RESPONSE TO REVIEWER 1

Thank you for your comments.

1. In the section describing IMS-dataset you might want to explain a bit more in detail what instruments the dataset is based on.

IMS uses an often-changing list of instruments and models to build its dataset. We have added some examples of instruments that are used in Section 2.1.1.

Line 100: “The maps are produced by a trained analyst using visible imagery from a collection of geostationary (e.g. GOES, MeteoSat) and polar orbiting (e.g. AVHRR, MODIS, SAR) satellite instruments, with additional information from microwave sensors (e.g. DMSP, AMSR, AMSU), surface observations (e.g. SNOTEL), and models (e.g. SNODAS) (Helfrich et al., 2007).”

2. There is a fractional snow extent product from Globsnow/Sen3app projects that might be also worth a look and included in the comparison. For 2015 it is based on VIIRS (Suomi-NPP) data. The data and information are available here: http://www.globsnow.info/index.php?page=SE or here: http://sen3app.fmi.fi/index.php?page=Fractional_Snow_Cover_Extent_- _NH&style=main

We have looked at the fractional snow extent product from Globsnow/Sen3app as suggested, and have decided to exclude it from this work. This product does not provide snow cover information when clouds are present in the VIIRS observations. As a result, there is no information on snow cover for approximately a third of the TEMPO domain in 2015. Therefore, the product is not appropriate for the study performed here.

3. In the conclusion you write: “However, the lack of confidence in snow identification has previously led many retrieval procedures to omit observations over snow. Increasing this confidence such that these observations could be included would not only improve spatial and temporal sampling, but also allow the inclusion of observations with higher quality information on the lower troposphere.” It would be useful to actually demonstrate this with an example or case study, perhaps based on OMI data. I mean, showing one OMI scene/orbit of NO2 retrievals, where the added value of this improved snow information would be visible. For example, an OMI orbit with snow-cover that was filtered out or somehow incorrectly flagged and would be improved using a more accurate knowledge of the snow cover (with the right AMFs and profiles) in the NO2 retrieval.

Thank you for this suggestion. We have included a figure (Figure 6) that shows how including observations over snow improves sampling and increases AMFs. This is explained in the text on Line 280 as follows:

“We next examine the effect on both spatial sampling and sensitivity to the lower troposphere of a retrieval data set if observations with surface snow are included rather than omitted. We use IMS to identify the presence of snow for OMI observations over North America in January 2015. We then use LIDORT to calculate AMFs for these observations using the corresponding snow-free (Sun et al., 2017) or snow-covered (O'Byrne et al., 2010) surface reflectance, and examine the results of either including or omitting snow-covered scenes. Figure 6 shows that including snow-covered scenes results in a significant (factor of 2.1) increase in observation frequency, particularly in the northern US and Canada. Additionally, including snow-covered scenes increases the average AMF by a factor of 2.7 in regions with occasional snow cover. The increase in AMF demonstrates that including snow-covered scenes increases the quality of information about the tropospheric NO_2 column by increasing the observation sensitivity to tropospheric NO_2."

4. Could you comment on how the increased sensitivity in the PBL might affect NO2 retrievals at relatively higher latitudes (where snow is very often present)? For example, how would those scattering weight profiles in Fig. 2 look like for higher SZA/or a different latitude? It might be less important for TEMPO but it is relevant for OMI/TROPOMI missions to improve retrieval at high latitudes in autumn-winter.

We have added a scattering weight profile for a high latitude location in Figure 2.

5. There is this paper by Vasilkov et al. about BRDF and OMI retrievals you might need to mention/discuss in your paper: Vasilkov, A., Qin, W., Krotkov, N., Lamsal, L., Spurr, R., Haffner, D., Joiner, J., Yang, E.-S., and Marchenko, S.: Accounting for the effects of surface BRDF on satellite cloud and trace-gas retrievals: a new approach based on geometry-dependent Lambertian equivalent reflectivity applied to OMI algorithms, Atmos. Meas. Tech., 10, 333-349, https://doi.org/10.5194/amt-10-333-2017, 2017.

We have added a mention to this paper in the introduction (Line 59):

"Correspondingly, surface snow may be mistaken for cloud, leading to errors in cloud fraction and pressure estimates used in trace gas retrievals (Lin et al., 2015; O'Byrne et al., 2010; Vasilkov et al., 2017)."

and in the conclusion, as follows (Line 316):

"This could potentially include Bidirectional Reflectance Distribution Functions (BRDF) that describe reflection at different viewing angles, as this effect has been shown to have significant impact on retrieved NO_2 columns (Vasilkov et al., 2017)"

RESPONSE TO REVEIWER 2

Thank you for your comments.

1. The assessment of different snow cover data sets is carried out for the entire year of 2015. This approach of using the full year data may cause biases in the metrics. The authors admit "All data sets have high accuracy numbers, owing to a high number of true negatives during the summer months" (Line 220). I think that the assessment of the snow cover data sets should be done on a seasonal basis and the metrics for different seasons should be compared. It would be particularly interesting to assess the snow data sets for spring when melting snow occurs.

We have included a table in the Appendix that gives evaluations of the snow data sets by season, and now include the following text on Line 245:

"Data sets were also evaluated by season with similar results (Appendix Table A1). All data sets have weaker performance metrics during the spring melt season, which has been observed in past evaluations (Frei et al., 2012). IMS has the highest *F* score in winter and autumn but is slightly outperformed by MAIAC in spring."

However, keeping in mind that the goal of this work is to evaluate data sets for informing retrieval algorithms, and as most retrieval algorithms would likely choose a single data set to provide snow information throughout the year, we continue to focus on the full year data.

2. In my opinion, results of the RT simulations shown in Fig. 2 and 3 do not provide new significant information. Effects of surface reflectance on trace gas retrievals have been studied theoretically (see O'Bryne et al., JGR, 2010; Lin et al., ACP, 2015; Vasilkov et al., 2017 and references there).

We respectfully contend that Figures 2-3 do provide important information here. They illustrate how changes in snow cover affect the observation sensitivity to NO_2. Indeed Reviewer 1 expressed interest in Figure 2.

Figure 2 of the manuscript (showing the scattering weights for a single solar zenith angle and a single NO2 profile) is not conclusive because the NO2 sensitivity to surface reflectance substantially depends on tropospheric NO2 profiles (see Fig. 13 in Vasilkov et al., AMT, 2017).

It is true that the *column NO_2 sensitivity* depends on tropospheric NO_2 profiles. However, the scattering weights in Figure 2 represent the *sensitivity of backscattered radiation* to surface reflectance, which is independent of NO_2 profile. We have taken care to clarify this in the text (Line 201):

"Figure 2 shows the sensitivity of backscattered radiation (scattering weights) over snow-covered and snow-free surfaces …"

Figure 3 compares AMFs for snow-covered and snow-free conditions for January 2013. The snow-free conditions are absolutely unrealistic for January. That is why I doubt that useful information can be derived from this comparison.

We clarified that the figure is for the observation geometry of January. The Figure 3 “snow-free conditions” plot shows AMF values in the case that snow is not present during a given observation. It is not meant to suggest that snow is never present in January in North America. As snow-covered scenes are often omitted in retrieval algorithms, the resulting data sets are essentially “snow-free”, and thus a snow-free map of AMF does provide important context.

I think that the text and figures related to the RT simulations can be removed without the loss of significant material. To some extent, this is supported by the title of the manuscript because the RT simulations are not mentioned in the title.

We have strengthened the material covering snow and AMFs by including Figure 6, which shows how including snow-covered scenes improves the quantity and quality of retrieval data sets. We have changed the title to reflect this as well. Together with Figures 2-3 we feel that this is new, significant information.

Specific comments Line 24. The quantity “F” is not defined here.

We have removed the mention of *F* here. It is defined later in the abstract.

Line 52. It is worthwhile to mention that uncertainties in surface reflectance also lead to uncertainties in the cloud fraction and pressure retrievals which affect the NO2 retrievals (Vasilkov et al., AMT, 2017).

We now mention this effect in the introduction (Line 59):

“Correspondingly, surface snow may be mistaken for cloud, leading to errors in cloud fraction and pressure estimates used in trace gas retrievals (O'Byrne et al., 2010; Lin et al., 2015; Vasilkov et al., 2017).”

Line 162. Indeed, snow reflectivity is almost spectrally independent in UV/Vis. However, the maps in Fig. 1 include snow-free regions. For such regions, ground reflectivity does depend on wavelength, so reflectivity at 354 nm may not be used for 440 nm.

The snow reflectivity (for 354 nm) is only used when snow is present. Snow-free regions use the MODIS CMG Gap-Filled Snow-Free Products at 470 nm, which are at a wavelength closer to the 440 nm used in the AMF calculation. We have clarified this in the text and in Figure 1.

Line 174. Please clarify “the most reliable source is used”.

As stated, the GHCN-D data set includes information from multiple sources. GCHN-D provides a priority ranking of these sources. We have added a citation to this line which provides additional information.

Line 185. Please explain why the F score is most relevant for TEMPO.

This is now clarified in the text (Line 192) as follows:

“The *F* score balances recall (which accounts for false negatives) and precision (which accounts for false positives) to measure correct classification of snow without the influence of frequent snow-free periods, and therefore is the metric which is most relevant for TEMPO”

Line 190. Where does the OMI cloud fraction come from? How is the cloud fraction determined for snow-covered and partially snow-covered scenes?

We no longer use the OMI cloud fraction in this work. From line 199:

"We assume cloud-free conditions in the AMF calculations, as the impact of surface reflectance on retrieved cloud fractions is beyond the scope of this paper."

Line 235. Is it correct that the MODIS products perform better at coarser resolution? Table 1 shows F=0.46 and 0.54 for the 4 km resolution while Table 2A shows F=0.45 and 0.53 for the 25 km resolution.

Yes, MODIS products do perform better when regridded to 4km than at their native resolution of 0.05°, where F=0.37 and 0.43. However as pointed out by the reviewer, the benefit of regridding does not continue to improve if the resolution is further decreased. This has been clarified in the text (Line 250):

"…MODIS Aqua and Terra products perform better when regridded from their native 0.05° resolution to a 4 km resolution as it reduces the number of grid boxes missing observations due to cloud…"

Reference to McLinden et al., ACP, 2014 is missing.

This has been fixed.

Figure 1. The caption states "reflectivity at visible wavelengths". The 354 nm wavelength (used for the upper panel) is not a visible wavelength. The lower panel is not informative because the color scale is not appropriate for it.

