with tropical cloud reflectance variability reported in Barker et al. (2017), and mid-tropospheric water vapor mixing ratio in the tropics from the AIRS instrument, e.g. Kahn et al. (2011). At finer spatial resolutions there is also evidence of scale breaks dependent on altitude. Consequently the scaling exponent is dependent on the length scale range calculated (Kahn et al., 2011).

The additive offset c defined the variance at zero distance, i.e. the irreducible measurement error for each spectrum. This implied a noise equivalent change in liquid water fraction of approximately 7.5% for tropical and 10-11% for extra tropical clouds. The addend and prefactor coefficients differed significantly between extra tropical and tropical clouds, and to a much smaller degree between the two extra tropical populations. All three populations were subject to Hyperion's biased sampling of ocean an land, and of instrument noise conditions dominated by solar zenith angles. We would not ascribe differences between Northern and Southern Hemispheres to cloud scaling since these discrepancies were small relative to their separation from the tropical model. In all cases, measurement error was a minor contributor to observed variance — most variability arose from spatially-correlated structure. Outside the tropics, measurement error accounted for just 25% of variance observed at GCM grid scales of 100 km. The remaining 75% was therefore attributable to spatial structure at subgrid scales.

5 Conclusions

This study reports the first global high spatial resolution survey of cloud thermodynamic phase. The Hyperion imaging spectrometer provides two novel contributions beyond prior records: first, highly diagnostic spectral features permits accurate discrimination of mixed phase clouds; and second, horizontal scales down to 30 m capture the characteristic spatial scaling relationships of the thermodynamic phase field. Aggregate seasonal and latitudinal changes of cloud thermodynamic phase generally corroborate observations by other sensing modalities, such as those of AIRS. Variogram analysis reveals a noise-equivalent measurement error of 7.5-11% in the liquid thickness fraction for different latitudinal zones. Spatial correlations follow a power law relationship with approximately 50% measurable variance determined at length scales of 6km. Significant spatial variability appears at scales far below the resolution of typical GCMs.

We note an important caveat to these results. The Hyperion datasets were spatially biased and strongly favored land mass over ocean. Insofar as the latitudinal distributions show asymmetries across northern and southern hemispheres, this may be related to the spatial distribution of land mass in the southern hemisphere midlatitude areas. Southern hemisphere observations were often acquired over islands, which would exhibit a more oceanic influence on cloud cover.

Despite this qualification, the Hyperion results generally corroborate existing global datasets where there is overlap, a useful complement to instruments measuring cloud phase with polarization or thermal emission. They provide a "first of a kind" observational record at sub-kilometer scales. These scales are critical for advancing subgrid parameterizations in climate GCMs, including the Wegener-Bergeron-Findeisen time scale of the growth of ice crystals (Tan and Storelvmo, 2016), and numerous

