

A neural network cloud mask for PACE OCI

A.M. Sayer^{1,2}, I. T. Carroll^{1,2}, and A. V. Semenov^{2,3}

¹Goddard Earth Sciences Technology and Research (GESTAR) II, University of Maryland
Baltimore County, Baltimore, MD, USA

²Ocean Ecology Laboratory, Mail code 616, NASA Goddard Space Flight Center, Greenbelt,
MD, USA

³Science Application International Corp, Lanham, MD, USA

Corresponding author: Andrew Sayer (andrew.sayer@nasa.gov)

Key Points:

- We train a neural network to the MODIS cloud mask using MODIS reflectance inputs.
- The network is applied to OCI observations using similar spectral bands.
- The strengths and limitations are in line with physical expectations.

Version: 1.0

Release Date: TBD

DOI: TBD

Abstract

Cloud identification is the first step in most satellite data processing. Algorithms to detect clouds vary from sensor to sensor based on their differing spectral, spatial, radiometric, and orbital/angular sampling capabilities. The MODerate resolution Imaging SpectroRadiometer (MODIS) has a well-understood cloud mask with a more than 20 year time series. MODIS has some commonality with the Ocean Color Instrument (OCI) on the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission, including 11 spectral bands with close equivalents. We train a simple multi-layer perceptron neural network (NN) with a single hidden layer to emulate the MODIS standard cloud mask (MYD35 data stream), using these MODIS solar reflectance inputs, and achieve approximately 96% classification match. We then apply this MODIS-trained neural network to OCI data directly to provide the OCI cloud mask currently used in the PACE data processing stream. We find a high degree of consistency between the three masks (MYD35, MODIS NN, and OCI). Currently identified weaknesses of the NN approach include (1) misidentification of optically-thick aerosol plumes (e.g. smoke, dust) as clouds; (2) under-detection of optically-thin cirrus clouds, particularly in the tropics; (3) poorer agreement over bright snow and ice surfaces; and (4) some over-screening of rivers and coastlines. These limitations are in line with physical expectations.

Plain Language Summary

If you want to examine clouds from space, you need to know what pixels of an image contain clouds; if you want to look at other things like the Earth's surface, you need to know where the clouds are not. Figuring this out is a process called cloud masking and is mostly based on the facts that clouds tend to be bright, white, textured, and cold. The MODIS instruments have flown in space for more than 20 years and have a cloud masking approach which is very well documented and understood. We want to translate that to a mask for the new OCI instrument on NASA's PACE satellite, which has some common characteristics with MODIS. We trained a neural network to emulate the MODIS cloud mask, which is about 95% accurate compared to the original. We then apply this neural network to OCI data. We find similar patterns of cloudiness, which indicates that the mask works quite well for OCI, and can be used by others to identify which parts of an OCI image may be useful for their purposes. However there are some limitations, such as a difficulty in telling clouds from dust or smoke, at seeing very transparent clouds, and at seeing clouds over snow and ice. These differences are expected.

Keywords: clouds, cloud mask, MODIS, OCI, PACE

Version Description

This is Version 1.0 of the neural network PACE OCI cloud mask algorithm. It describes the algorithm as distributed within PACE OCI version 3.0 data and onwards.

1 Introduction

This ATBD describes a cloud masking algorithm for the Ocean Color Instrument (OCI) on the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission, which launched on February 8 2024. Identifying which pixels contain clouds and which do not, often called cloud masking, is generally the first step in satellite data processing because algorithms to determine cloud properties should only be applied to cloudy scenes and vice-versa. This algorithm draws from the skill of the well-understood cloud mask approach from the MODerate resolution Imaging Spectroradiometer (MODIS) instruments, specifically the MODIS on the Aqua satellite. This was chosen because the two share some common bands and have similar pixel sizes, orbit times, and solar and view geometries.

PACE is in a Sun-synchronous polar orbit with an Equatorial local solar crossing time of 1 pm for the ascending (daytime) node. OCI (Meister *et al.*, 2024) is a broad-swath (2,600 km) passive imaging radiometer, with a pixel horizontal size of 1.2 km at the sub-satellite point. To decrease the proportion of the swath affected by strong Sun glint over water, OCI tilts rearwards by approximately 20° in the Southern hemisphere and forwards by the same amount in the Northern hemisphere (with the tilt change happening around the Equator). OCI measures continuous spectra from the ultraviolet (UV, 315 nm) to near-infrared (NIR, 895 nm) with 2.5 nm spacing (improving to 1.25 nm spacing from 640 to 715 nm and 740 to 775 nm) and a full width half maximum (FWHM) of approximately 5 nm throughout. It also measures at seven discrete shortwave infrared (SWIR) bands centered near 940, 1038, 1378, 1250, 1615, 2130, and 2260 nm. Each scan provides a single line of pixels of data in each band.

Aqua launched in May 2002 and is in a Sun-synchronous polar orbit with an Equatorial local solar crossing time of 1:30 pm for the ascending (daytime) node. MODIS (Barnes *et al.*, 2003) is a broad-swath (2,300 km) passive imaging radiometer, with a pixel horizontal size of 0.25 to 1 km at the sub-satellite point dependent on band. Unlike OCI, MODIS does not tilt. MODIS measures at 36 discrete bands from the visible to the thermal infrared (TIR); we use inputs aggregated to the common 1 km nominal spatial resolution provided in with MODIS MYD021KM level 1 product. Unlike OCI, a single scan consists of more than a single row of pixels: 10 of the nominal 1 km pixels, 20 of the nominal 0.5 km pixels, and 40 of the nominal 0.25 km pixels.

Due to Earth's shape and the scan geometry, both MODIS and OCI imagery have a bow-tie distortion whereby pixels become larger (especially in the across-track direction) and overlap away from the sub-satellite point (which is why we describe pixel sizes for both instruments as "nominal"). We note MODIS also flies on the Terra platform, although that has a daytime (descending) crossing time of 10:30 am, so we use only Aqua data due to its greater similarity in crossing time (and hence solar zenith angle) to PACE.

In developing the cloud mask algorithm, we first use MODIS solar reflectances to train a neural network (NN) that emulates the MODIS cloud mask. We select bands that MODIS and OCI both sample similarly (noting that no two sensors have identical spectral response functions). In this

way, the MODIS-trained neural network can also be applied to PACE OCI measurements. We evaluate the NN against the source MODIS cloud masks by applying it to MODIS observations not used in training, and to OCI data, for the first year of the PACE mission (March 2024 to February 2025).

The algorithm is part of the operational PACE data processing chain, and the associated CLDMASK product currently has Provisional maturity. It is an input to the downstream codes used to retrieve cloud height and optical properties, as provided in the OCI cloud (CLD) product suite. It is also used as an optional mask in the land surface reflectance (SFREFL) and vegetation index (LANDVI) product suites, and is provided as an input for polarimeter retrievals in the level 1c ancillary data files.

2 Context/Background

The initial development work for this algorithm (training a neural network to emulate the MODIS cloud mask) was done by undergraduate intern Taha Al-Nufaili (a student at University of Maryland, College Park) in summer 2022. This was heavily refined by Ian Carroll, who also developed a stand-alone Python application of the network to process OCI inputs. Alex Semenov translated the Python code into C as part of the operational PACE data processing chain. Andrew Sayer is the cloud subject matter expert who conceptualized the approach, prepared the labelled data for NN training and validation, and led the algorithm evaluation.

The work builds on the MODIS cloud mask (specifically, that of the current Collection 6.1), as applied to MODIS Aqua measurements to generate the MYD35 data product. The previous Collection 5 algorithm was described and evaluated by Ackerman et al. (2008) and Frey et al. (2008); refinements for Collection 6 (also applied for Collection 6.1) are given in Section 3(a) of Baum et al. (2012). Essentially, the MYD35 algorithm applies several series of spectral tests (a mixture of thresholds and spectral differences or ratios) and ancillary data which are suited to identify various types of clouds from solar and TIR bands. The thresholds come from a combination of radiative transfer results and expertise gleaned from extensive evaluation analyses over the previous several decades. These tests are grouped according to the types of clouds they are expected to be useful to identify. The results from these groups are then combined to provide an overall level of confidence that a pixel is cloudy, expressed as a four-way classification (“confidently cloudy”, “probably cloudy”, “probably clear”, and “confidently clear”). The overwhelming majority of pixels belong to one of the “confidently” classes, and many users (including this work) collapse this four-way to a binary classification whereby “confidently cloudy” and “probably cloudy” pixels are treated as overcast (cloud-covered), and “probably clear” and “confidently clear” pixels are treated as clear-sky (free from clouds).