The figure caption now specifies "UV-Visible" instead of only "visible" wavelengths. We have also changed the colour scale.

Figure 2. The corresponding NO2 profiles should be shown. Surface reflectivities should be specified. What is the viewing zenith angle of observations?

Surface reflectivities and zenith angles are now included in Figure 2. We have edited the text at Line 201 to better distinguish between the sensitivity of backscattered radiation to lower troposphere NO_2 (i.e. scattering weights) and the sensitivity of the NO_2 column to lower troposphere NO_2 (i.e. AMFs). Figure 2 focuses on how the scattering weights themselves (which do not depend on the NO_2 profile) are affected by reflectivity, and thus we do not include the corresponding NO_2 profiles for the sake of clarity.

"Figure 2 shows the sensitivity of backscattered radiation (scattering weights) over snow-covered and snow-free surfaces for two locations … This shows that satellite observed backscattered radiation is up to five times as sensitive to NO_2 in the boundary layer in the presence of snow, due to the increased absorption by NO_2 in the lower troposphere when the surface reflects more sunlight."

Appendix. Please explain why some numbers for the CMC and NISE data sets are slightly different in Tables A1 and A2. The spatial resolution of the data sets is same for both tables.

Thank you for noticing this. There were some errors in the Appendix tables that have been corrected. In Table A3 (previously A2), all products were regridded to a common 25km resolution. For NISE, this is slightly different than its native 25km grid, hence a small difference in its *F* score (0.51 to 0.52).

Assessing snow extent data sets over North America to inform and improve trace gas retrievals from solar backscatter

Matthew J. Cooper[1], Randall V. Martin[1,2], Alexei I. Lyapustin[3], and Chris A. McLinden[4]

1. Department of Physics and Atmospheric Science, Dalhousie University, Halifax, Nova Scotia, Canada.

2. Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts, USA

3. NASA Goddard Space Flight Center, Greenbelt, MD, USA

4. Air Quality Research Division, Environment and Climate Change Canada, Toronto, Ontario, Canada

Abstract

Accurate representation of surface reflectivity is essential to tropospheric trace gas retrievals from solar backscatter observations. Surface snow cover presents a significant challenge due to its variability and thus snow-covered scenes are often omitted from retrieval data sets; however, the high reflectance of snow is advantageous for trace gas retrievals. We first examine the implications of surface snow on retrievals from the upcoming TEMPO geostationary instrument for North America. We use a radiative transfer model to examine how an increase in surface reflectivity due to snow cover changes the sensitivity of satellite retrievals to NO_2 in the lower troposphere. We find that a substantial fraction (>50%) of the TEMPO field of regard can be snow covered in January, and that the average sensitivity to the tropospheric NO_2 column substantially increases (doubles) when the surface is snow covered.

We then evaluate seven existing satellite-derived or reanalysis snow extent products against ground station observations over North America to assess their capability of informing surface conditions for TEMPO retrievals. The Interactive Multisensor Snow and Ice Mapping System (IMS) had the best agreement with ground observations (accuracy=93%, precision=87%, recall=83%~~, *F*=85%~~). Multiangle Implementation of Atmospheric Correction (MAIAC) retrievals of MODIS observed radiances had high precision (90% for Aqua and Terra), but underestimated the presence of snow (recall=74% for Aqua, 75% for Terra). MAIAC generally outperforms the standard MODIS products (precision=51%, recall=43% for Aqua;

precision=69%, recall=45% for Terra). The Near-real-time Ice and Snow Extent (NISE) product had good precision (83%) but missed a significant number of snow covered pixels (recall=45%). The Canadian Meteorological Centre (CMC) Daily Snow Depth Analysis Data set had strong performance metrics (accuracy=91%, precision=79%, recall=82%~~, *F*=81%~~). We use the *F* score, which balances precision and recall, to determine overall product performance (*F* = 85%, 82(82)%, 81%, 58%, 46(54)% for IMS, MAIAC Aqua(Terra), CMC, NISE, MODIS Aqua(Terra) respectively) for providing snow cover information for TEMPO retrievals from solar backscatter observations. We find that using IMS to identify snow cover and enable inclusion of snow-covered scenes across North America in January can increase both the number of observations by a factor of 2.1 and the average sensitivity to the tropospheric NO_2 column by a factor of 2.7.

1. Introduction

Satellite observations of solar backscatter are widely used as a source of information on atmospheric trace gases (Richter and Wagner, 2011). These observations have provided valuable information on vertical column densities of O_3, NO_2, SO_2, CO, HCHO, CH_4 and other important trace gases in the troposphere (Fishman et al., 2008). Satellite observations of trace gases have been used to assess air quality (Duncan et al., 2014; Martin, 2008) and to gain insight into atmospheric processes including emissions (Streets et al., 2013), lifetimes (Beirle et al., 2011; Fioletov et al., 2015; de Foy et al., 2015; Valin et al., 2013), and deposition (Geddes and Martin, 2017; Nowlan et al., 2014). The utility of these observations is dependent on their quality, and thus ensuring retrieval accuracy is essential.

Previous studies have found that retrieved NO_2 vertical column densities are highly sensitive to errors in assumed surface reflectance (Boersma et al., 2004; Lamsal et al., 2017; Martin et al., 2002). Much of this error sensitivity results from observation sensitivity to trace gases in the lower troposphere. The observation sensitivity is accounted for in the air mass factor (AMF) conversion of observed line-of-sight “slant columns” to vertical column densities. Uncertainties in surface reflectance are a significant contributor to AMF uncertainty.

Existing reflectivity climatologies (e.g. Kleipool et al., 2008; Koelemeijer et al., 2003; Liang et al., 2002; Herman and Celarier, 1997) do not represent snow cover well, since the statistical methods to exclude reflective clouds from the climatologies also exclude variable snow cover; Correspondingly, surface snow may be mistaken for cloud, leading to errors in cloud fraction and pressure estimates used in trace gas retrievals (Lin et al., 2015; O'Byrne et al., 2010; Vasilkov et al., 2017). Therefore, snow cover is particularly challenging to retrievals. Misrepresenting surface snow cover can lead to large errors (20-50%) in retrieved NO_2 columns over broad regions with seasonal snow cover (O'Byrne et al., 2010). For this reason, observations over snow are often omitted to avoid potential errors. This limits the ability of satellite retrieved data sets to offer adequate temporal and spatial sampling in winter months. Additionally, over highly reflective surfaces such as snow observation sensitivity to the lower troposphere is larger and has less dependence on *a priori* NO_2 profiles ~~over highly reflective surfaces such as snow~~ (Lorente et al., 2017; O'Byrne et al., 2010); Thus, omitting snow-covered scenes means omitting the observations with the greatest sensitivity to the lower troposphere. This could be remedied by using a product that would allow for snow cover identification to be done with confidence.

Several data products provide information on snow extent using surface station observations, satellite observed radiances, or visible imagery. Previous evaluations have found it difficult to determine which of these products is definitively the best, partly due to differences in resolution. Most products are more consistent during the winter months when persistent, deep snow is present (Frei et al., 2012; Frei and Lee, 2010). However, disagreements are common during accumulation and melting seasons, over mountains, and under forest canopies. These evaluations have largely focused on local or regional snow cover, or included only cloud-free observations.

The upcoming geostationary Tropospheric Emissions: Monitoring of Pollution (TEMPO) satellite instrument will provide hourly observations of air quality relevant trace gases over North America at an unprecedented spatial and temporal resolution (Zoogman et al., 2017). As is the case for all nadir satellite retrievals, the quality of these observations will depend on the accuracy of the surface reflectance used in the retrieval. As a significant portion of the observed domain experiences snow cover, an accurate representation of snow cover is needed. Current plans to deal with snow cover for TEMPO are to rely on external observations.

In this work, we examine the importance of accurate snow identification by using a radiative transport model to evaluate how the vertical sensitivity of a satellite retrieval is impacted by surface reflectance. We then assess seven snow extent products that are expected to continue to be operational during the TEMPO mission using in situ observations across North America with the intent of determining which product is best suited for providing snow cover information for TEMPO and other future satellite retrievals. Finally, we combine radiative transfer model results with a snow extent product to show how including snow-covered scenes improves both the quantity and quality of information in a retrieval data set.

2. Data and Algorithms

2.1. Gridded Snow Products

2.1.1. IMS

One of the most widely used sources of snow extent data is the Interactive Multisensor Snow and Ice Mapping System (IMS). IMS provides daily, near-real-time maps of snow and sea ice cover in the Northern Hemisphere at 4km resolution (Helfrich et al., 2007). The maps are produced by a trained analyst using visible imagery from a collection of geostationary (e.g. GOES, MeteoSat) and polar orbiting (e.g. AVHRR, MODIS, SAR) satellite instruments, with additional information from microwave sensors (e.g. DMSP, AMSR, AMSU), surface observations (e.g. SNOTEL), and models (e.g. SNODAS) (Helfrich et al., 2007). By using multiple sources of information with different spatial resolution and temporal sampling, IMS can minimize interference from clouds.

2.1.2. MODIS

A second commonly used snow and ice product is derived from MODIS satellite observations from the Terra and Aqua satellites (Hall and Riggs, 2007). Terra and Aqua have sun-synchronous, near polar orbits with overpass times of 1030 and 1330 hr respectively. Snow cover is calculated using a Normalized Difference Snow Index (NDSI), which examines the difference between observed radiation at visible wavelengths (where snow is highly reflective) and short IR wavelengths (where there is little reflection from snow). Observations are made at 500 m spatial resolution and aggregated to produce daily snow cover fractions on a 0.05° resolution grid. Past evaluations of the standard MODIS snow product show good agreement in cloud-free conditions but often snow is misidentified as cloud (Hall and Riggs, 2007; Yang et al., 2015).

The Multiangle Implementation of Atmospheric Correction (MAIAC) algorithm is ~~also derived from~~another algorithm processing MODIS observations. MAIAC retrievals uses radiances observed by the MODIS Aqua and Terra satellites to provide atmospheric and surface products including snow detection on a 1 km grid (Lyapustin et al., 2011a, 2011b, 2012). While the NDSI used by the standard MODIS product is also used by MAIAC as one of the criteria, the overall snow and cloud detection in MAIAC are different from the standard MODIS algorithm (Lyapustin et al., 2008).