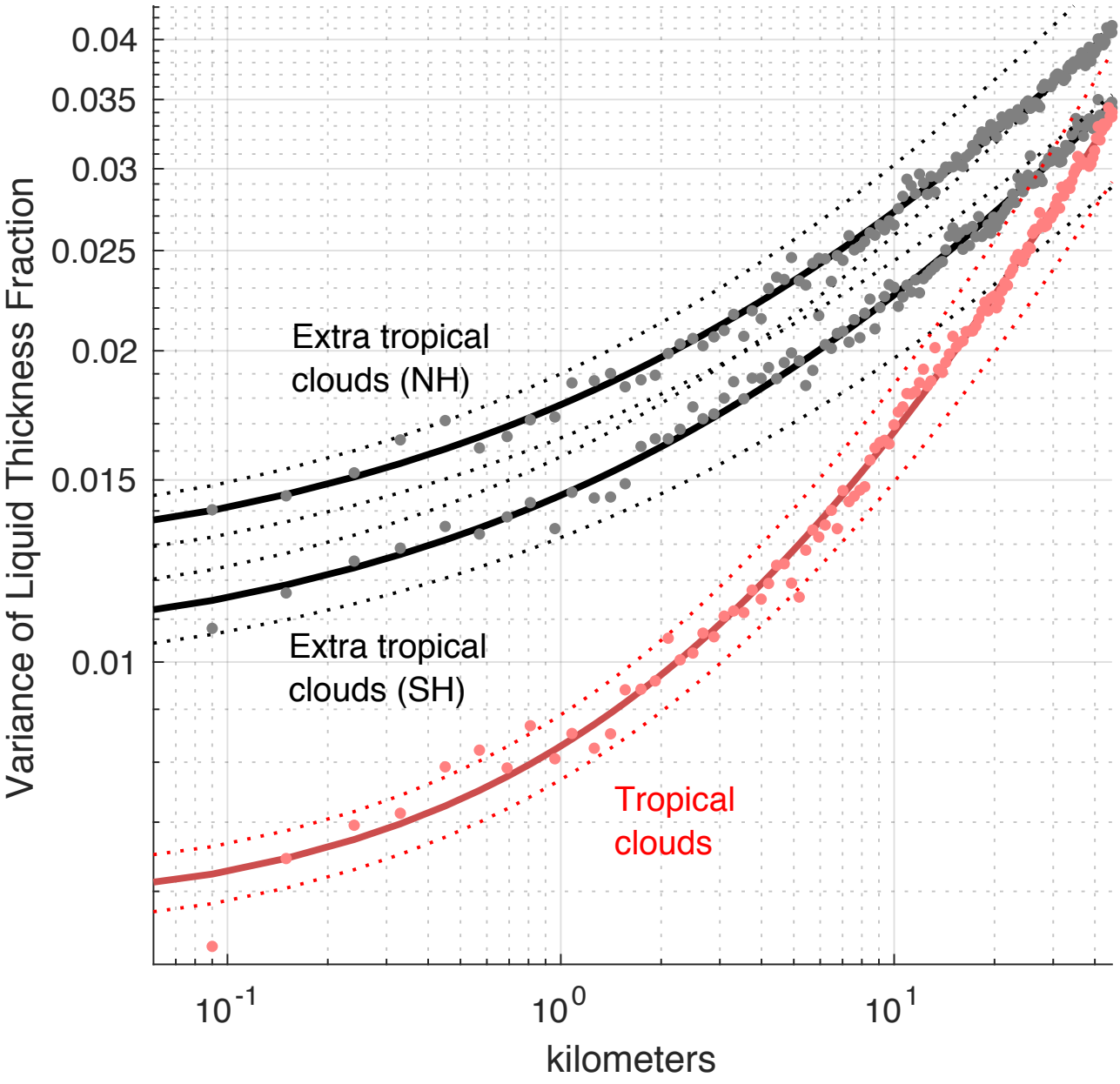

Figure 9. Variogram for all clouds in the Hyperion dataset, showing the variance as a function of separation between points. The curve suggests a power law relationship. Dotted lines indicate 95% confidence intervals on the coefficients.