3 Algorithm Description

3.1 Scientific Theory

Clouds can be detected from space at solar wavelengths because they scatter and absorb sunlight, which in general results in increased brightness and more limited spectral variation (a spectral whitening) compared to a cloud-free pixel. At TIR wavelengths (observed by MODIS but not OCI), they can be detected because they are generally colder than Earth's surface. TIR bands are particularly well suited for detecting clouds at night (where there is no solar illumination to reflect), and (during the daytime) when the brightening from scattering is small (e.g. semi-transparent cirrus clouds), or contrast with the underlying surface is limited (e.g. above a snow or ice-covered surface). Note, that the algorithm described here for OCI is only applied to daytime observations (defined as a solar zenith angle smaller than 82°) as the PACE observatory does not collect science data at night.

The main underlying assumption behind most (if not all) cloud masking algorithms is that a satellite pixel can be classified in a binary way as covered by a cloud or cloud-free. In reality, particularly near the edge of a cloud, a pixel is likely to be partially cloudy; sub-pixel inhomogeneity becomes an increasingly large complication as satellite pixels get larger. While this issue has been discussed and quantification has been attempted for some time (e.g. DiGirolamo & Davies, 1997; Koren *et al.*, 2008; Jones *et al.*, 2012; Kotarba, 2020), in practice it is accepted as a reasonable approximation for many end uses, at least for imagers with footprints of order 1 km or finer: downstream cloud retrieval algorithms retrieve pixel-effective properties while downstream non-cloud (e.g. surface, aerosol) retrieval algorithms tend to be clear-sky conservative and prefer to reject partially cloudy pixels. The other limitation is that there is no strict definition for what amount of atmospheric water constitutes a cloud, and different sensors (as well as different observing contexts for a given sensor) can have varying sensitivities to the minimum robustly detectable amount of cloud water. Thus, a sensor's observation characteristics and location of a pixel on the Earth can affect the proportion of pixels identified as cloudy (*ibid.*).

The other assumptions inherent to this specific algorithm are that (1) the MODIS cloud mask product MYD35 provides a sufficiently high-quality data set that it is suitable to use as a cloud mask; (2) that there is enough information in the common bands between MODIS and OCI to replicate the cloud mask with a NN; and (3) that MODIS and OCI are similar enough that transfer learning of the NN from MODIS to OCI would provide a cloud mask of sufficient quality to be usable for downstream applications. We feel that assumption (1) is validated by the wide use of the MODIS cloud mask for over 20 years, and thorough knowledge of its strengths and limitations. Assumptions (2) and (3) are evaluated through this document. One relevant factor here is that OCI's pixels are slightly larger (1.2 km vs. 1 km, at the sub-satellite points) which means that more pixels contain clouds (would increase apparent cloud cover) but edges or tenuous clouds may be less detectable (they are smeared out over the slightly larger pixel size).

One limitation of the current approach is that, by only using 11 spectral bands that MODIS and OCI observe similarly, the algorithm is not making full use of OCI's full capabilities. We expect that several spectral regions monitored by OCI (including but not limited to UV wavelengths and gas absorption features such as the O₂ A-band) will provide additional skill for cloud masking and cloud-aerosol discrimination. These will be investigated in the future.

We also note that, importantly, by training to the MODIS cloud mask we inherit its weaknesses as well as its strengths. Validation against the spaceborne CALIOP lidar and CloudSat radar (which unfortunately ended operations prior to PACE's launch) and similar airborne and ground-based systems has shown that MODIS's cloud mask has around 85-90% accuracy overall for daytime data, is most accurate for daytime over-water scenes, and can have difficulties identifying optically-thin cirrus clouds and over snow/ice (e.g. Ackerman *et al.*, 2008; Wang *et al.*, 2016; Kotarba, 2020). Although an exact optical depth threshold is hard to state because it depends on solar and view geometries and the underlying surface, Ackerman *et al.* (2008) suggested 0.4 as a guideline value because around 90% of clouds missed by MODIS in their comparison against ground-based lidar (over land) had optical depths below that.

Additionally, although it contains 'callback' tests to try and recover aerosol-laden pixels mistakenly identified as clouds, some overscreening of optically-thick smoke and dust plumes is also known to be an issue (e.g. Brennan *et al.*, 2005). This would be a problem for use of such a cloud mask as an input to aerosol property retrieval (as events of interest would be screened out). Conversely, this misclassification might be an asset for a retrieval of surface properties as optically-thick opaque aerosol plumes obscure the surface features in a similar manner to clouds and so would need to be screened out.

3.2 Algorithm description

3.2.1 Labelled data

The first step in development of a NN to emulate the MODIS cloud mask was to obtain suitable labelled data for training and validation. We used cloud mask and processing path information from the MYD35 data product, calibrated top of atmosphere (TOA) reflectance data from the MYD021KM product (which aggregates MODIS' spectral bands to a common 1 km at nadir swath), and geolocation and geometry data from the MYD03 product. Specific variables extracted from each are shown in Table 1.

We used MODIS Aqua data from the 1st and 15th of each calendar month from the year 2015. This provides a sampling of the seasonal variability of cloudiness, solar geometries, and helps keep the labelled pixels more independent of one another (as the typical cloud/weather system is shorter-lived than the ~15 day gap between sampled days). We first subset to daytime pixels (defined as a solar zenith angle smaller than 82°) and those where the cloud mask was processed. We also exclude any pixels where MODIS bands 1-7 lacked valid TOA reflectance data, which mostly affects band 6 (centered near 1.6 μm), for which around half of MODIS Aqua's detectors are dead, but also catches occasional glitches or saturation in other bands.

MODIS bands 8 and 9 (near 412 and 443 nm, respectively) saturate at radiance levels seen for optically thick clouds at moderate solar zenith angles. We do not discard these as it would lead to a sampling bias; instead, we fill any saturated values with the TOA reflectance for band 3 (near 470 nm) which has a much higher saturation limit. The rationale for this is that the bands

seldom saturate except for cases of thick cloud, for which the true spectral signal should be fairly flat (as cloud scattering is spectrally neutral across the visible and thick clouds are above the bulk of the spectrally-curved Rayleigh scattering signal), and the spectral interval from 412 to 470 nm is not too large.

Table 1. Variables extracted from MODIS files for NN training and validation

| Product | Variable | Contents |
|----------|----------------------|---|
| MYD021KM | EV_250_Aggr1km_RefSB | TOA reflectance of native 250 m MODIS bands (1 and 2) aggregated to nominal 1 km footprint. |
| | EV_500_Aggr1km_RefSB | TOA reflectance of native 500 m MODIS bands (3 to 7) aggregated to nominal 1 km footprint. |
| | EV_1km_RefSB | TOA reflectance of native 1 km MODIS bands (8, 9, 18, and 26) used. |
| MYD03 | solarazimuth | Solar azimuth angle, degrees clockwise from North. |
| | solarzenith | Solar zenith angle, degrees from a nadir view. |
| | sensorazimuth | Sensor azimuth angle, degrees clockwise from North. |
| | sensorzenith | Sensor zenith angle, degrees clockwise from a nadir view. |
| | latitude | Pixel latitude, degrees North. |
| | longitude | Pixel longitude, degrees East. |
| MYD35 | cloud_mask | Cloud mask and processing path information (surface cover type and daytime flag). |

Next, we randomly select 0.1% (one per thousand) of pixels from the remaining data set on a granule-by-granule basis. This strategy means that the training and validation data broadly reflect the same underlying distribution as the satellite samples in flight. We use a fixed random number initialization seed so that the code can be rerun to obtain the same pixel selection in case of, for example, bugs or needing to extract additional quantities. Over the full set of labelled data, this gives 5,257,292 pixels, which we split as 90% for training, 8% for validation, and 2% for evaluation. The cloud mask, collapsed to a binary cloudy/not cloudy categorization, is the label compared against the output from the NN. We take the following features as inputs to the NN, choosing based on OCI-obtainable features used by the source MODIS cloud mask and those our own expertise suggests may be useful for training:

- Solar and view geometry; specifically, the solar zenith angle, viewing zenith angle, and relative azimuth angle. The latter is defined as the solar minus sensor azimuth with 180° subtracted (a standard convention in NASA Earth science due to the reference frames), transformed to lie in the closed interval $[0^\circ, 180^\circ]$ assuming azimuthal symmetry.