2.1.3. NISE

The Near-real-time Ice and Snow Extent (NISE) provides daily updated snow cover extent information on a 25x25 km grid (Nolin et al., 2005). NISE uses microwave measurements from the Special Sensor Microwave Imager/Sounder (SSM/I) on a sun-synchronous, quasi-polar orbit to observe how microwave radiation emitted by soil is scattered by snow. Products based on microwave measurements such as NISE are known to miss wet and thin snow, as wet snow emits microwave radiation similar to soil, and thin snow does not provide sufficient scattering.

2.1.4. CMC

The Canadian Meteorological Centre (CMC) Daily Snow Depth Analysis Data is a statistical interpolation of snow depth measurements from 8,000 surface sites across Canada and U.S. interpolated using a snow pack model (Brasnett, 1999). Unlike the aforementioned satellite products that provide snow extent, CMC provides snow depths. Daily snow maps are produced at 25 km resolution. As it a reanalysis product, there is a time delay in availability. The CMC snow depths show good agreement with independent observations over midlatitudes and is considered an improvement over previous snow depth climatologies (Brown et al., 2003).

2.2 Surface observations

These snow identification products are evaluated against surface station observations from the Global Historical Climatology Network-Daily (GHCN-D) database, an amalgamation of daily climate records from over 80,000 surface stations worldwide (Menne et al., 2012a). Most observations over Canada and the United States are collected by government organizations (Environment and Climate Change Canada and NOAA National Climatic Data Center, respectively) with additional measurements from smaller observation networks. While the focus of the database is collecting temperature and precipitation measurements, many stations (1,279 in Canada, 13,932 in United States in 2015 used here) also offer snow depth measurements.

A subset of the surface stations included in GHCN-D may also be used in the CMC reanalysis. It is difficult to definitively know which stations are used, as CMC does not routinely archive this information. However, we estimate that only 5% of the GHCN-D stations used here are located within 0.1° of a possible CMC station, and thus GHCN-D has sufficient independent information sources to evaluate the CMC product.

2.3 Radiative transfer calculations

The sensitivity of satellite observations of NO_2 to its vertical distribution is calculated here using the LIDORT radiative transfer model (Spurr, 2002). The model is used to calculate scattering weights, which quantify the sensitivity of backscattered solar radiation to NO_2 at different altitudes (Martin et al., 2002; Palmer et al., 2001). The observation sensitivity to lower tropospheric NO_2 is represented by the air mass factor. Air mass factors for OMI satellite observations in January 2013 are calculated as a useful analog for future TEMPO observations as both instruments are spectrometers observing reflected sunlight at UV to visible wavelengths. AMFs are calculated at 440 nm, at the centre of the NO_2 retrieval window for OMI and TEMPO where NO_2 has strong absorption features. Vertical NO_2 profiles, and other trace gas and aerosol profiles needed for the AMF calculation shown here, are obtained from a simulation of the GEOS-Chem chemical transport model version 11-01 (www.geos-chem.org).

Figure 1 shows maps of snow-free and snow-covered reflectances used here. Snow-free surface reflectance at 470 nm is provided by Nadir BRDF-Adjusted reflectances from the MODIS CMG Gap-Filled Snow-Free Products (Sun et al., 2017). Reflectivities at 354 nm for snow-covered scenes are derived from OMI observations as described by O'Byrne et al. (2010). While this wavelength is different than the 440 nm wavelength used to calculate AMFs, snow reflectivity has weak spectral dependence in UV-Visible wavelengths (Feister and Grewe, 1995; O'Byrne et al., 2010). Snow can increase surface reflectance by over a factor of 10 in central North America where short vegetation is readily covered by snow.

3. Methods

Here we test daily snow cover products for 2015. Snow products are regridded from their native resolutions to a common 4 km grid (similar to the spatial resolution of TEMPO). A grid box is considered to be snow covered if any observations within that box are snow covered. MAIAC, NISE, and IMS give only a yes/no flag for presence of snow. MODIS products provide a pixel snow fraction, and we consider any pixels with nonzero snow fractions as snow covered. Any CMC grid box with nonzero snow depth is considered snow covered.

GHCN-D surface measurements are used as the ground "truth" for evaluating the satellite and reanalysis snow data products tested here. If measurements from multiple surface data networks exist in the same grid box, the most reliable source is used per the priority order given by GHCN-D (Menne et al., 2012b). If observations from multiple surface stations within the most reliable network within a grid box disagree on the presence of snow on a given day, that day is excluded from the evaluation.

We assess the snow data sets using metrics that are commonly used for evaluating binary data sets (Rittger et al., 2013). These metrics are based on the possible outcomes for identifying snow: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Accuracy measures the likelihood that a grid box, with snow or without, is correctly classified:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precision is the probability that a region identified as snow-covered has snow:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall is the likelihood that snow cover is detected when present:

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

The F score balances recall (which accounts for false negatives) and precision (which accounts for false positives) to measure correct classification of snow without the influence of frequent snow-free periods, and therefore is the metric which is most relevant for TEMPO:

$$F = 2 * \frac{precision * recall}{precision + recall} \quad (4)$$

4. Results

We first examine the effect of surface reflectivity on retrieval sensitivity by using the LIDORT radiative transfer model to calculate NO_2 air mass factors for both snow-free and snow-covered scenarios over North America. We calculate air mass factors ~~for all cloud-free (OMI cloud fraction < 20%) OMI NO_2 observations~~ over North America in January 2013. We assume cloud-free conditions in all AMF calculations, as the impact of surface reflectance on retrieved cloud fractions is beyond the scope of this paper.

Figure 2 shows the sensitivity of backscattered radiation (scattering weights) over snow-covered and snow-free~~- retrieval sensitivity (scattering weights)~~surfaces for two~~a~~ locations; a midlatitude location (~~in the~~ US Midwest, ~~(~~42°N, ~~100~~99°W) with a solar zenith angle of 60~~5~~° and at a high latitude location (Northern Canada, 58°N, 76°W) with a solar zenith angle of 79°. The ~~mean~~ snow-covered scattering weights are ~~is~~ greater than the ~~mean~~ snow-free scattering weights throughout the troposphere, by factors of 2.0~~2~~ (2.7) below 5 km, 2.~~6~~7 (3.7) below 2 km, and ~~3.1~~2.6 (5.3) below 1 km at the mid (high) latitude location. This shows that ~~a~~ satellite observed backscattered radiation~~ation~~ is up to ~~three~~ five times as sensitive to NO_2 in the boundary layer in the presence of snow, due to the increased absorption by NO_2 in the lower troposphere when the surface reflects more sunlight.

Figure 3 shows the distribution of AMF values over ~~the TEMPO field of regard~~North America with and without reflectance from snow. The snow-free AMF distribution is unimodal with a median of 1.2. Allowing for the presence of snow introduces a second mode with a median of 3.2~~0~~. Mean AMFs increase by a factor 2.0 in the presence of snow, indicating an overall doubling in the sensitivity to tropospheric NO_2 over snow covered surfaces across North America. The impact is larger over polluted regions, as mean AMFs increase by a factor of 2.2 in regions where NO_2 columns exceed 1×10^{15} molec/cm^2. Maps of AMF with and without snow cover for January 2013 show that AMF values increase over 69% of the land surface within the TEMPO domain.

We next examine the snow datasets to identify the one most suited for the TEMPO retrieval algorithm. Figure 4 shows the spatial distribution of false positives and false negatives in the data sets. In all data sets, both false positives and negatives are most frequent over mountainous regions, particularly in the Rocky Mountain region, consistent with previous validation studies (Chen et al., 2012, 2014; Frei et al., 2012; Frei and Lee, 2010). These errors are often attributed to differences in representativeness, as snow cover in mountain regions is often spatially inhomogeneous, and thus *in situ* measurements may not be representative of the pixel. A slight increase in the number of false positives in IMS over mid-western and prairie regions may result from crop regions with high snow-free albedos being mistaken for snow in visible imagery (Chen et al., 2012; Yang et al., 2015). NISE, MODIS Aqua, and MODIS Terra have more false negatives overall, especially in the Great Lakes and New England regions. False positives are less frequent than false negatives in all data sets. IMS and CMC have the lowest frequency of false negatives. NISE and MAIAC have the lowest frequency of false positives.

Figure 5 shows the metrics used to evaluate data set performance. Table 1 summarizes these results. All data sets have high accuracy numbers, owing largely to a high number of true negatives during the summer months. MODIS Aqua and Terra have low recall and *F* scores. When only observations with MODIS cloud fractions less than 20% are used, MODIS has better agreement with the ground stations (*F* statistic increases from 0.38 to 0.49 at native resolution for Aqua, 0.43 to 0.63 for Terra), however this reduces the number of usable MODIS observations by up to 60%. NISE has high precision but low recall, indicating that while areas classified as snow-covered by NISE are likely correct, many snow-covered regions are missing in the data set. This is consistent with evaluations by McLinden et al. (2014) and O'Byrne et al. (2010). Although CMC, IMS, and MAIAC products show an increase in frequency of false negatives over the Rocky Mountains, they retain a high precision in this region due to frequent snow cover. While MAIAC Aqua/Terra have high accuracy and precision, lower recall values indicate that they are conservative in identifying the presence of snow. This is possibly a consequence of the method used for identifying cloud, which may incorrectly classify fresh snowfall as cloud (Lyapustin et al., 2008). Data sets were also evaluated by season with similar results (Appendix Table A1). All data sets have weaker performance metrics during the spring melt season, which has been observed in past evaluations (Frei et al., 2012). IMS has the highest *F* score in winter and autumn but is slightly outperformed by MAIAC in spring. Data sets were also evaluated at their native resolutions and at a common 25 km resolution (Appendix Tables A2-3). Results are similar at each resolution with two exceptions: MODIS Aqua and Terra products perform better when regridded from their native 0.05° resolution to a 4 km ~~coarser~~ resolution as it reduces the number of grid boxes missing observations due to cloud, and MAIAC Aqua and Terra perform better at their native resolution than at either 4 km or 25 km as degrading the spatial resolution results in a loss of information.

For all data sets, recall is generally low in two regions: along the Pacific coastline where snow depths are relatively thin, and in the south when snow is rare and generally short lived. Thin snow is likely to be less homogenous across a pixel and more likely to be obscured by forest canopies or tall grasses, and thus is difficult to observe from satellite imagery. Short lived snow in the south is likely to be missed by satellite observations, especially since clouds are often present. However, as IMS uses multiple observations at multiple times of day in addition to incorporating ground station data, it is more likely to find snow in these cases than other satellite products (Hall et al., 2010). Overall, IMS has best agreement with *in situ* observations, with the highest accuracy, recall, and *F* statistic and relatively high precision.

