# PACE OCI Trace Gas Retrievals

**Zachary Fasnacht[1], Joanna Joiner[2], Eric Bucsela[3], Matthew Bandel[1], Fei Liu[4], Lok Lamsal[5], and Nickolay Krotkov[2]**

[1]Science & Technology Corporation
[2]NASA Goddard Space Flight Center
[3]Astralis Technologies
[4]GESTAR, Morgan State University
[5]Bureau of Ocean Energy Management

Corresponding author: Zachary Fasnacht (zachary.t.fasnacht@nasa.gov)

**Version:** 1.0
**Release Date:** 2026/02/19
**DOI:** TBD

## Abstract

Satellite based observations of trace gases such as nitrogen dioxide ($NO_2$) and ozone ($O_3$) are important for monitoring of air quality and improving regional chemistry transport models. Additionally, accounting for trace gases is important for atmospheric corrections necessary for retrieval of ocean color, clouds, aerosols, clouds and more. While not designed for the retrieval of trace gases, the spectral range of PACE OCI covers the absorption window of several trace gases including $NO_2$ (405-465 nm) and $O_3$ (Huggins Bands ~310-360 nm).While traditional approaches for trace gas retrievals can be computationally expensive and require significant effort to develop, machine learning methods can be more efficient and easier to develop. Here we utilize an algorithm that uses principal component analysis (PCA) along with machine learning (ML) to retrieve trace gases for PACE OCI.

## Plain Language Summary

This PACE OCI algorithm estimates two important atmospheric gases, nitrogen dioxide ($NO_2$) and ozone ($O_3$). Both gases play an important role in air quality as well as Earth's climate. The method applies a machine learning model to the light reflected by the Earth and measured by the satellite. To train and test the algorithm, retrievals from the TROPOMI instrument are used, which has been extensively validated. This allows for a reliable $NO_2$ and $O_3$ retrieval from PACE OCI, even though the instrument was not designed for trace gas retrievals.

**Keywords: Trace Gases, Nitrogen Dioxide, Ozone, PACE OCI**

## Version Description

This is version 1.0 of the PACE OCI Trace Gas ATBD

# 1 Introduction

In the 1970's the Backscatter Ultraviolet Instrument was launched onboard the Nimbus-4 satellite to provide the first measurements of ozone ($O_3$) from space (Mateer and Hudson 1973). Satellite retrievals of nitrogen dioxide ($NO_2$) began in the mid 1990's with the launch of the Global Ozone Monitoring Experiment (GOME) (Richter and Burrows 2002). Over the past few decades, spaceborne measurements of $NO_2$ and $O_3$ have advanced substantially in spatial resolution, with instruments such as the Ozone Monitoring Instrument (OMI, spatial resolution of 12x24 $km^2$ at nadir), the Tropospheric Monitoring Instrument (TROPOMI, spatial resolution 3.5x5.5 $km^2$ at nadir), and more recently the Geostationary Environment Spectrometer (GEMS, spatial resolution of 7x8 $km^2$ at center of field of regard) and the Tropospheric Emissions: Monitoring of Pollution (TEMPO, spatial resolution of 2x4.75 $km^2$ at center of field of regard).

Several approaches are commonly used for trace gas retrievals from these atmospheric instruments. One of the most widely applied techniques is differential optical absorption spectroscopy (DOAS) (Burrows et al. 1999, Richter and Burrows 2002). In the DOAS method, narrow-band spectral structures are fitted to measured radiances to derive the slant column density (SCD), which is then converted to vertical column density (VCD) using an air mass factor (AMF). The AMF depends on multiple parameters, including surface reflectivity, *a priori* vertical profiles, surface pressure, and cloud properties.

An alternative approach, particularly for $O_3$, is a Total Ozone Mapping Spectrometer (TOMS)-type algorithm, which relies on ratios of UV radiances at different wavelengths and forward model radiative transfer calculations to estimate total column $O_3$ (Krueger 1989, McPeters et al. 1996).