- The cloud mask, collapsed to a binary cloudy/not cloudy categorization.
- The surface cover type, encoded as a one-hot array (i.e. zero or one) for categories “Water”, “Sun glint”, “Land”, “Desert”, and “Snow/Ice”. Each pixel belongs to only one category. This is inherited directly from the MODIS cloud mask processing path, except that we include pixels classified as “Coastal” in MYD35 as part of “Land” (PACE does not have a specific coastal pixel designation in its available land mask, and coastal pixels in MODIS are typically land-dominant).
- Spectral TOA reflectance in 11 bands: MODIS bands 8, 9, 3, 4, 1, 2, 5, 6, 7, 18, and 26, centered near 412, 443, 470 550, 645, 865, 936, 1240, 1640, and 2130 nm respectively.
- 3x3 pixel standard deviation of TOA reflectance at 4 bands (470, 645, 865, and 2130 nm). For granule edges, the box shrinks as necessary (to a minimum of 2x2 pixels at granule corners).

Figure 1 shows maps of where pixels in the different surface categorizations lie (aggregated to a 2.5° grid). Figures 2, 3, and 4 show histograms of solar/view geometry, TOA reflectance, and TOA reflectance standard deviation, respectively, for each surface class. The overall cloud fraction (to two decimal places) of pixels in each of the water, glint, land, desert, and snow/ice categories are 0.79, 0.61, 0.70, 0.42, and 0.25, respectively.

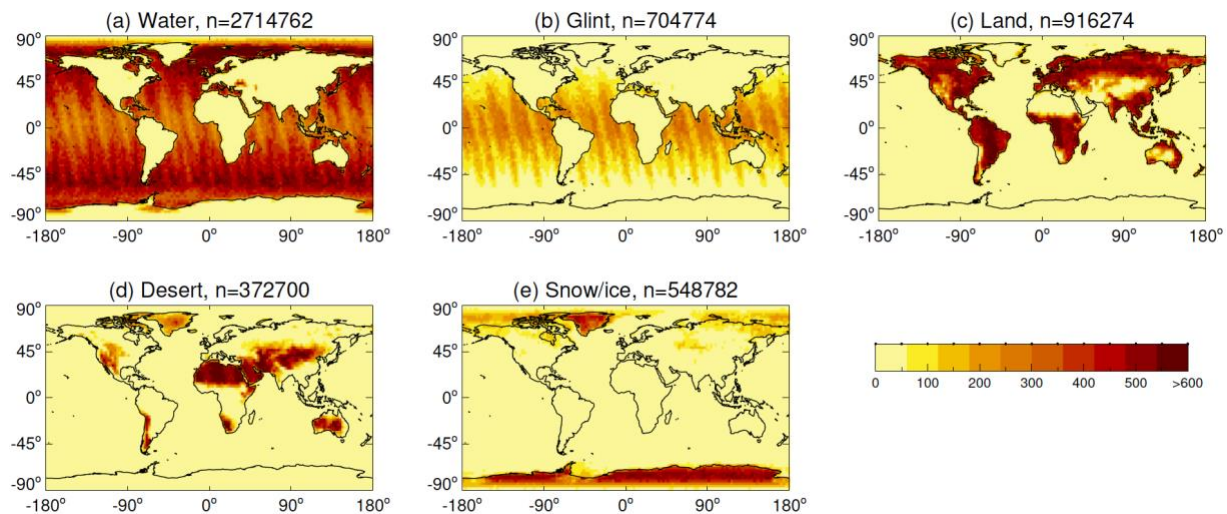


Figure 1. Geographic distribution of labelled data, split by surface type. The total number (n) of points in each category is given in the panel titles. Axes show latitude and longitude.

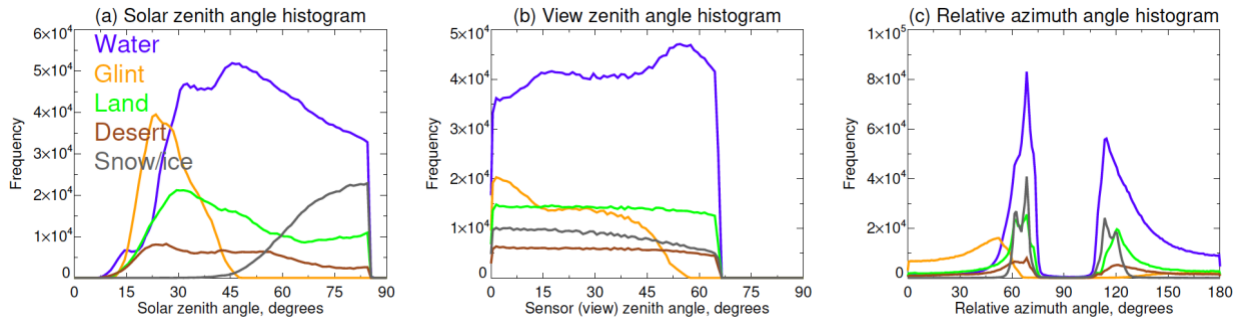


Figure 2. Histograms of solar and viewing geometry from pixels extracted as labelled data, for each surface type in Figure 1. The bin size is 1 degree.