While CMC also has strong performance metrics, it is important to consider the information source used to describe snow extent in each product. Products based on satellite observations are advantageous when assessing how surface reflectivity affects backscattered radiation observed from space. For example, thin snow, or snow obscured by tree canopies, may not affect the observed brightness from space, but would be considered snow-covered by a product based on surface observations (e.g. CMC). Also, the reflectivity of a snow-covered surface decreases over time as the snow ages (Warren and Wiscombe, 1980); This effect would not be captured by snow depth measurements. And while snow depth has been used as an indicator of brightness (Arola et al., 2003), it can not account for snow aging or canopy effects. IMS is based on visible satellite imagery and thus determines snow extent based on brightness from space, which is more applicable to satellite retrievals. And while most satellite-based products rely on observations made at a single overpass time and viewing geometry, IMS has the advantage of incorporating observations from multiple satellites with differing measurement times and geometries, including both geostationary and low Earth orbits. These reasons, in addition to a strong agreement with in situ measurements and near-real-time updates, make IMS best suited for informing TEMPO retrievals.

We next examine the effect on both spatial sampling and sensitivity to the lower troposphere of a retrieval data set if observations with surface snow are included rather than omitted. We use IMS to identify the presence of snow for OMI observations over North America in January 2015. We then use LIDORT to calculate AMFs for these observations using the corresponding snow-free (Sun et al., 2017) or snow-covered (O'Byrne et al., 2010) surface reflectance, and examine the results of either including or omitting snow-covered scenes. Figure 6 shows that including snow-covered scenes results in a significant (factor of 2.1) increase in observation frequency, particularly in the northern US and Canada. Additionally, including snow-covered scenes increases the average AMF by a factor of 2.7 in regions with occasional snow cover. The increase in AMF demonstrates that including snow-covered scenes increases the quality of information about the tropospheric NO_2 column by increasing the observation sensitivity to tropospheric NO_2.

5. Conclusion

An accurate representation of snow cover is essential to ensuring satellite retrieval accuracy, including those from TEMPO. Radiative transfer model calculations indicate that NO_2 retrievals over reflective snow-covered surfaces are more than twice as sensitive to NO_2 in the boundary layer than over snow-free surfaces~~, with the greatest increases in sensitivity occurring over polluted regions~~. This makes snow an attractive surface over which to observe tropospheric NO_2. However, the lack of confidence in snow identification has previously led many retrieval procedures to omit observations over snow. We show that ~~I~~increasing this confidence such that these observations could be included ~~would~~ not only improves spatial and temporal sampling, but also allows the inclusion of observations with higher quality information on the lower troposphere.

We evaluated seven snow extent data sets to determine their usefulness for informing satellite retrievals of trace gas from solar backscatter observations. All products were more likely to misidentify snow over mountains or where snow cover is thin or short lived. IMS had the best agreement with *in situ* observations (F=0.85), and as a satellite based, operational, daily updated product, it is well suited for informing TEMPO satellite retrievals. The low recall value (0.45) for NISE indicated that a significant number of snow covered pixels are missed. The standard MODIS products showed medium precision and low recall owing to cloud contamination. The MAIAC products had the highest precision (0.90 for both Aqua and Terra) of those tested, but is conservative in ascribing the presence of snow (recall=0.74 for Aqua, 0.75 for Terra). CMC had strong performance metrics (F=0.81), but as a reanalysis product based on ground observations it may not appropriately represent how a surface snow reflectivity would affect TEMPO observed radiances.

Future work should investigate snow reflectance products ~~, potentially including Bidirectional Reflectance Distribution Functions (BRDF) that describe reflection at different viewing angles,~~ that could be used when snow is detected. This could potentially include Bidirectional Reflectance Distribution Functions (BRDF) that describe reflection at different viewing angles, as this effect has been shown to have significant impact on retrieved NO_2 columns (Vasilkov et al., 2017). A retrieval algorithm that combines daily snow detection from IMS with a climatology of snow reflectance has the potential to greatly improve upon current methodologies.

6. Data Availability

IMS (National Ice Center, 2008), NISE (Brodzik and Stewart, 2016), MODIS Aqua (Hall and Riggs, 2016a), MODIS Terra (Hall and Riggs, 2016b), and CMC (Brown and Brasnett, 2010) data are available from the NASA National Snow and Ice Data Center (http://nsidc.org). MAIAC Collection 6 re-processing of MODIS data started in September 2017 and is expected to be completed by the end of year. This study used MAIAC data currently available via ftp at NASA Center for Climate Simulations (NCCS): ftp://maiac@dataportal.nccs.nasa.gov/DataRelease/. GHCN-D data are available from the NOAA National Climatic Data Center (Menne et al., 2012b; www.ncdn.noaa.gov). Code for calculating scattering weights and air mass factors, and snow-covered surface reflectances used here are available at http://fizz.phys.dal.ca/~atmos. Snow-free surface reflectances are available at ftp://rsftp.eeos.umb.edu/data02/Gapfilled/. The GEOS-Chem chemical transport model used here is available at www.geos-chem.org.

7. References

Arola, A., Kaurola, J., Koskinen, L., Tanskanen, A., Tikkanen, T., Taalas, P., Herman, J. R., Krotkov, N. and Fioletov, V.: A new approach to estimating the albedo for snow-covered surfaces in the satellite UV method, J. Geophys. Res., 108(D17), 4531, doi:10.1029/2003JD003492, 2003.

Beirle, S., Boersma, K. F., Platt, U., Lawrence, M. G. and Wagner, T.: Megacity emissions and lifetimes of nitrogen oxides probed from space., Science, 333(6050), 1737–9, doi:10.1126/science.1207824, 2011.

Boersma, K. F., Eskes, H. J. and Brinksma, E. J.: Error analysis for tropospheric NO2 retrieval from space, J. Geophys. Res. Atmos., 109(D4), 2004.

Brasnett, B.: A Global Analysis of Snow Depth for Numerical Weather Prediction, J. Appl. Meteorol., 38(6), 726–740, doi:10.1175/1520-0450(1999)038<0726:AGAOSD>2.0.CO;2, 1999.

Brodzik, M. J. and Stewart, J. S.: Near-Real-Time SSM/I-SSMIS EASE-Grid Daily Global Ice Concentration and Snow Extent, Version 5, , doi:http://dx.doi.org/10.5067/3KB2JPLFPK3R, 2016.

Brown, R. D. and Brasnett, B.: Canadian Meteorological Centre (CMC) Daily Snow Depth Analysis Data, Version 1, , doi:http://dx.doi.org/10.5067/W9FOYWH0EQZ3, 2010.

Brown, R. D., Brasnett, B. and Robinson, D.: Gridded North American monthly snow depth and snow water equivalent for GCM evaluation, Atmosphere-Ocean, 41(1), 1–14, doi:10.3137/ao.410101, 2003.

Chen, C., Lakhankar, T., Romanov, P., Helfrich, S., Powell, A. and Khanbilvardi, R.: Validation of NOAA-Interactive Multisensor Snow and Ice Mapping System (IMS) by Comparison with Ground-Based Measurements over Continental United States, Remote Sens., 4(12), 1134–1145, doi:10.3390/rs4051134, 2012.

Chen, X., Jiang, L., Yang, J. and Pan, J.: Validation of ice mapping system snow cover over southern China based on Landsat Enhanced Thematic Mapper Plus imagery, J. Appl. Remote Sens., 8(1), 84680, doi:10.1117/1.JRS.8.084680, 2014.

Duncan, B. N., Prados, A. I., Lamsal, L. N., Liu, Y., Streets, D. G., Gupta, P., Hilsenrath, E., Kahn, R. A., Nielsen, J. E., Beyersdorf, A. J., Burton, S. P., Fiore, A. M., Fishman, J., Henze, D. K., Hostetler, C. A., Krotkov, N. A., Lee, P., Lin, M., Pawson, S., Pfister, G., Pickering, K. E., Pierce, R. B., Yoshida, Y. and Ziemba, L. D.: Satellite data of atmospheric pollution for U.S. air quality applications: Examples of applications, summary of data end-user resources, answers to FAQs, and common mistakes to avoid, Atmos. Environ., 94, 647–662, doi:10.1016/j.atmosenv.2014.05.061, 2014.

Feister, U. and Grewe, R.: Spectral albedo measurements in the UV and visible region over different types of surfaces, Photochem. Photobiol., 62(4), 736–744, doi:10.1111/j.1751-1097.1995.tb08723.x, 1995.

Fioletov, V. E., McLinden, C. A., Krotkov, N. and Li, C.: Lifetimes and emissions of SO2 from point sources estimated from OMI, Geophys. Res. Lett., 42(6), 1969–1976, doi:10.1002/2015GL063148, 2015.

Fishman, J., Al-Saadi, J. A., Creilson, J. K., Bowman, K. W., Burrows, J. P., Richter, A., Chance, K. V., Edwards, D. P., Martin, R. V., Morris, G. A., Pierce, R. B., Ziemke, J. R., Schaack, T. K., Thompson, A. M., Fishman, J., Al-Saadi, J. A., Creilson, J. K., Bowman, K. W., Burrows, J. P., Richter, A., Chance, K. V., Edwards, D. P., Martin, R. V., Morris, G. A., Pierce, R. B., Ziemke, J. R., Schaack, T. K. and Thompson, A. M.: Remote Sensing of Tropospheric Pollution from Space, Bull. Am. Meteorol. Soc., 89(6), 805–821, doi:10.1175/2008BAMS2526.1, 2008.

de Foy, B., Lu, Z., Streets, D. G., Lamsal, L. N. and Duncan, B. N.: Estimates of power plant NO x emissions and lifetimes from OMI NO 2 satellite retrievals, Atmos. Environ., 116, 1–11, 2015.

Frei, A. and Lee, S.: A comparison of optical-band based snow extent products during spring over North America, Remote Sens. Environ., 114(9), 1940–1948, doi:10.1016/j.rse.2010.03.015, 2010.

Frei, A., Tedesco, M., Lee, S., Foster, J., Hall, D. K., Kelly, R. and Robinson, D. A.: A review of global satellite-derived snow products, Adv. Sp. Res., 50(8), 1007–1029, doi:10.1016/j.asr.2011.12.021, 2012.