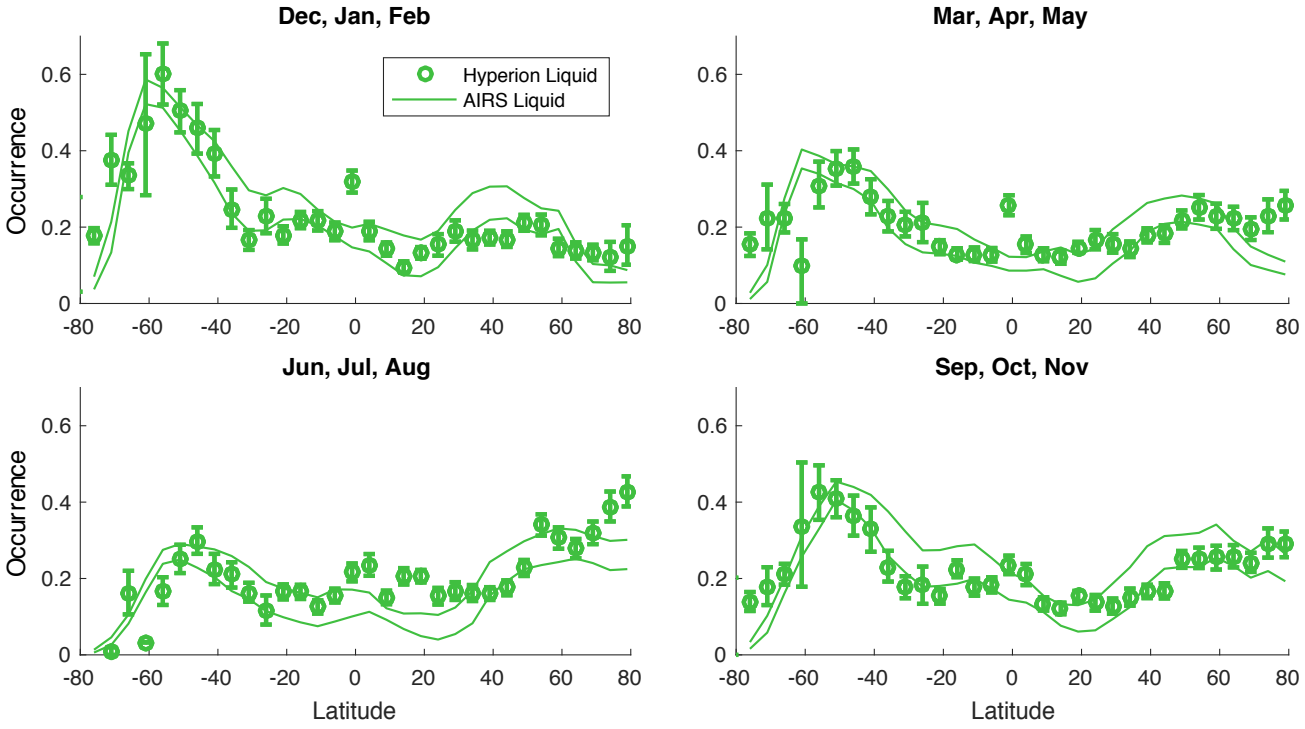

Figure A1. A comparison similar to Figure 7 without normalization across latitudes. Two thin lines represent AIRS retreivals using the standard effective cloud fraction (ECF) retrieval threshold of 0.5 and 0.1.

other temperature-dependent microphysical processes that control ice and liquid water partitioning (Ceppi et al., 2016), and further extending the observational record to a new extreme of spatial resolution.

Data availability. All Hyperion data can be downloaded via http://earthexplorer.usgs.gov

Appendix A: Non-Normalized AIRS Comparisons

Figures A1 and A2 show AIRS zonal averages for liquid and ice, respectively, with the correction of Jin and Nasiri (2014) but without normalization across latitudes. Thin lines indicate AIRS frequencies for ECF Thresholds of 0.1 and 0.5. This indicates the sensitivity of the AIRS result to this choice. The absolute abundance of the ice phase shows the largest disparities. This is expected, and accountable to very thin ice clouds to which AIRS is far more sensitive.