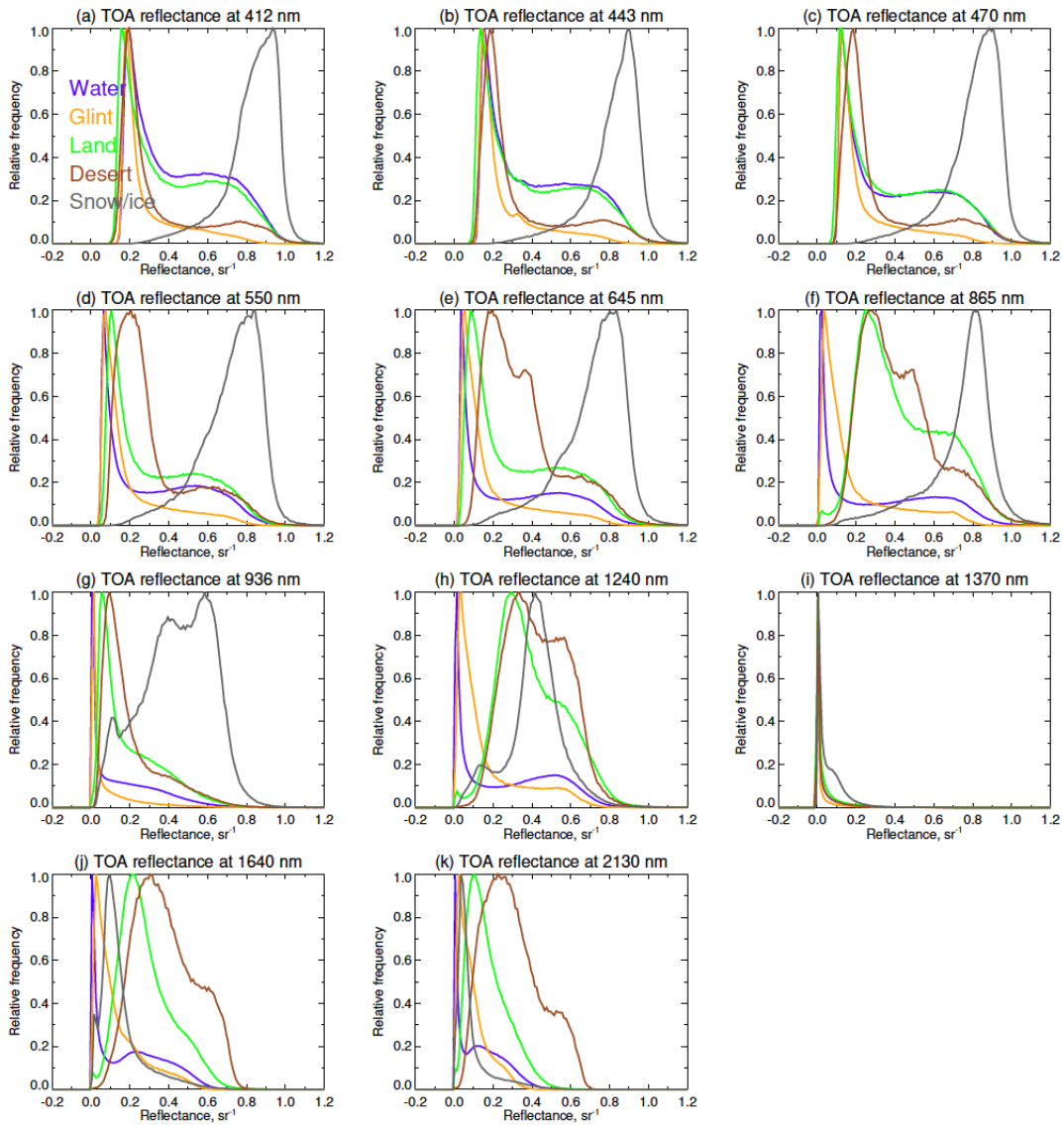


Figure 3. Histograms of TOA reflectance at 11 bands from pixels extracted as labelled data, for each surface type in Figure 1. The bin size is 0.01.

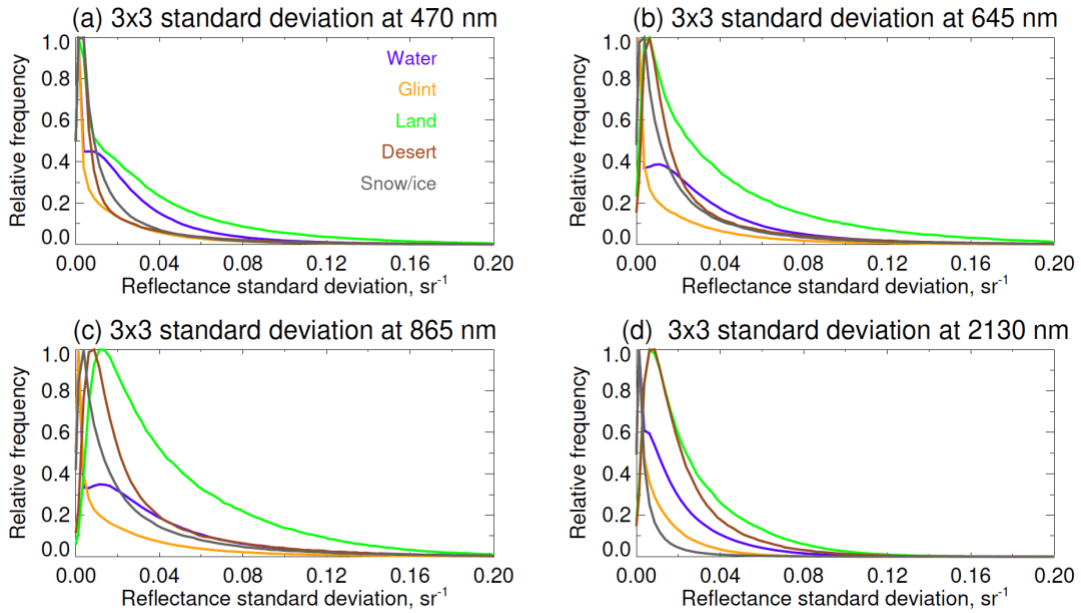


Figure 4. Histograms of TOA reflectance 3x3 pixel standard deviation from four bands from pixels extracted as labelled data, for each surface type in Figure 1. The bin size is 0.0025.

3.2.2 Neural network description and training

The model is a multi-layer perceptron (MLP) type NN with a single hidden layer having 256 nodes and a sigmoid activation function. This simple architecture had an accuracy as good or better than several more elaborate MLPs and reached our training cut-off, set as 50 epochs without decrease in loss computed on the validation data, in fewer epochs. We used the logistic loss, or the cross-entropy of the labels with the NN outputs. Training used the Adam optimizer (Kingma and Ba, 2015), a learning rate of 3×10^{-4} , and a batch size of 512. Figure 5 shows the change in loss over training epochs, i.e. sequential passes through all batches of the training data. We constructed and trained the model with Tensorflow 2.14.1.

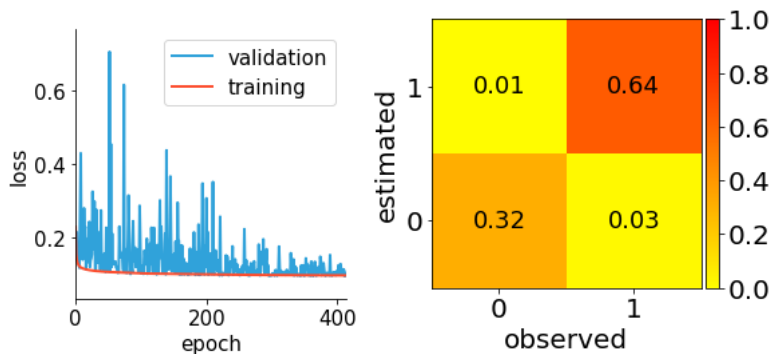


Figure 5. Cross-entropy loss versus training epoch (left) for the training (red) and validation (blue) splits of the labelled data. Confusion matrix (right) on the test split of the labelled data. In the confusion matrix 0 indicates a classification of cloud-free, and 1 of cloudy.

After training, the network achieves 96% accuracy on the withheld test subset of data overall. Split by surface type, these numbers are 97%, 95%, 96%, 96%, and 92% for water, glint, land, desert, and snow/ice, respectively.

3.2.3 Transfer learning to OCI

Running the MODIS-trained network on OCI inputs is fairly straightforward. The geometry features have direct equivalents. TOA reflectances are obtained from OCI bands close to MODIS band centers (listed in Table 2). Pixel standard deviations are computed in the same way as for MODIS. The exception is the surface categorization. The OCI L1b files contain a binary land/water mask, used to define these two classes. The water category is overwritten as glint if the HIGLINT flag is set (Wang & Bailey, 2001). Both land and water categories are overwritten as snow/ice if the land snow or sea ice (respectively) variables from the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2) reanalysis (Gelaro *et al.*, 2017), or GEOS-IT and GEOS-FP in the case of near real time processing, are above 0.5. These data sets are available at 0.5° latitude, 0.625° longitude, and 1 hour temporal resolution and are spatiotemporally interpolated to the location of each pixel. Finally, we use the MODIS-based (Friedl *et al.*, 2002) International Geosphere Biosphere Project (IGBP) surface climatology classification at 0.05° latitude-longitude resolution to overwrite land pixels as desert, for IGBP-defined deserts (“barren” surface class). We note that one limitation is that this is a static map based on observations from the year 2004.

In routine processing, the algorithm is applied to all daytime OCI pixels where the inputs are available. Due to OCI’s SWIR focal plane misalignment, the 13 rows along the western edge of each swath are not processed. Figure 6 shows an example of the source MYD35 cloud mask, the NN applied to MODIS, and the application to OCI data (from an orbit 80 minutes prior to Aqua) over parts of eastern North America and the Atlantic Ocean. The non-processed pixels are visible as a strip of grey on the western edge of the OCI swath; also apparent is OCI’s broader swath than MODIS.

We additionally produce a “dilated” cloud mask for OCI. This is set to 1 if a pixel is cloudy or any of its immediate neighbors are cloudy, and as such identifies both cloudy and cloud-adjacent pixels. This provides a stricter cloud mask for downstream algorithms that may be particularly sensitive to e.g. undetected cloud fragments.