Geddes, J. A. and Martin, R. V.: Global deposition of total reactive nitrogen oxides from 1996 to 2014 constrained with satellite observations of NO2 columns, Atmos. Chem. Phys., 17(16), 10071–10091, doi:10.5194/acp-17-10071-2017, 2017.

Hall, D. . and Riggs, G. A.: MODIS/Aqua Snow Cover Daily L3 Global 0.05Deg CMG, Version 6, , doi:http://dx.doi.org/10.5067/MODIS/MYD10C1.006, 2016a.

Hall, D. K. and Riggs, G. A.: Accuracy assessment of the MODIS snow products, Hydrol. Process., 21(12), 1534–1547, doi:10.1002/hyp.6715, 2007.

Hall, D. K. and Riggs, G. A.: MODIS/Terra Snow Cover Daily L3 Global 0.05Deg CMG, Version 6, , doi:http://dx.doi.org/10.5067/MODIS/MOD10C1.006, 2016b.

Hall, D. K., Fuhrmann, C. M., Perry, L. B., Riggs, G. A., Robinson, D. A. and Foster, J. L.: A Comparison of Satellite-Derived Snow Maps with a Focus on Ephemeral Snow in North Carolina, in 67th Eastern Snow Conference, Hancock MA USA., 2010.

Helfrich, S. R., McNamara, D., Ramsay, B. H., Baldwin, T. and Kasheta, T.: Enhancements to, and forthcoming developments in the Interactive Multisensor Snow and Ice Mapping System (IMS), Hydrol. Process., 21(12), 1576–1586, doi:10.1002/hyp.6720, 2007.

Herman, J. R. and Celarier, E. A.: Earth surface reflectivity climatology at 340-380 nm from TOMS data, J. Geophys. Res. Atmos., 102(D23), 28003–28011, doi:10.1029/97JD02074, 1997.

Kleipool, Q. L., Dobber, M. R., de Haan, J. F. and Levelt, P. F.: Earth surface reflectance climatology from 3 years of OMI data, J. Geophys. Res., 113(D18), D18308, doi:10.1029/2008JD010290, 2008.

Koelemeijer, R. B. A., de Haan, J. F. and Stammes, P.: A database of spectral surface reflectivity in the range 335–772 nm derived from 5.5 years of GOME observations, J. Geophys. Res., 108(D2), 4070, doi:10.1029/2002JD002429, 2003.

Lamsal, L. N., Janz, S. J., Krotkov, N. A., Pickering, K. E., Spurr, R. J. D., Kowalewski, M. G., Loughner, C. P., Crawford, J. H., Swartz, W. H. and Herman, J. R.: High-resolution NO2 observations from the Airborne Compact Atmospheric Mapper: Retrieval and validation, J. Geophys. Res. Atmos., 122(3), 1953–1970, doi:10.1002/2016JD025483, 2017.

Liang, S., Fang, H., Chen, M., Shuey, C. J., Walthall, C., Daughtry, C., Morisette, J., Schaaf, C. and Strahler, A.: Validating MODIS land surface reflectance and albedo products: methods and preliminary results, Remote Sens. Environ., 83(1–2), 149–162, doi:10.1016/S0034-4257(02)00092-5, 2002.

Lin, J.-T., Liu, M.-Y., Xin, J.-Y., Boersma, K. F., Spurr, R., Martin, R. and Zhang, Q.: Influence of aerosols and surface reflectance on satellite NO2 retrieval: seasonal and spatial characteristics and implications for NOx emission constraints, Atmos. Chem. Phys., 15(19), 11217–11241, doi:10.5194/acp-15-11217-2015, 2015.

Lorente, A., Folkert Boersma, K., Yu, H., Dörner, S., Hilboll, A., Richter, A., Liu, M., Lamsal, L. N., Barkley, M., De Smedt, I., Van Roozendael, M., Wang, Y., Wagner, T., Beirle, S., Lin, J.-T., Krotkov, N., Stammes, P., Wang, P., Eskes, H. J. and Krol, M.: Structural uncertainty in air mass factor calculation for NO2 and HCHO satellite retrievals, Atmos. Meas. Tech., 10(3), 759–782, doi:10.5194/amt-10-759-2017, 2017.

Lyapustin, A., Wang, Y. and Frey, R.: An automatic cloud mask algorithm based on time series of MODIS measurements, J. Geophys. Res., 113(D16), D16207, doi:10.1029/2007JD009641, 2008.

Lyapustin, A., Martonchik, J., Wang, Y., Laszlo, I. and Korkin, S.: Multiangle implementation of atmospheric correction (MAIAC): 1. Radiative transfer basis and look-up tables, J. Geophys. Res., 116(D3), D03210, doi:10.1029/2010JD014985, 2011a.

Lyapustin, A., Wang, Y., Laszlo, I., Kahn, R., Korkin, S., Remer, L., Levy, R. and Reid, J. S.: Multiangle implementation of atmospheric correction (MAIAC): 2. Aerosol algorithm, J. Geophys. Res., 116(D3), D03211, doi:10.1029/2010JD014986, 2011b.

Lyapustin, A. I., Wang, Y., Laszlo, I., Hilker, T., G.Hall, F., Sellers, P. J., Tucker, C. J. and Korkin, S. V.: Multi-angle implementation of atmospheric correction for MODIS (MAIAC): 3. Atmospheric correction, Remote Sens. Environ., 127, 385–393, doi:10.1016/j.rse.2012.09.002, 2012.

Martin, R. V, Chance, K., Jacob, D. J., Kurosu, T. P., Spurr, R. J. D., Bucsela, E., Gleason, J. F., Palmer, P. I., Bey, I. and Fiore, A. M.: An improved retrieval of tropospheric nitrogen dioxide from GOME, J. Geophys. Res. Atmos., 107(D20), 2002.

Martin, R. V.: Satellite remote sensing of surface air quality, Atmos. Environ., 42(34), 7823–7843, doi:10.1016/j.atmosenv.2008.07.018, 2008.

McLinden, C. A., Fioletov, V., Boersma, K. F., Kharol, S. K., Krotkov, N., Lamsal, L., Makar, P. A., Martin, R. V., Veefkind, J. P. and Yang, K.: Improved satellite retrievals of NO2 and SO2 over the Canadian oil sands and comparisons with surface measurements, Atmos. Chem. Phys., 14(7), 3637–3656, doi:10.5194/acp-14-3637-2014, 2014.

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E., Houston, T. G., Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E. and Houston, T. G.: An Overview of the Global Historical Climatology Network-Daily Database, J. Atmos. Ocean. Technol., 29(7), 897–910, doi:10.1175/JTECH-D-11-00103.1, 2012a.

Menne, M. J., Durre, I., Korzeniewski, B., McNeal, S., Thomas, K., Yin, X., Anthony, S., Ray, R., Vose, R. S., Gleason, B. E. and Houston, T. G.: Global Historical Climatology Network - Daily (GHCN-Daily), Version 3.22, , doi:http://doi.org/10.7289/V5D21VHZ, 2012b.

National Ice Center: IMS Daily Northern Hemisphere Snow and Ice Analysis at 1 km, 4 km, and 24 km Resolutions, Version 1, , doi:http://dx.doi.org/10.7265/N52R3PMC, 2008.

Nolin, A., Armstrong, R. and Maslanik, J.: Near real-time SSM/I EASE-grid daily global ice concentration and snow extent, Digit. Media, Natl. Snow Ice Data Center, Boulder, CO, USA, 2005.

Nowlan, C. R., Martin, R. V., Philip, S., Lamsal, L. N., Krotkov, N. A., Marais, E. A., Wang, S. and Zhang, Q.: Global dry deposition of nitrogen dioxide and sulfur dioxide inferred from space-based measurements, Global Biogeochem. Cycles, 28(10), 1025–1043, doi:10.1002/2014GB004805, 2014.

O'Byrne, G., Martin, R. V., van Donkelaar, A., Joiner, J. and Celarier, E. A.: Surface reflectivity from the Ozone Monitoring Instrument using the Moderate Resolution Imaging

Spectroradiometer to eliminate clouds: Effects of snow on ultraviolet and visible trace gas retrievals, J. Geophys. Res., 115(D17), D17305, doi:10.1029/2009JD013079, 2010.

Palmer, P. I., Jacob, D. J., Chance, K. and Martin, R. V: Air mass factor formulation for spectroscopic measurements from satellites' Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment, J. Geophys. Res., 106(D13), 14,514-539,550, 2001.

Richter, A. and Wagner, T.: The Use of UV, Visible and Near IR Solar Back Scattered Radiation to Determine Trace Gases, pp. 67–121, Springer, Berlin, Heidelberg., 2011.

Rittger, K., Painter, T. H. and Dozier, J.: Assessment of methods for mapping snow cover from MODIS, Adv. Water Resour., 51, 367–380, doi:10.1016/j.advwatres.2012.03.002, 2013.

Spurr, R. J. D.: Simultaneous derivation of intensities and weighting functions in a general pseudo-spherical discrete ordinate radiative transfer treatment, J. Quant. Spectrosc. Radiat. Transf., 75(75), 129–175 [online] Available from: www.elsevier.com/locate/jqsrt (Accessed 20 July 2017), 2002.

Streets, D. G., Canty, T., Carmichael, G. R., de Foy, B., Dickerson, R. R., Duncan, B. N., Edwards, D. P., Haynes, J. A., Henze, D. K. and Houyoux, M. R.: Emissions estimation from satellite retrievals: A review of current capability, Atmos. Environ., 77, 1011–1042, 2013.

Sun, Q., Wang, Z., Li, Z., Erb, A. and Schaaf, C. B.: Evaluation of the global MODIS 30 arc-second spatially and temporally complete snow-free land surface albedo and reflectance anisotropy dataset, Int. J. Appl. Earth Obs. Geoinf., 58, 36–49, doi:10.1016/j.jag.2017.01.011, 2017.

Valin, L. C., Russell, A. R. and Cohen, R. C.: Variations of OH radical in an urban plume inferred from NO2 column measurements, Geophys. Res. Lett., 40(9), 1856–1860, 2013.