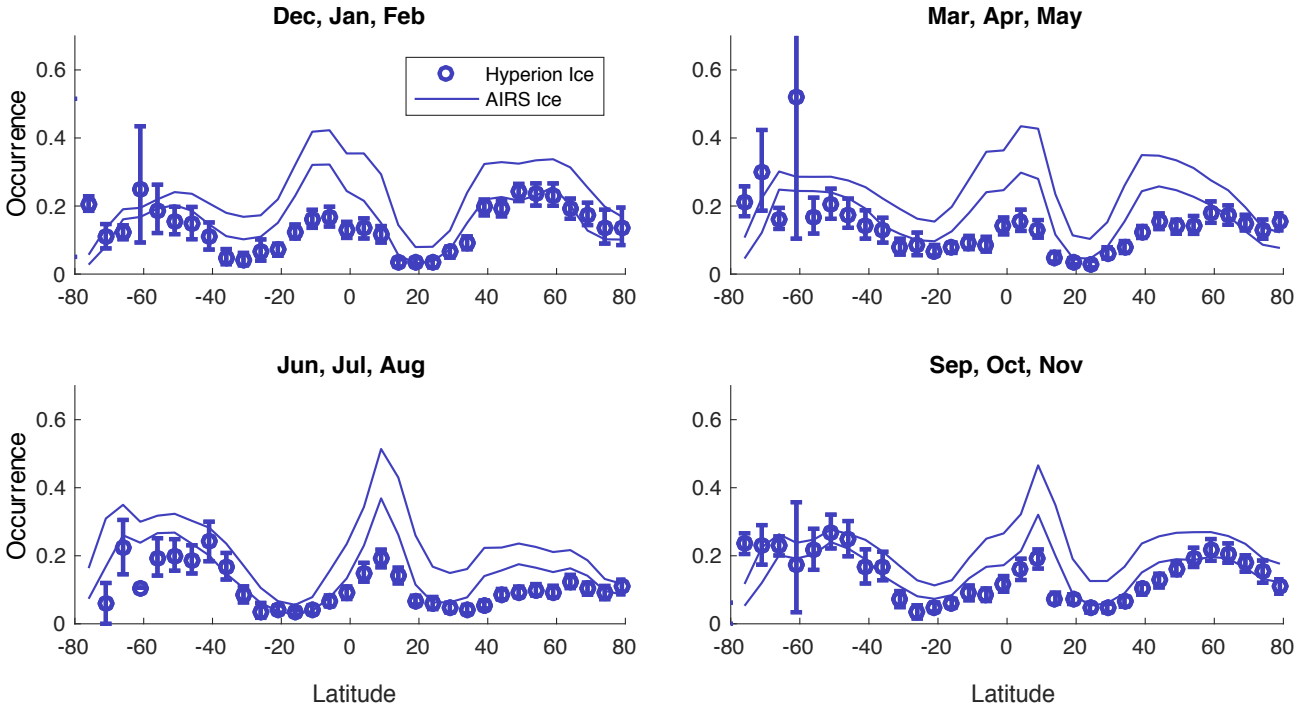

Figure A2. As in Figure A1, for the ice phase.

Competing interests. No author has competing interests.

Acknowledgements. This research was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. We thank the Hyperion Team of Goddard Space Flight Center for their assistance in acquiring and interpreting the data. We thank Bo-Cai Gao for radiative transfer methods used in part of the Hyperion analysis. US Government support acknowledged.

References

Barker, H. W., Qu, Z., Bélair, S., Leroyer, S., Milbrandt, J. A., and Vaillancourt, P. A.: Scaling properties of observed and simulated satellite visible radiances, Journal of Geophysical Research: Atmospheres, 122, 9413–9429, 2017.

Bender, F. A., Ramanathan, V., and Tselioudis, G.: Changes in extratropical storm track cloudiness 1983–2008: observational support for a poleward shift, Climate Dynamics, 38, 2037–2053, 2012.

Boardman, J. W. and Kruse, F. A.: Analysis of Imaging Spectrometer Data Using N-Dimensional Geometry and a Mixture-Tuned Matched Filtering Approach, Geoscience and Remote Sensing, IEEE Transactions on, 49, 4138–4152, 2011.

Brown, L. D. and Levine, M.: Variance estimation in nonparametric regression via the difference sequence method, The Annals of Statistics, 35, 2219–2232, 2007.

Ceppi, P. and Hartmann, D. L.: Connections between clouds, radiation, and midlatitude dynamics: A review, Current Climate Change Reports, 1, 94–102, 2015.

Ceppi, P., Hartmann, D. L., and Webb, M. J.: Mechanisms of the negative shortwave cloud feedback in middle to high latitudes, Journal of Climate, 29, 139–157, 2016.

Chylek, P., Robinson, S., Dubey, M., King, M., Fu, Q., and Clodius, W.: Comparison of near-infrared and thermal infrared cloud phase detections, Journal of Geophysical Research: Atmospheres (1984–2012), 111, 2006.

Dudhia, A.: Oxford University Reference Forward Model (RFM), http:// www.atm.ox.ac.uk/RFM/, 2014.