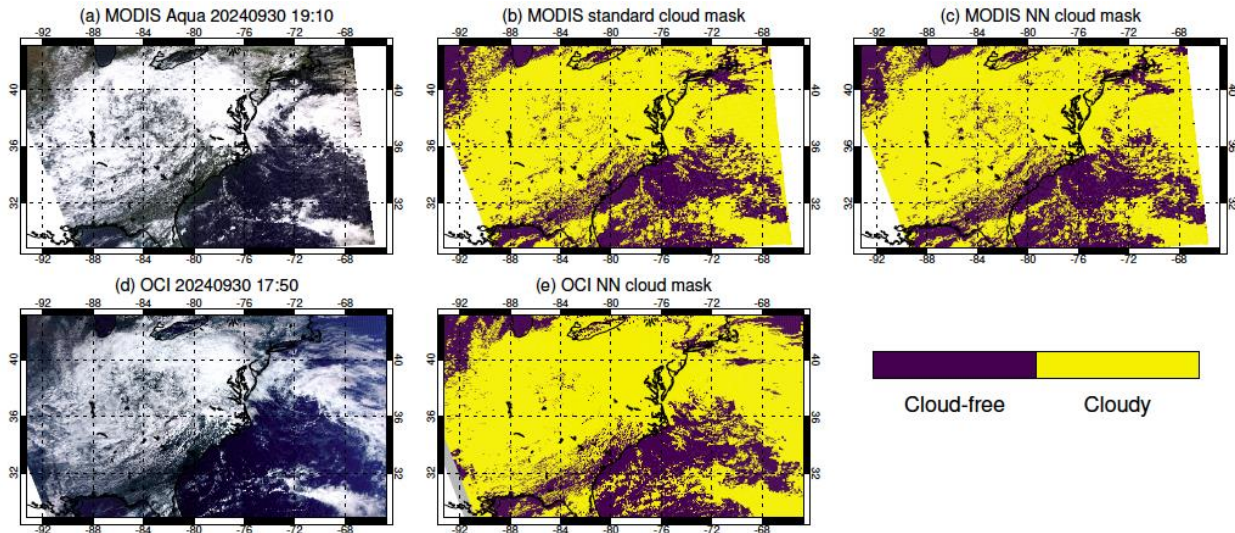


Figure 6. Example true color images (left) and cloud mask classification (middle, right) for MODIS (top) and OCI (bottom) for a scene on September 30 2024. At this zoom level only a subset of scan lines are mapped so the rows of dead detectors in the MODIS NN cloud mask image are not visible.

3.3 Algorithm Input Variables

Table 2 lists the data used as input values. Data sources are provided in Section 6.

3.4 Algorithm Output Variables

The data resulting from running the algorithm are listed in Table 3. These data sets are distributed in the PACE OCI CLDMASK suite.

4 Algorithm Usage Constraints

The algorithm is based on MODIS inputs and we chose bands with a close analogue in PACE OCI. As a result, the algorithm as implemented can't be directly applied to other sensors, although by adjusting the MODIS bands used similar networks could be trained for other passive imaging radiometer instruments. The algorithm is applicable to all daytime PACE OCI pixels, regardless of underlying surface type or cloud cover. It is not applicable where one or more of the required bands are not available, for example the 13 rows along the western edge of each swath where OCI's SWIR bands are not available due to focal plane misalignment.

Table 2. Input variables required for OCI data processing. OCI variable names reference the data set names as stored in OCI level 1B files (which combine data from multiple wavelengths by focal plane). Band indices count from zero.

| Name | Long name | Unit |
|----------------------|---|------------------|
| rhot_blue index 41 | TOA reflectance at 412.5 nm | sr ⁻¹ |
| rhot_blue index 53 | TOA reflectance at 442.3 nm | sr ⁻¹ |
| rhot_blue index 64 | TOA reflectance at 469.7 nm | sr ⁻¹ |
| rhot_blue index 96 | TOA reflectance at 549.9 nm | sr ⁻¹ |
| rhot_red index 20 | TOA reflectance at 644.7 nm | sr ⁻¹ |
| rhot_red index 150 | TOA reflectance at 864.6 nm | sr ⁻¹ |
| rhot_swir index 0 | TOA reflectance at 939.7 nm | sr ⁻¹ |
| rhot_swir index 2/3 | TOA reflectance at 1250.4 or 1248.6 nm dependent on gain saturation | sr ⁻¹ |
| rhot_swir index 4 | TOA reflectance at 1378.2 nm | sr ⁻¹ |
| rhot_swir index 5/6 | TOA reflectance at 1619.6 or 1618.0 nm dependent on gain saturation | sr ⁻¹ |
| rhot_swir index 7 | TOA reflectance at 2130.6 nm | sr ⁻¹ |
| solar_zenith | Solar zenith angle | ° |
| sensor_zenith | View zenith angle | ° |
| solar_azimuth | Solar azimuth angle | ° |
| sensor_azimuth | View azimuth angle | ° |
| watermask | Land/water flag | - |
| L2_flags | Ocean Colour level 2 processing flags | - |
| IGBP_Land_Cover_Type | IGBP surface type classification dataset | - |
| FRSNO | MERRA2 fractional snow cover over land | - |
| SEAICE | MERRA2 fractional sea ice cover over water | - |

Table 3. *Output quantities provided within the level 2 PACE OCI CLDMASK file.*

| Name | Long name | Unit |
|--------------------|--|------|
| cloud_flag | Cloud mask | - |
| cloud_flag_dilated | Cloud or cloud-adjacent pixel mask | - |
| l2_flags | L2 flags (giving information about processing path and status) | - |

5 Performance Assessment

At present, we have performed limited evaluation of the approach by comparing to the source MODIS Aqua cloud mask (MYD35) that the neural network was trained on. To this end, we have processed MODIS data from Jan 1, 2024 to Feb 28, 2025 with the algorithm. The first part of our evaluation is a direct pixel-level comparison between the MODIS NN and source MYD35. This also provides approximately one calendar year of overlap since the start of routine PACE science data flow (Mar 5, 2024).

As a direct pixel-level comparison between OCI and MODIS is not possible (they are on different satellites with different orbital characteristics), for the second part of our evaluation we have focused on global comparisons of fractional cloud cover. To calculate this, we averaged pixels to a 1° equal angle grid on a daily basis, and report the fraction of processed pixels which were identified as cloudy (colloquially referred to as “cloud fraction” and commonly reported in satellite level 3 products). For MODIS, we selected only those pixels processed by both MYD35 and our MODIS NN. In order to decrease sampling-related discrepancies, we only compare grid cells seen by both MODIS and OCI on a given day with at least 100 pixels contributing to the MODIS aggregate. We note pixel counts are not available within the OCI level 3 cloud fraction product.