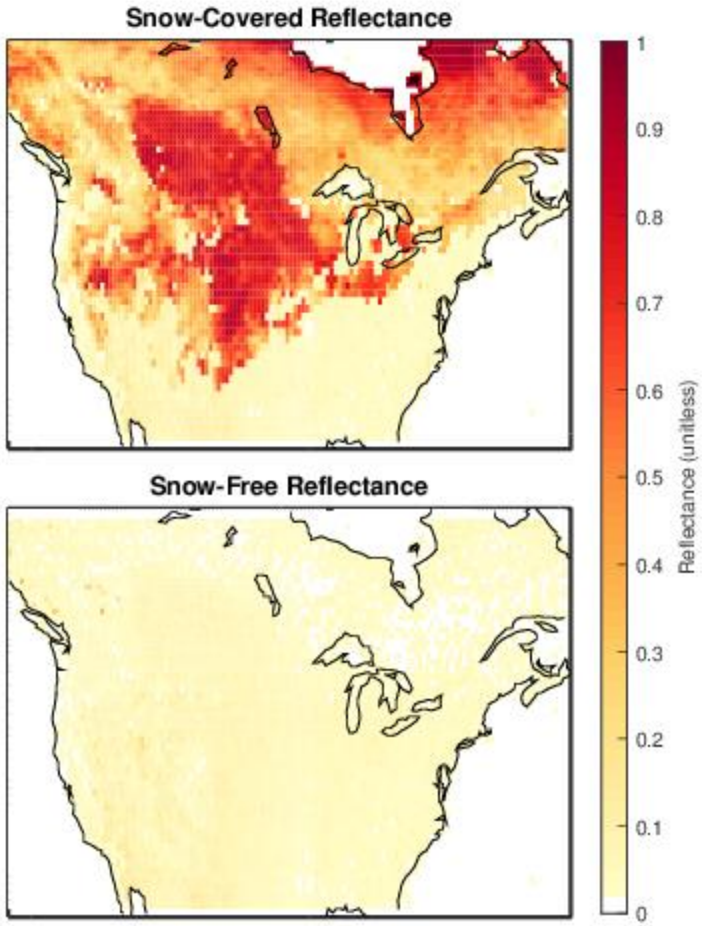
Vasilkov, A., Qin, W., Krotkov, N., Lamsal, L., Spurr, R., Haffner, D., Joiner, J., Yang, E.-S. and Marchenko, S.: Accounting for the effects of surface BRDF on satellite cloud and trace-gas retrievals: a new approach based on geometry-dependent Lambertian equivalent reflectivity applied to OMI algorithms, Atmos. Meas. Tech., 10(1), 333–349, doi:10.5194/amt-10-333-2017, 2017.

Warren, S. G. and Wiscombe, W. J.: A Model for the Spectral Albedo of Snow. II: Snow

Containing Atmospheric Aerosols, J. Atmos. Sci., 37(12), 2734–2745, doi:10.1175/1520-0469(1980)037<2734:AMFTSA>2.0.CO;2, 1980.

Yang, J., Jiang, L., Ménard, C. B., Luojus, K., Lemmetyinen, J. and Pulliainen, J.: Evaluation of snow products over the Tibetan Plateau, Hydrol. Process., 29(15), 3247–3260, doi:10.1002/hyp.10427, 2015.

Zoogman, P., Liu, X., Suleiman, R. M., Pennington, W. F., Flittner, D. E., Al-Saadi, J. A., Hilton, B. B., Nicks, D. K., Newchurch, M. J., Carr, J. L., Janz, S. J., Andraschko, M. R., Arola, A., Baker, B. D., Canova, B. P., Chan Miller, C., Cohen, R. C., Davis, J. E., Dussault, M. E., Edwards, D. P., Fishman, J., Ghulam, A., González Abad, G., Grutter, M., Herman, J. R., Houck, J., Jacob, D. J., Joiner, J., Kerridge, B. J., Kim, J., Krotkov, N. A., Lamsal, L., Li, C., Lindfors, A., Martin, R. V., McElroy, C. T., McLinden, C., Natraj, V., Neil, D. O., Nowlan, C. R., O′Sullivan, E. J., Palmer, P. I., Pierce, R. B., Pippin, M. R., Saiz-Lopez, A., Spurr, R. J. D., Szykman, J. J., Torres, O., Veefkind, J. P., Veihelmann, B., Wang, H., Wang, J. and Chance, K.: Tropospheric emissions: Monitoring of pollution (TEMPO), J. Quant. Spectrosc. Radiat. Transf., 186, 17–39, doi:10.1016/j.jqsrt.2016.05.008, 2017.


Snow-Covered Reflectance
Snow-Free Reflectance
Reflectance (unitless)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

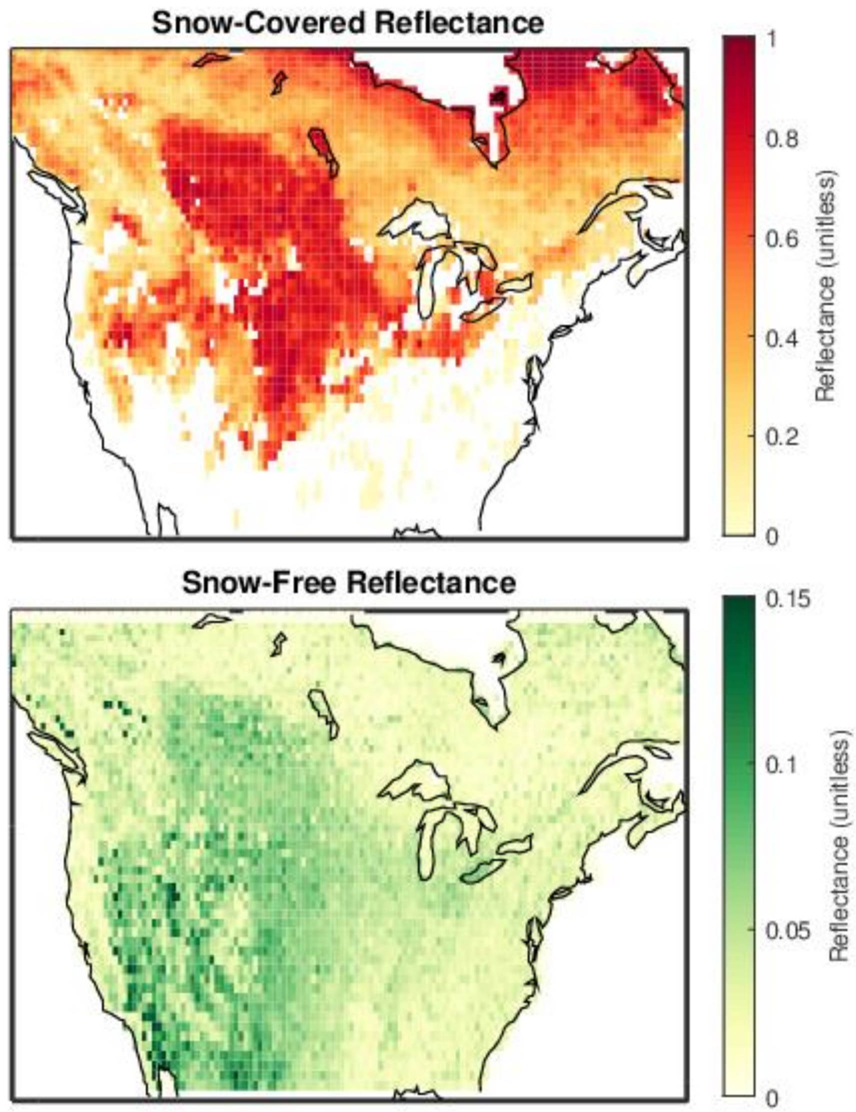
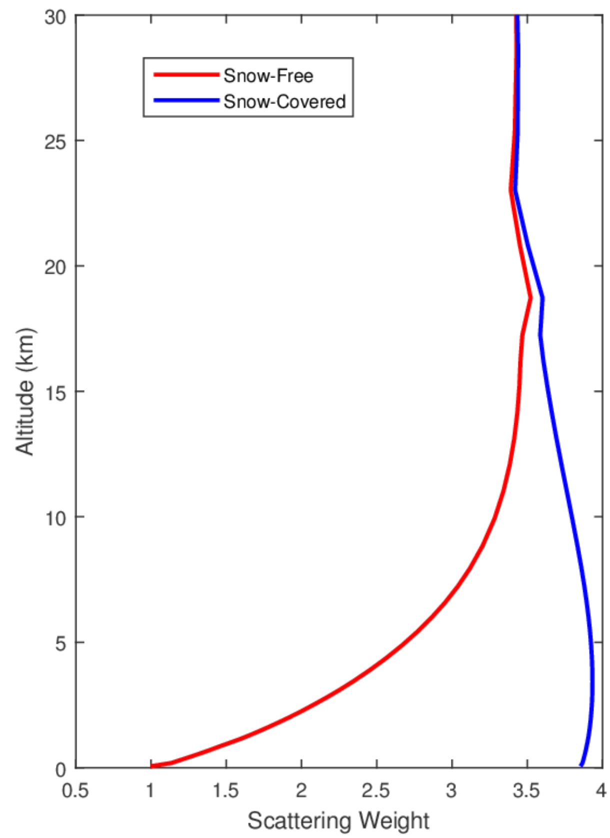


Figure 1: Surface reflectivity at UV-visible wavelengths for snow-covered and snow-free conditions for January 2013. White space in top panel indicates no snow reflectance information is available.