Ehrlich, A., Bierwirth, E., Wendisch, M., Gayet, J.-F., Mioche, G., Lampert, A., and Heintzenberg, J.: Cloud phase identification of Arctic boundary-layer clouds from airborne spectral reflection measurements: test of three approaches, Atmospheric Chemistry and Physics, 8, 7493–7505, https://doi.org/10.5194/acp-8-7493-2008, http://www.atmos-chem-phys.net/8/7493/2008/, 2008.

Eldering, A., O'Dell, C. W., Wennberg, P. O., Crisp, D., Gunson, M. R., Viatte, C., Avis, C., Braverman, A., Castano, R., Chang, A., Chapsky, L., Cheng, C., Connor, B., Dang, L., Doran, G., Fisher, B., Frankenberg, C., Fu, D., Granat, R., Hobbs, J., Lee, R. A. M., Mandrake, L., McDuffie, J., Miller, C. E., Myers, V., Natraj, V., O'Brien, D., Osterman, G. B., Oyafuso, F., Payne, V. H., Pollock, H. R., Polonsky, I., Roehl, C. M., Rosenberg, R., Schwandner, F., Smyth, M., Tang, V., Taylor, T. E., To, C., Wunch, D., , and Yoshimizu, J.: The Orbiting Carbon Observatory-2: first 18 months of science data products, Atmospheric Measurement Techniques, 10, 549–563, 2017.

Gao, B.-C. and Goetz, A. F.: Retrieval of equivalent water thickness and information related to biochemical components of vegetation canopies from AVIRIS data, Remote sensing of environment, 52, 155–162, 1995.

Garrigues, S., Allard, D., Baret, F., and Weiss, M.: Quantifying spatial heterogeneity at the landscape scale using variogram models, Remote Sensing of Environment, 103, 81–96, 2006.

Griffin, M. K., hua K. Burke, H., Mandl, D., and Miller, J.: Cloud Cover Detection Algorithm for EO-1 Hyperion Imagery, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery IX, Proceedings of SPIE, 5093, 2003.

Hirakata, M., Okamoto, H., Hagihara, Y., Hayasaka, T., and Oki, R.: Comparison of global and seasonal characteristics of cloud phase and horizontal ice plates derived from CALIPSO with MODIS and ECMWF, Journal of Atmospheric and Oceanic Technology, 31, 2114–2130, 2014.

Hu, Y., Rodier, S., Xu, K.-m., Sun, W., Huang, J., Lin, B., Zhai, P., and Josset, D.: Occurrence, liquid water content, and fraction of super-cooled water clouds from combined CALIOP/IIR/MODIS measurements, Journal of Geophysical Research: Atmospheres, 115, 2010.

Jin, H. and Nasiri, S. L.: Evaluation of AIRS Cloud-Thermodynamic-Phase Determination with CALIPSO, Journal of Applied Meteorology and Climatology, 53, 1012–2027, 2014.

Kahn, B., Irion, F., Dang, V., Manning, E., Nasiri, S., Naud, C., Blaisdell, J., Schreier, M., Yue, Q., Bowman, K., et al.: The Atmospheric Infrared Sounder version 6 cloud products, Atmospheric Chemistry and Physics, 14, 399, 2014.

Kahn, B. H., Teixeira, J., Fetzer, E., Gettelman, A., Hristova-Veleva, S., Huang, X., Kochanski, A., Köhler, M., Krueger, S., Wood, R., et al.: Temperature and water vapor variance scaling in global models: Comparisons to satellite and aircraft data, Journal of the Atmospheric Sciences, 68, 2156–2168, 2011.

Kokhanovsky, A.: Optical properties of terrestrial clouds, Earth-Science Reviews, 64, 189–241, 2004.

Kou, L., Labrie, D., and Chylek, P.: Refractive indices of water and ice in the 0.65-to 2.5-μm spectral range, Applied Optics, 32, 3531–3540, 1993.

Lawson, C. L. and Hanson, R. J.: Solving least squares problems, 161, 1974.

Marcotte, D.: Fast Variogram Computation with FFT, Computers and Geosciences, 22, 1175–1186, 1996.