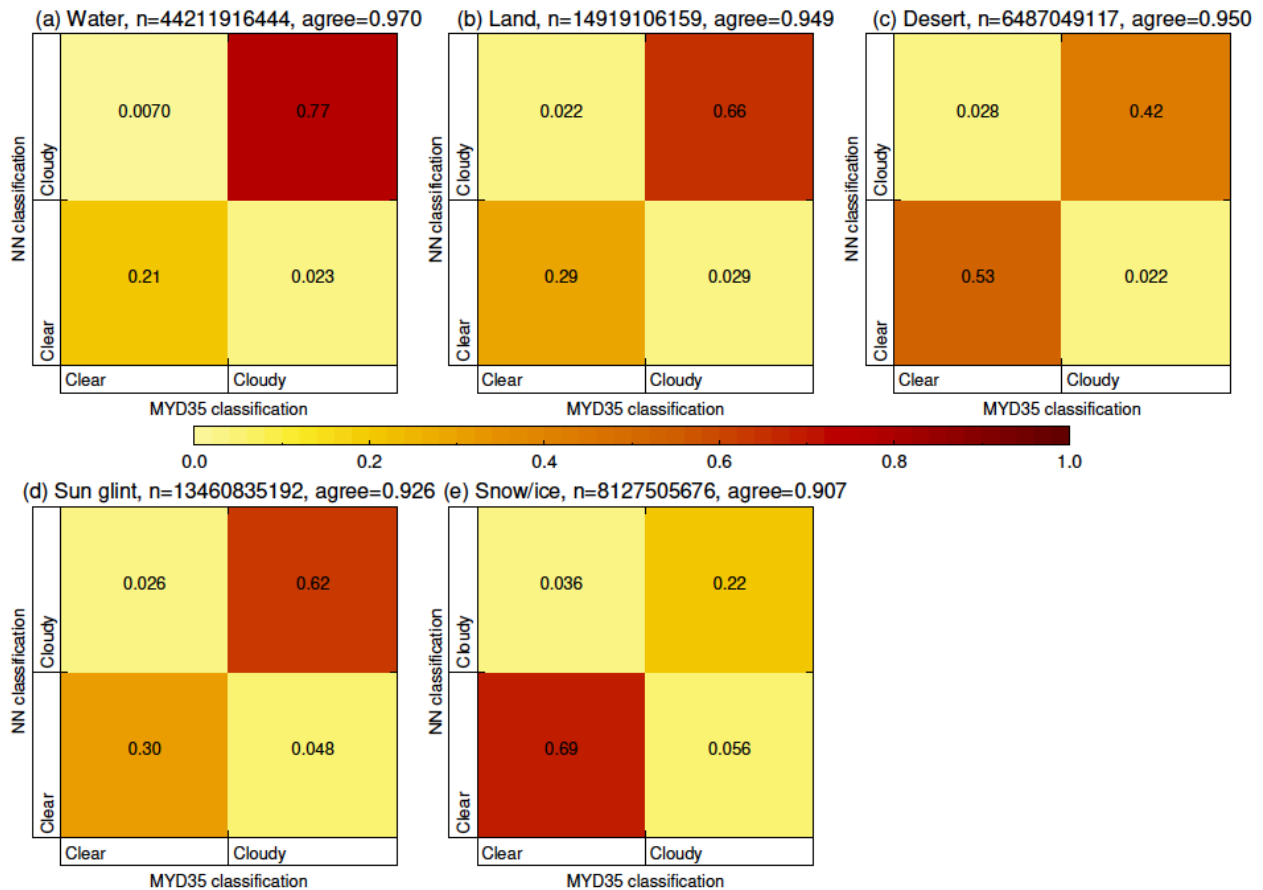


Figure 7. Confusion matrices for each surface categorization for common pixels processed by MYD35 and the MODIS NN, for all daytime granules from January 2024 to February 2025. Colors indicate the relative abundance of points in each category.

Figure 7 is a summary of the performance of the NN against MYD35, applied directly to MODIS observations from January 2024-February 2025. Fractional accuracies for the surface classes are 0.970, 0.949, 0.950, 0.926, and 0.907, for water, land, desert, Sun glint, and snow/ice, respectively. Overall classification accuracy is 0.952. In all classes there is a tendency for the NN to detect slightly fewer clouds than seen in MYD35 (i.e. more errors of omission than of commission). This is expected, because the NN does not use the thermal bands which provide skill particularly for clouds over snow and Sun glint (where surface-cloud contrast in solar reflectances is lower) and optically-thin clouds. This explains the slightly lower accuracy over Sun glint and snow/ice than other surfaces.

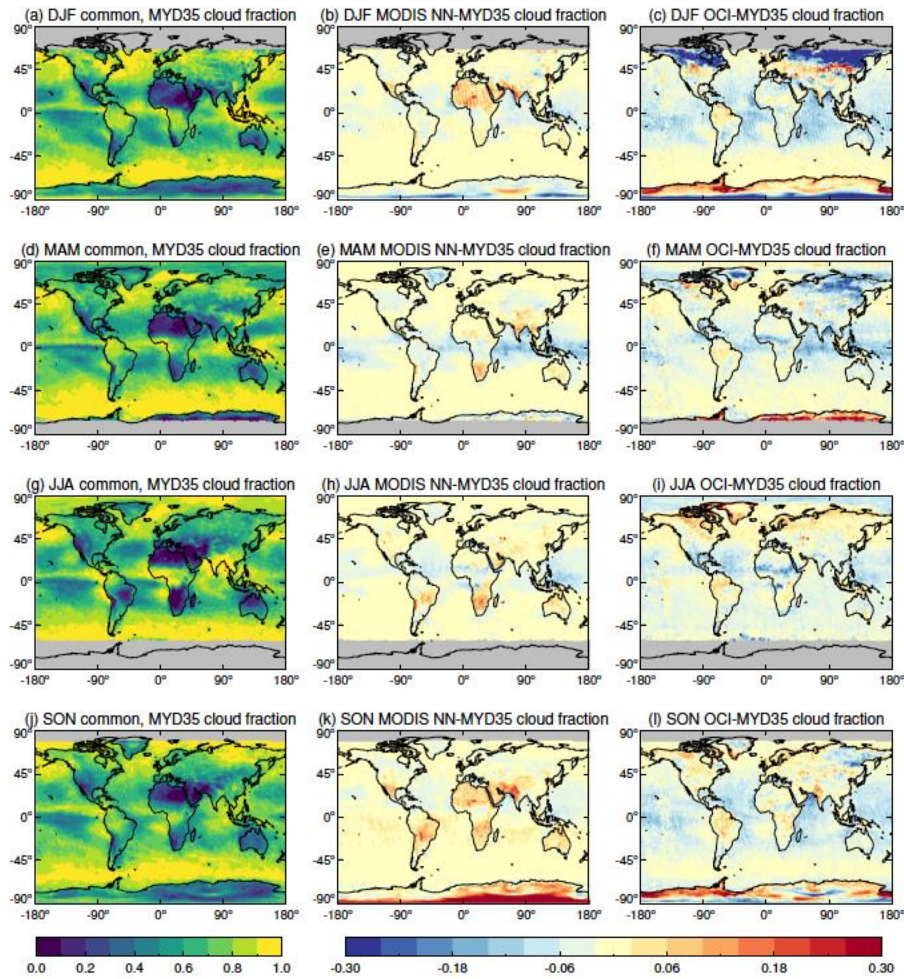


Figure 8. Seasonal (rows top to bottom) comparison of cloud fractions from MYD35, the NN applied to MODIS, and the NN applied to OCI, for March 2024 to February 2025. The left column shows the MYD35 cloud fraction. The middle and right columns show the MODIS NN – MYD35, and NN applied to OCI – MYD35, respectively. All data shown at 1° only considering grid cells commonly sampled on a daily basis. Grid cells with fewer than 10 days contributing in a given season are shaded in grey.

A mapped comparison of cloud cover is shown in Figure 8. For commonly-sampled MYD35 and MODIS NN data (middle column), in many cases the overall cloud fraction difference is within ± 0.03 . More negative differences (false negatives) tend to be associated with tropical oceans. More positive differences (false positives) tend to be associated with deserts, and with times and locations of heavy aerosol loading. This suggests that the aforementioned MODIS tendency for misidentification of optically thick aerosols as clouds could be exacerbated in the NN. Larger cloud fraction differences are also seen over polar land.

Extending this comparison to OCI (right column), the picture is a little more muddled. Overall the OCI-MYD35 difference is more negative, which would be consistent with cloud edges and

fragments being detected slightly less often in OCI's slightly coarser pixels; a 1.2 vs. 1 km nominal pixel corresponds to about a 44% increase in pixel area for OCI. There are, however, also areas where OCI detects more clouds. This could be different cloud types becoming more detectable at OCI's coarser resolution (the converse of the previous argument) or could be related to differences between MODIS and OCI spectral response functions becoming important in some cases. Differences are also larger over snow-covered surfaces (e.g. northern high latitudes during boreal winter, and both poles), implying difficulties over snow.

It is also possible that some of these differences are sampling noise as there are maximum ~90 days in a season contributing to the comparisons. This is not an issue for the MYD35 to MODIS NN comparison because those use the same source pixels so there is no sampling difference. At this time, the reasons are unclear, and we will revisit the comparison when additional months of data are available and sampling-related differences would be expected to diminish.