Snow-Free
Snow-Covered
Altitude (km)
Scattering Weight

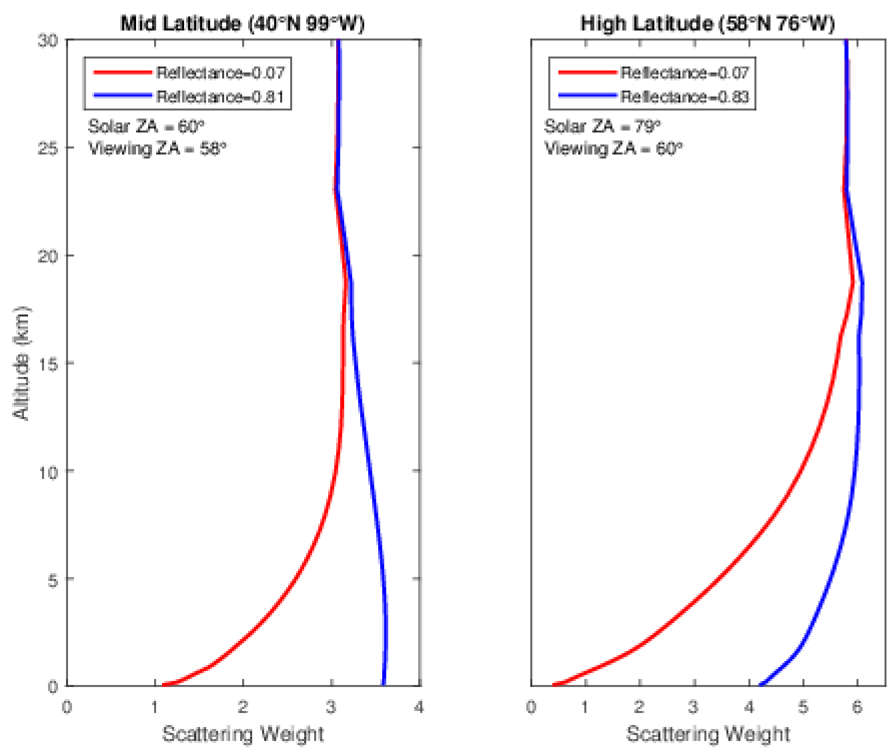
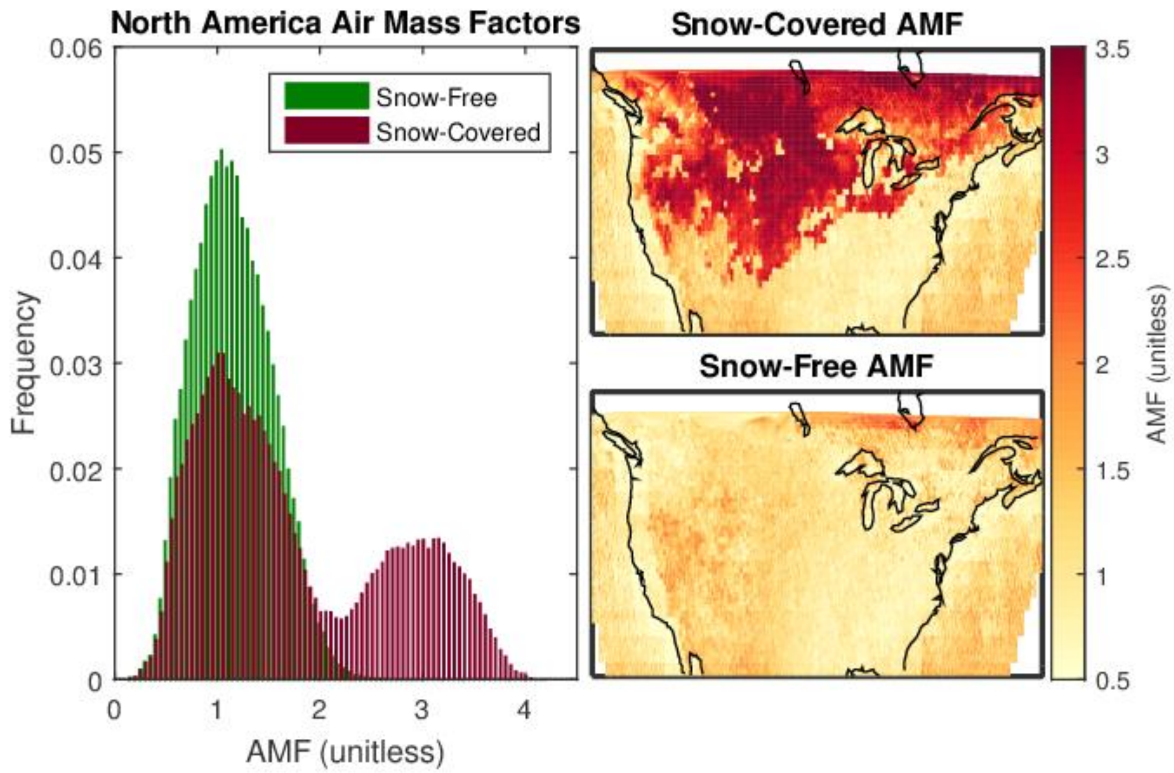

Mid Latitude (40°N 99°W)
Reflectance=0.07
Reflectance=0.81
Solar ZA = 60°
Viewing ZA = 58°
High Latitude (58°N 76°W)
Reflectance=0.07
Reflectance=0.83
Solar ZA = 79°
Viewing ZA = 60°
Altitude (km)
Scattering Weight
Scattering Weight

Figure 2: Observation sensitivity to NO_2. Scattering weight profiles calculated for cloud-free OMI NO_2 retrievals, with and without surface snow cover, for January 2013 at (Left) 42° N, ~~100~~99° ~~-~~W~~ for January 2013~~ with a solar zenith angle (ZA) of ~~65~~60° and (Right) 58° N, 76° W with a solar zenith angle of 79°.~~.~~



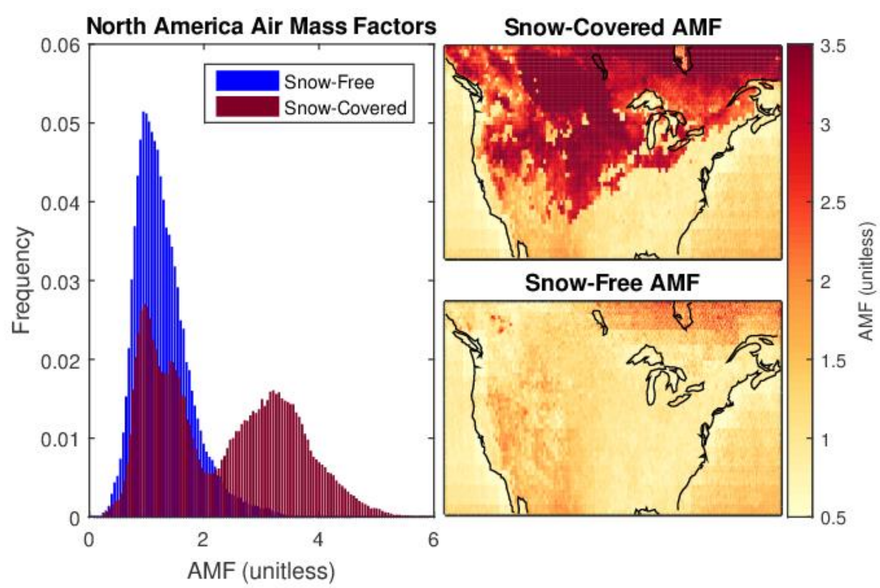


Figure 3: (Left) Distribution of Air Mass Factors (AMFs) calculated for OMI NO_2 retrievals over North America for observation geometry of January 2013, with and without surface snow cover. (Right) Maps of AMF for snow-covered and snow-free conditions.

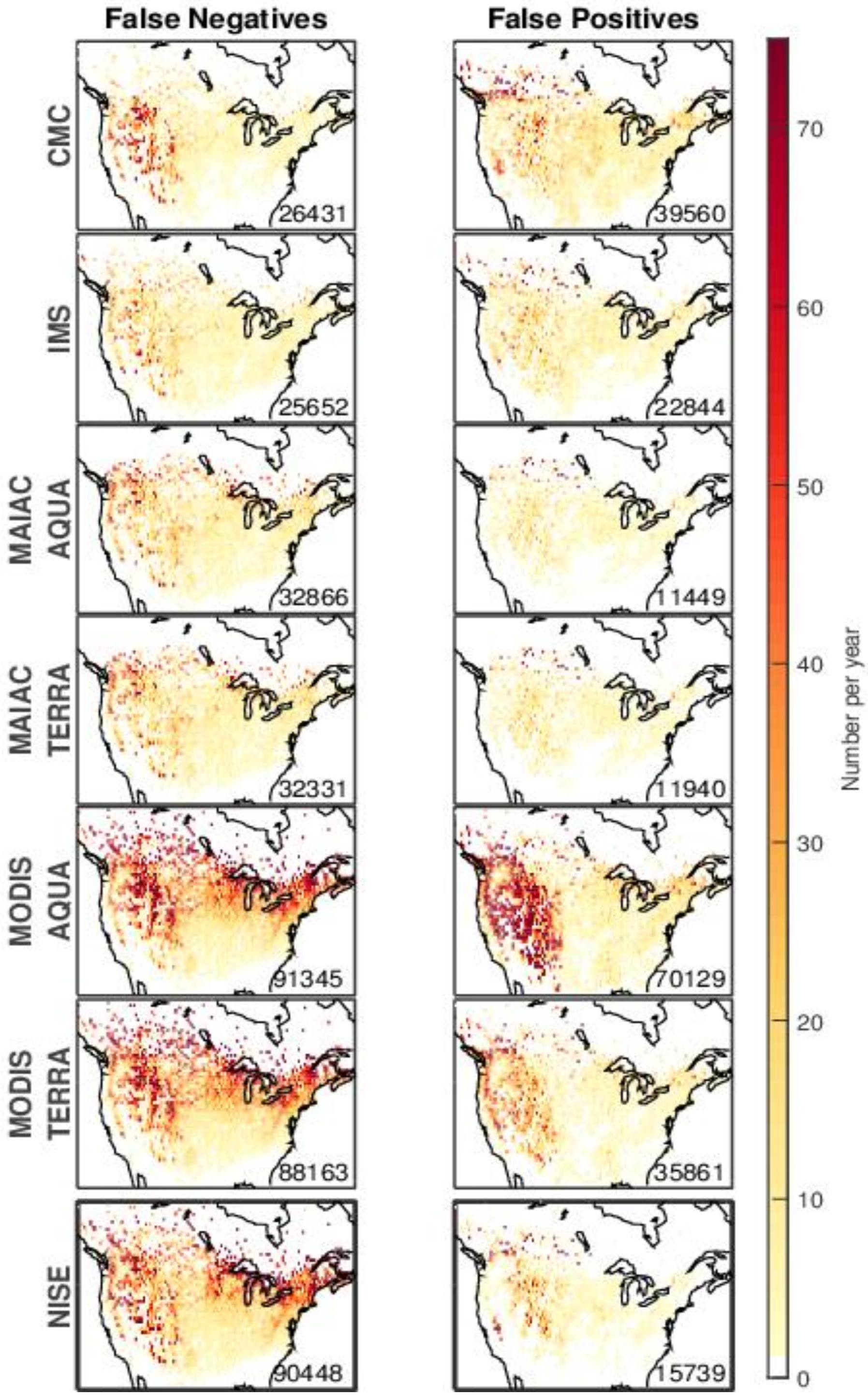


Figure 4: Number of false positive (FP) and false negative (FN) snow attributions by the snow data sets in 2015. All data sets are evaluated at 4 km resolution. Total number of false snow attributions inset. White space indicates no ground stations present.