Martins, J. V., Marshak, A., Remer, L. A., Rosenfeld, D., Kaufman, Y. J., Fernandez-Borda, R., Koren, I., Correia, A. L., Zubko, V., and Artaxo, P.: Remote sensing the vertical profile of cloud droplet effective radius, thermodynamic phase, and temperature, Atmospheric Chemistry and Physics, 11, 9485–9501, https://doi.org/10.5194/acp-11-9485-2011, http://www.atmos-chem-phys.net/11/9485/2011/, 2011.

Marvel, K., Zelinka, M., Klein, S. A., Bonfils, C., Caldwell, P., Doutriaux, C., Santer, B. D., and Taylor, K. E.: External influences on modeled and observed cloud trends, Journal of Climate, 28, 4820–4840, 2015.

McCoy, D. T., Hartmann, D. L., Zelinka, M. D., Ceppi, P., and Grosvenor, D. P.: Mixed-phase cloud physics and Southern Ocean cloud feedback in climate models, Journal of Geophysical Research: Atmospheres, 120, 9539–9554, 2015.

Mouroulis, P., Green, R. O., Van Gorp, B., Moore, L. B., Wilson, D. W., and Bender, H. A.: Landsat swath imaging spectrometer design, Optical Engineering, 55, 015 104–015 104, 2016.

Norris, J. R., Allen, R. J., Evan, A. T., Zelinka, M. D., O'Dell, C. W., and Klein, S. A.: Evidence for climate change in the satellite cloud record, Nature, 536, 72–75, 2016.

Pagano, T. S., Aumann, H. H., Hagan, D. E., and Overoye, K.: Prelaunch and in-flight radiometric calibration of the Atmospheric Infrared Sounder (AIRS), IEEE transactions on geoscience and remote sensing, 41, 265–273, 2003.

Pilewskie, P. and Twomey, S.: Cloud phase discrimination by reflectance measurements near 1.6 and 2.2 μm, Journal of the atmospheric sciences, 44, 3419–3420, 1987.

Rothman, L., Gordon, I., Babikov, Y., Barbe, A., Chris Benner, D., Bernath, P., Birk, M., Bizzocchi, L., Boudon, V., Brown, L., et al.: The HITRAN2012 molecular spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 130, 4–50, 2013.

Rowe, P. M., Neshyba, S., and Walden, V.: Radiative consequences of low-temperature infrared refractive indices for supercooled water clouds, Atmospheric Chemistry and Physics, 13, 11 925–11 933, 2013.

Tan, I. and Storelvmo, T.: Sensitivity Study on the Influence of Cloud Microphysical Parameters on Mixed-Phase Cloud Thermodynamic Phase Partitioning in CAM5, Journal of the Atmospheric Sciences, 73, 709–727, 2016.

Tan, I., Storelvmo, T., and Zelinka, M. D.: Observational constraints on mixed-phase clouds imply higher climate sensitivity, Science, 352, 224–227, 2016.

Thompson, D. R., Gao, B.-C., Green, R. O., Roberts, D. A., Dennison, P. E., and Lundeen, S.: Atmospheric correction for global mapping spectroscopy: ATREM advances for the HyspIRI preparatory campaign, Remote Sensing of Environment, pp. –, https://doi.org/http://dx.doi.org/10.1016/j.rse.2015.02.010, http://www.sciencedirect.com/science/article/pii/S0034425715000607, 2015.

Thompson, D. R., McCubbin, I., Gao., B.-C., Green, R. O., Matthews, A. A., Mei, F., Meyer, K., Platnick, S., Schmid, B., Tomlinson, J., , and Wilcox, E.: Measuring cloud thermodynamic phase with shortwave infrared imaging spectroscopy, Journal of Geophysical Research Atmospheres, 121, 2016.

Von Neumann, J.: Distribution of the ratio of the mean square successive difference to the variance, Ann. Math. Statist., 12, 367–395, 1941.

Wasserman, L.: All of Nonparametric Statistics, Springer-Verlag, New York, 2006.