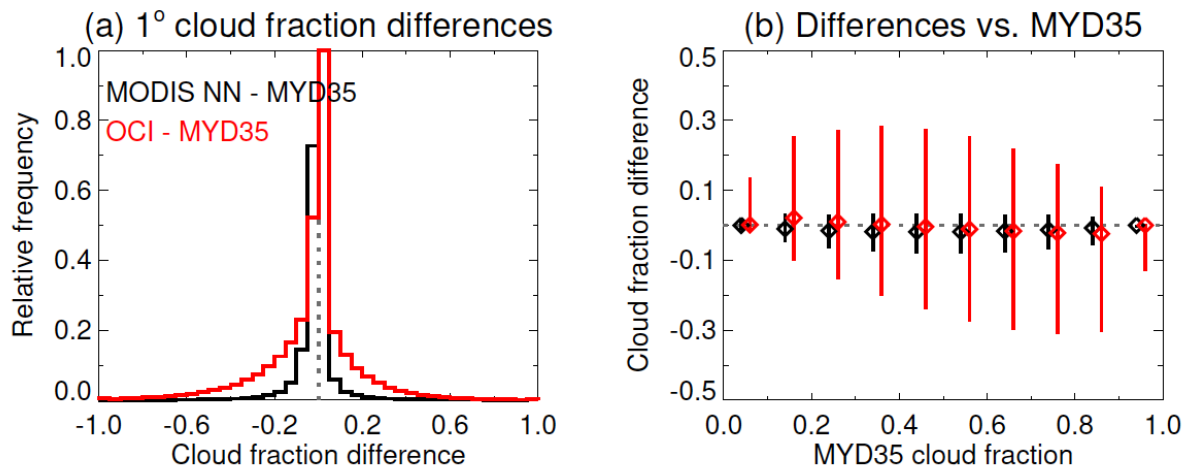


Figure 9. Differences between daily cloud fraction at 1° from the NN applied to MODIS (black) and OCI (red), relative to MYD35. Panel (a) shows a simple difference histogram with a bin size of 0.05. Panel (b) shows median (diamonds) and central one standard deviation of data (vertical lines) as a function of MYD35 cloud fraction with a bin size of 0.1. In (b), the MODIS NN and OCI data are offset slightly along the x axis for visual clarity.

Figure 9 takes a closer look at cloud fraction from individual 1° daily grid cells. For the NN applied to MODIS, the modal difference is in the bin -0.05 to 0 (small and negative), and almost all grid cells have cloud fractions agreeing within ± 0.1 . For OCI, the mode is from 0 to 0.05 (small and positive) and there are more significant positive and negative tails in the distribution. As discussed previously, some of this is due to differences in cloud structure between the (non-coincident) MODIS and OCI overpasses, which is not an issue for the MODIS NN – to – MYD35 comparison. Figure 9(b) shows that the MODIS NN minus MYD35 difference is independent of MYD35 cloud fraction. For OCI, differences are very close to zero for both fully/almost clear (MYD35 cloud fraction under 0.1) and fully/almost cloudy (MYD35 cloud fraction over 0.9) grid cells. For grid cells containing broken cloudiness, there is a tendency for a positive offset where

cloud fraction is below 0.5 and a negative offset where cloud fraction is above 0.5. This indicates some likely dependence on the structure of the underlying cloud fields.

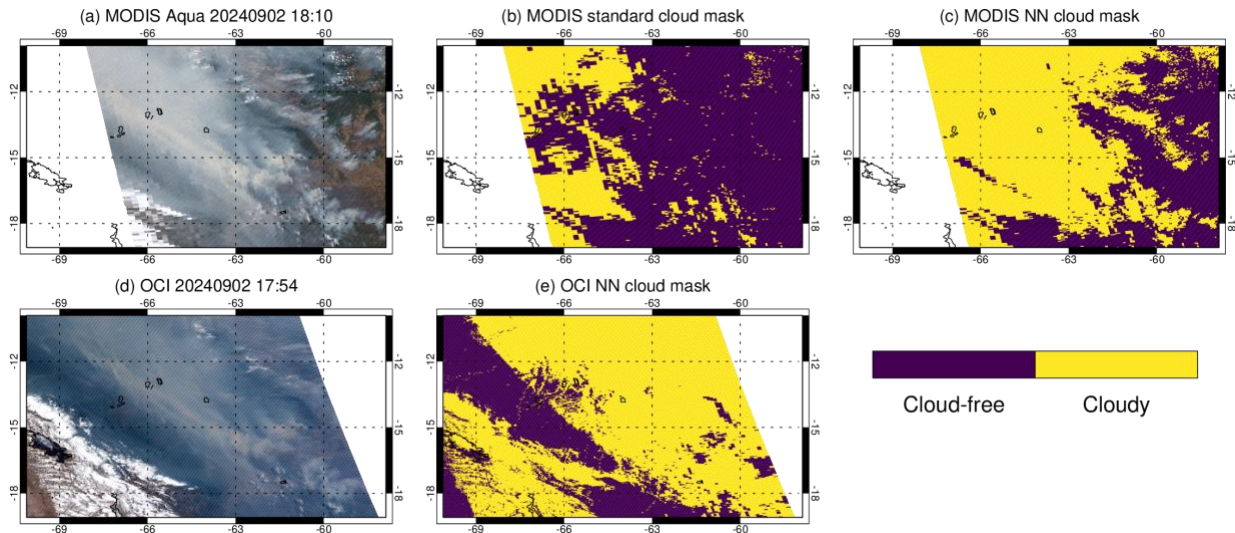


Figure 10. As Figure 6, except for a scene over South America from September 2 2024. A thick plume of smoke is seen traveling south through the image, with smaller isolated fires to the east and the Andes mountains to the west.

Here, we show some case studies from individual orbits of data. Figure 10 shows MODIS and OCI scenes during the Amazonian biomass burning season, from September 2 2024. A dense plume of wildfire smoke travels against the Andes, with additional small fire plumes to the eastern part of the region. Some of the thick smoke layer has clouds embedded within it. In this case, none of the cloud masks appear perfect. The MYD35 standard cloud mask (Figure 10(b)) captures the embedded clouds, but overscreens some of the dense smoke. The NN applied to MODIS has a more severe overscreening problem. The NN applied to MODIS also overscreens much of the thick smoke, but correctly identifies some of the optically-thinner smoke nearer the Andes which MODIS incorrectly flags. We note that some of the differences may be due to the differing geometries of the sensors for this scene (on the western edge of the MODIS swath and eastern edge of OCI). Additionally, while difficult to tell from this example, the OCI retrievals often show a false “border” of cloudy pixels around coastlines and along rivers. This is likely due to their higher spatial variability of reflectance (a characteristic of clouds) and lack of a “coastal” category in the current version of this algorithm.

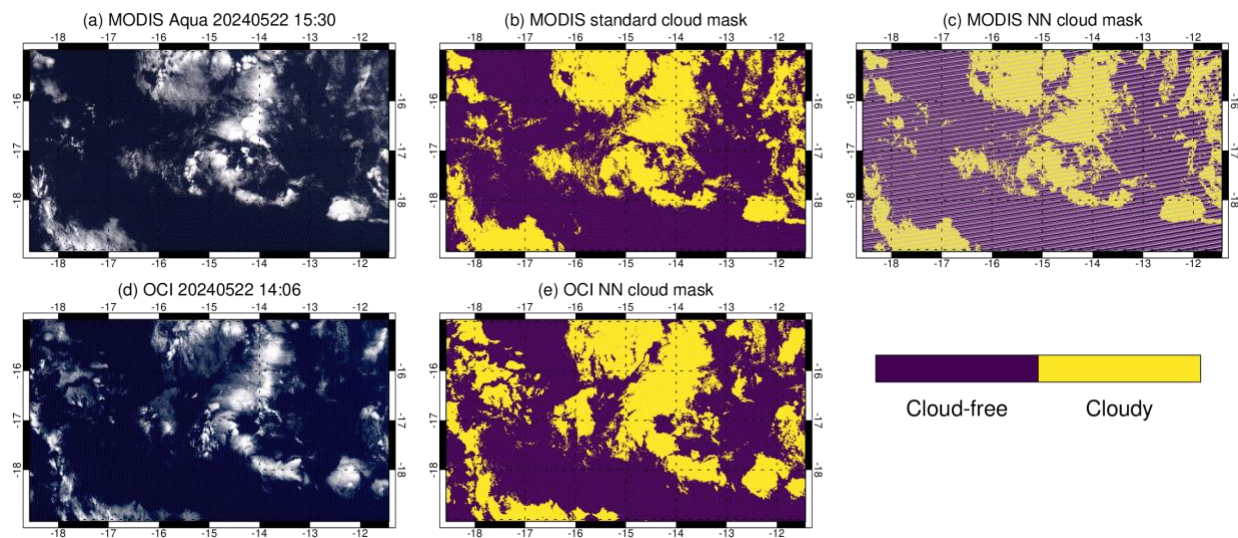


Figure 11. As Figure 6, except for a scene over the tropical Atlantic Ocean from May 2 2024. Missing data (grey) in (c) are due to dead detectors on MODIS Aqua (not visible in other images due to the level of zoom in this scene).