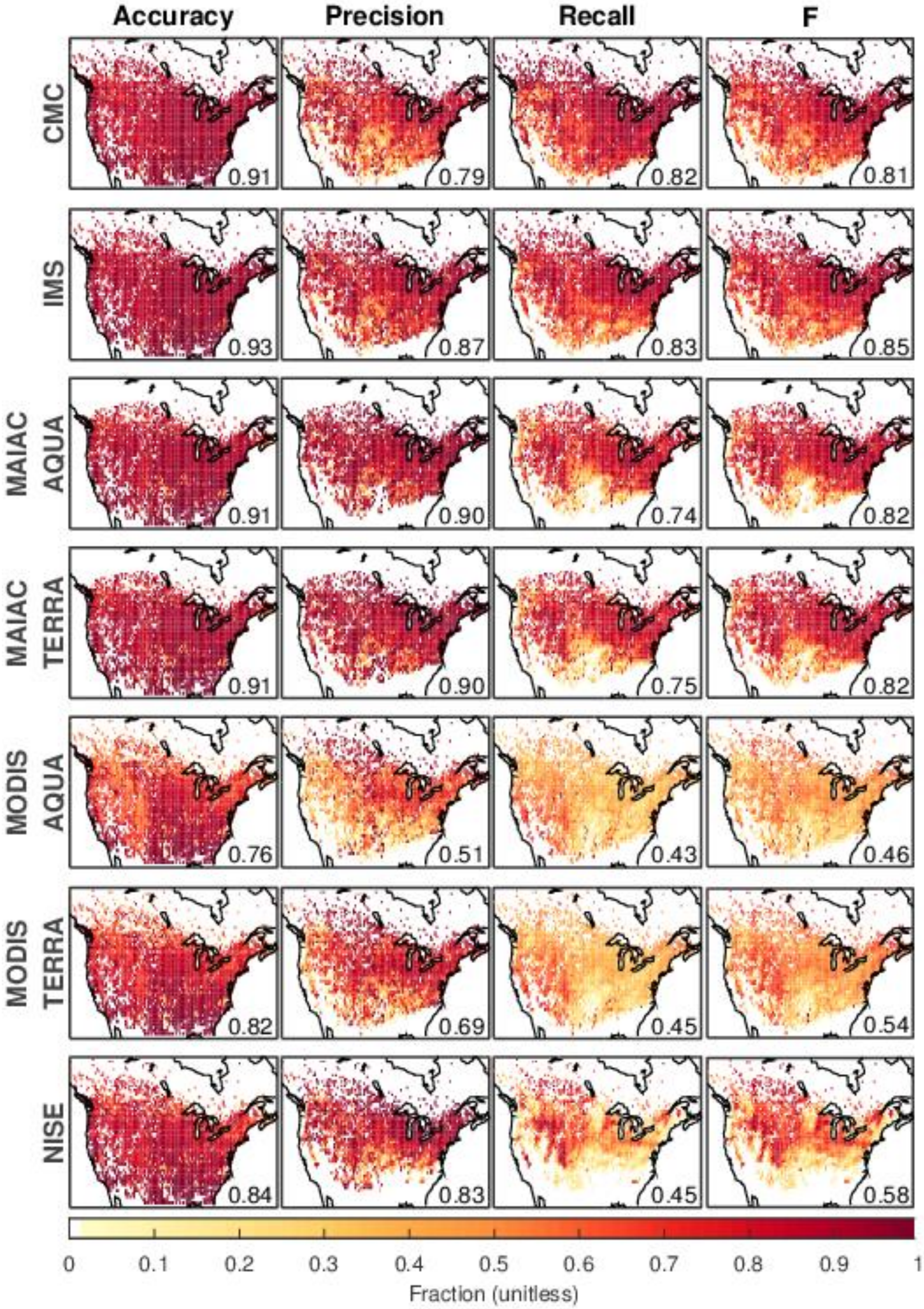


Figure 5: Statistical metrics to evaluate snow cover products. All data sets are gridded at 4 km resolution. White space indicates no ground stations present.

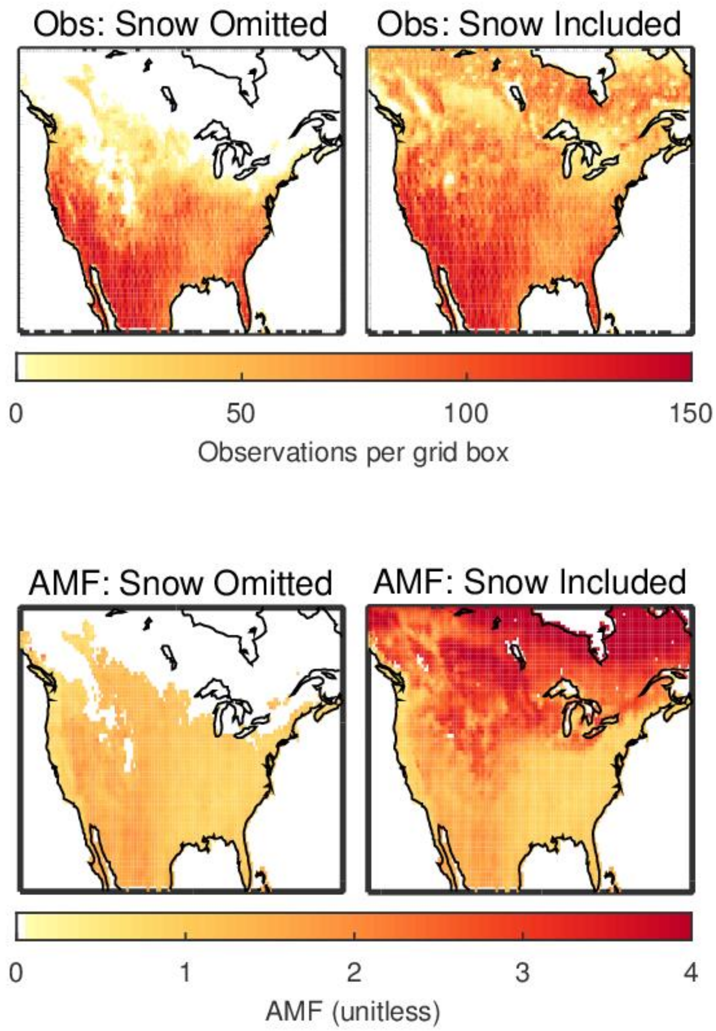


Figure 6: OMI observation frequency (top) and average AMFs (bottom) over North America in January using IMS to identify surface snow conditions. White space indicates a lack of observations.

	Accuracy	Precision	Recall	F
CMC	0.91	0.79	**0.83**	0.81
IMS	**0.93**	0.87	**0.83**	**0.85**
MAIAC AQUA	0.91	**0.90**	0.74	0.82
MAIAC TERRA	0.91	**0.90**	0.75	0.82
MODIS AQUA	0.76	0.51	0.43	0.46
MODIS TERRA	0.82	0.69	0.45	0.54
NISE	0.84	0.83	0.45	0.58

Table 1: ~~Metrics for evaluating~~Evaluation of daily snow extent data set performance for 2015. GHCN-D surface observations are used as “truth”. All products are regridded to a common 4 km resolution. The highest value for each metric is shown in bold.

Appendix

Months	Data Set	Accuracy	Precision	Recall	F
DJF	CMC	0.84	0.84	0.89	0.86
	IMS	**0.88**	0.90	**0.88**	**0.89**
	MAIAC AQUA	0.84	**0.93**	0.80	0.86
	MAIAC TERRA	0.84	0.92	0.80	0.86
	MODIS AQUA	0.58	0.84	0.34	0.48
	MODIS TERRA	0.60	0.88	0.37	0.52
	NISE	0.63	0.90	0.41	0.57
MAM	CMC	0.90	0.63	0.57	0.59
	IMS	**0.93**	0.74	**0.67**	0.70
	MAIAC AQUA	**0.93**	**0.81**	0.62	**0.71**
	MAIAC TERRA	**0.93**	**0.81**	0.63	**0.71**
	MODIS AQUA	0.86	0.43	0.39	0.41
	MODIS TERRA	0.89	0.62	0.40	0.49
	NISE	0.90	0.71	0.34	0.46
SON	CMC	0.91	0.73	**0.81**	0.76
	IMS	**0.92**	0.82	0.74	**0.78**
	MAIAC AQUA	0.91	**0.86**	0.60	0.71
	MAIAC TERRA	0.90	0.85	0.61	0.71
	MODIS AQUA	0.82	0.51	0.36	0.42
	MODIS TERRA	0.86	0.71	0.39	0.51
	NISE	0.85	0.85	0.25	0.39

Table A1: Evaluation of daily snow extent data set performance by season for 2015. GHCN-D surface observations are used as “truth”. All products are regridded to a common 4 km resolution. The highest value for each metric/season is shown in bold.

	Resolution	Accuracy	Precision	Recall	F
CMC	25 km	0.92	0.81	0.81	0.81
IMS	4 km	**0.93**	0.87	**0.83**	**0.85**
MAIAC AQUA	1 km	0.91	**0.91**	0.71	0.80
MAIAC TERRA	1 km	0.91	0.90	0.71	0.80
MODIS AQUA	0.05°	0.~~79~~77	0.~~79~~50	0.~~11~~30	0.~~19~~37
MODIS TERRA	0.05°	0.~~80~~81	0.~~79~~65	0.~~17~~32	0.~~28~~43
NISE	25 km	0.85	0.87	0.37	0.51

Table ~~A1~~A2: Evaluation of ~~Metrics for evaluating~~ daily snow extent data set performance for 2015. GHCN-D surface observations are used as “truth”. The highest value for each metric is shown in bold.

	Accuracy	Precision	Recall	F
CMC	0.92	0.81	0.81~~0~~	0.81~~0~~
IMS	**0.93**	0.84	**0.85~~4~~**	**0.84**
MAIAC AQUA	0.87	0.69	0.~~69~~73	0.~~69~~71
MAIAC TERRA	0.88~~7~~	0.68	0.~~68~~73	0.~~68~~71
MODIS AQUA	0.78	0.50~~49~~	0.41	0.45
MODIS TERRA	0.83	0.68	0.43	0.53
NISE	0.85	**0.~~86~~87**	0.37	0.~~51~~52

Table ~~A2~~A3: ~~Metrics for evaluating~~ Evaluation of daily snow extent data set performance for 2015. GHCN-D surface observations are used as “truth”. All products are regridded to a common 25 km resolution. The highest value for each metric is shown in bold.