Figure 11 shows a scene over the tropical Atlantic Ocean. Here, all algorithms appear to perform visually well, yet in the 84 minutes between the PACE and OCI overpasses the clouds have moved. This means that, on a daily 1° scale, the cloud fractions are quite different. This is an example of a real physical change related to sampling which results in a disagreement between retrievals from the two sensors and is a limitation of gridded analysis with non-coincident data. We note that for this scene, the two instruments observed with similar near-center-of-swath geometries.

6 Algorithm Implementation

6.1 Algorithm Availability

The code used in algorithm development is not yet publicly available. We are currently refining the approach and intend to make the code public and submit a peer-reviewed article once the data and algorithm are stable. The code used in routine OCI data processing, and documentation, are distributed within the Ocean Colour Science SoftWare (OCSSW) package: <https://oceandata.sci.gsfc.nasa.gov/ocssw>

6.2 Input Data Access

The current version 3.0 PACE OCI level 1b files (doi:10.5067/PACE/OCI/L1B/SCI/3), which are freely available via NASA Earthdata at <https://www.earthdata.nasa.gov/>, are the input data used to run this algorithm.

6.3 Output Data Access

The output PACE OCI CLDMASK product is available at level 2 via NASA Earthdata (doi:10.5067/PACE/OCI/L2/CLOUD_MASK/3.1). It is also available at level 3 resolution within the OCI CLOUD level 3 suite (doi: 10.5067/PACE/OCI/L3M/CLOUD/3.1). These data are distributed by the NASA OB.DAAC.

6.4 Important Related URLs

As well as Earthdata search, data can be queried and downloaded via the OB.DAAC at <https://oceancolor.gsfc.nasa.gov/data/find-data/> directly.

7 Open Research

At the present time, data and software used in developing this algorithm are available from the authors on reasonable request. As noted earlier, the approach is in refinement and when it is more stable and has undergone more evaluation the authors intend to submit a peer-reviewed publication detailing the approach. At this point data and code will be made freely available. The code necessary to run the algorithm on OCI, and the routinely-generated data products resulting from this, are already freely available as discussed in Section 6.

8 Acknowledgements

The authors are grateful to Taha Al-Nufaili, then an undergraduate student at the University of Maryland, College Park who prototyped the original iteration of this concept under co-authors Sayer and Carroll during his summer 2022 internship. The work of the authors was supported by PACE Project Science.

Contact Details

Sayer, Andrew M.

andrew.sayer@nasa.gov

Conceptualization, formal analysis, investigation, methodology, project administration, software, resources, supervision, validation, visualization, writing – original draft, writing – review & editing

Goddard Earth Sciences Technology and Research (GESTAR) II, University of Maryland
Baltimore County, Baltimore, MD, USA

Ocean Ecology Laboratory, Mail code 616, NASA Goddard Space Flight Center,
Greenbelt, MD, USA

Carroll, Ian T.

ian.t.carroll@nasa.gov

Conceptualization, data curation, formal analysis, investigation, methodology, project administration, software, resources, supervision, validation, visualization, writing – review & editing

Goddard Earth Sciences Technology and Research (GESTAR) II, University of Maryland Baltimore County, Baltimore, MD, USA

Ocean Ecology Laboratory, Mail code 616, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Semenov, Alexandr V.

alexandr.v.semenov@nasa.gov

Data curation, software, resources, writing – review & editing

Science Application International Corp, Lanham, MD, USA

Ocean Ecology Laboratory, Mail code 616, NASA Goddard Space Flight Center, Greenbelt, MD, USA

References

Ackerman, S. A., Holz, R. E., Frey, R., Eloranta, E. W., Maddux, B. C. and McGill, M. (2008), Cloud Detection with MODIS. Part II: Validation. *J. Atmos. Oceanic Technol.*, 25, 1073–1086. doi:10.1175/2007JTECHA1053.1

Barnes, W. L., X. Xiong, X., and Salomonson, V. V. (2003), Status of Terra MODIS and Aqua MODIS, *Adv. Space Res.*, 32 (11), 2099–2106. doi:10.1016/S0273-1177(03)90529-1

Baum, B. A., Menzel, W. P., Frey, R. A., Tobin, D. C., Holz, R. E., Ackerman, S. A., Heidinger, A. K., and Yang, P. (2012), MODIS Cloud-Top Property Refinements for Collection 6. *J. Appl. Meteor. Climatol.*, 51, 1145–1163. doi:10.1175/JAMC-D-11-0203.1

Brennan, J. I., Kaufman, Y. J., Koren, I., and Li, R.-R. (2005), Aerosol-cloud interaction-Misclassification of MODIS clouds in heavy aerosol. *IEEE Trans. Geosci. Remote Sens.*, 43(4), 911. doi:10.1109/TGRS.2005.844662

Di Girolamo, L., & R. Davies (1997), Cloud fraction errors caused by finite resolution measurements, *J. Geophys. Res.*, 102(D2), 1739–1756, doi:10.1029/96JD02663

Frey, R. A., Ackerman, S. A., Liu, Y., Strabala, K. I., Zhang, H., Key, J. R., and Wang, X. (2008), Cloud Detection with MODIS. Part I: Improvements in the MODIS Cloud Mask for Collection 5. *J. Atmos. Oceanic Technol.*, 25, 1057–1072. doi:10.1175/2008JTECHA1052.1

Friedl, M. A., McIver, D. K., Hodges, J.C.F, Zhang, X.Y, Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., and Schaaf, C. (2002), Global land cover mapping from MODIS: algorithms and early results. *Remote Sens. Environ.*, 83, 287–302. doi:10.1016/S0034-4257(02)00078-0

Gelaro, R. and Coauthors (2017), The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Climate*, 30, 5419–5454, doi:10.1175/JCLI-D-16-0758.1

Jones, A. L., Di Girolamo, L., and Zhao, G. (2012), Reducing the resolution bias in cloud fraction from satellite derived clear-conservative cloud masks. *J. Geophys. Res.*, 117, D12201. doi:10.1029/2011JD017195

Kingma, D. P., J. Ba. 2015. Adam: A Method for Stochastic Optimization. International Conference for Learning Representation. doi:10.48550/arXiv.1412.6980

Koren, I., Oreopoulos, L., Feingold, G., Remer, L. A., and Altaratz, O. (2008), How small is a small cloud? *Atmos. Chem. Phys.*, 8, 3855–3864. doi:10.5194/acp-8-3855-2008

Kotarba, A. Z. (2020), Calibration of global MODIS cloud amount using CALIOP cloud profiles. *Atmos. Meas. Tech.*, 13, 4995–5012. doi:10.5194/amt-13-4995-2020

Meister, G. Knuble, J. J., Gliese, U., Bousquet, R., Chemerys, L. H., Choi, H., Eplee, R. E., Estep, R. H., Gorman, E. T., Kitchen-McKinley, S., Kubalak, D., Lee, S., McClain, C., McIntire, J. W., Patt, F. S., Rhodes, Z. and Werdell, P. J. (2024), The Ocean Color Instrument (OCI) on the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) Mission: System Design and Prelaunch Radiometric Performance, *IEEE Trans. Geosci. Remote Sens.*, 62, 5517418. doi:10.1109/TGRS.2024.3383812

Wang, M., & Bailey, S. W. (2001), Correction of sun glint contamination on the SeaWiFS ocean and atmosphere products. *Appl. Opt.* 40, 4790–4798. doi:10.1364/AO.40.004790

Wang, T., E. J. Fetzer, S. Wong, B. H. Kahn, and Q. Yue (2016), Validation of MODIS cloud mask and multilayer flag using CloudSat-CALIPSO cloud profiles and a cross-reference of their cloud classifications. *J. Geophys. Res. Atmos.*, 121, 11,620–11,635, doi:10.1002/2016JD025239