Although the PACE OCI instrument was not originally designed for atmospheric trace gases, its spectral coverage in the UV and visible regions include wavelengths where key absorbers, such as $NO_2$ and $O_3$, exhibit strong absorption features. Compared to traditional atmospheric trace gas instruments, PACE OCI provides significantly higher spatial resolution, despite the coarser spectral resolution. As shown in Figure 1 (Fasnacht et al. 2025), the cross section of $NO_2$ in the visible contains many fine-scale spectral structures. While many of these features are smoothed at the PACE OCI spectral resolution, several prominent features are still detectable at the OCI resolution.

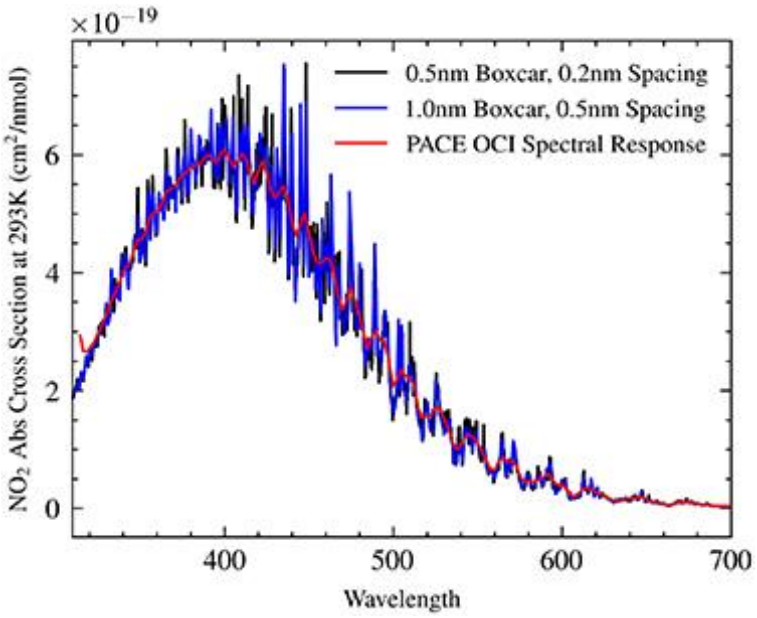


Fig. 1 Cross section of $NO_2$ at various spectral resolutions (Fasnacht et al. 2025)

To further assess the potential of using PACE OCI for trace gas retrievals, Joiner et al. 2023 simulated OCI measurements by convolving OMI radiances with the OCI spectral response function. As discussed previously, traditional trace gas retrieval methods involve computationally intensive forward model and spectral fitting. Alternatively, machine learning offers a more efficient approach by leveraging high quality retrievals from existing satellites as training data. In this approach, measured OCI reflectances containing multiple independent spectral signatures are decomposed into linearly independent principal components (PCs) that separate these spectral features. A neural network is then trained to learn the relationship between the coefficients of the PCs and physically meaningful quantities such as $NO_2$.

## 2 Algorithm Description

### 2.1 Scientific Theory

After the launch of PACE OCI, Fasnacht et al. (2025) further developed the approach described by Joiner et al. (2023) so it could be directly applied to PACE OCI. That approach was initially developed for $NO_2$ retrievals over land but has since been expanded to include $NO_2$ over water (further described in 3.1.1) and $O_3$ globally (further described in 3.1.3).

2.1.1 Training Data

To train the neural network, the extensively-validated TROPOMI trace gas retrievals (van Geffen 2020, van Geffen 2022) are co-located with the PACE OCI pixels. Since TROPOMI is spatially coarser than OCI, the OCI reflectances are averaged over the larger TROPOMI pixel. Additionally, due to short lifetimes of and dynamic transport of trace gases, we restrict the co-location to only include retrievals where PACE OCI and TROPOMI are within +/-2 hours of each other.

2.1.2 $NO_2$ Retrievals

For land pixels, the neural network is trained to estimate total vertical column (VCD) of $NO_2$ directly from TROPOMI observations. In traditional approaches, the retrieval is performed in two steps: (1) estimating the slant column density (SCD) from spectral fitting, and (2) applying an air mass factor (AMF) to convert the SCD into a vertical column. Our PCA+ML framework avoids this separation: for land scenes, it retrieves the VCD directly with performance comparable to the traditional two-step SCD+AMF method. As shown in Figure 2, given the decomposed principal components along with solar angles, the model is trained to learn the relationship between these inputs and the co-located TROPOMI $NO_2$ retrievals.

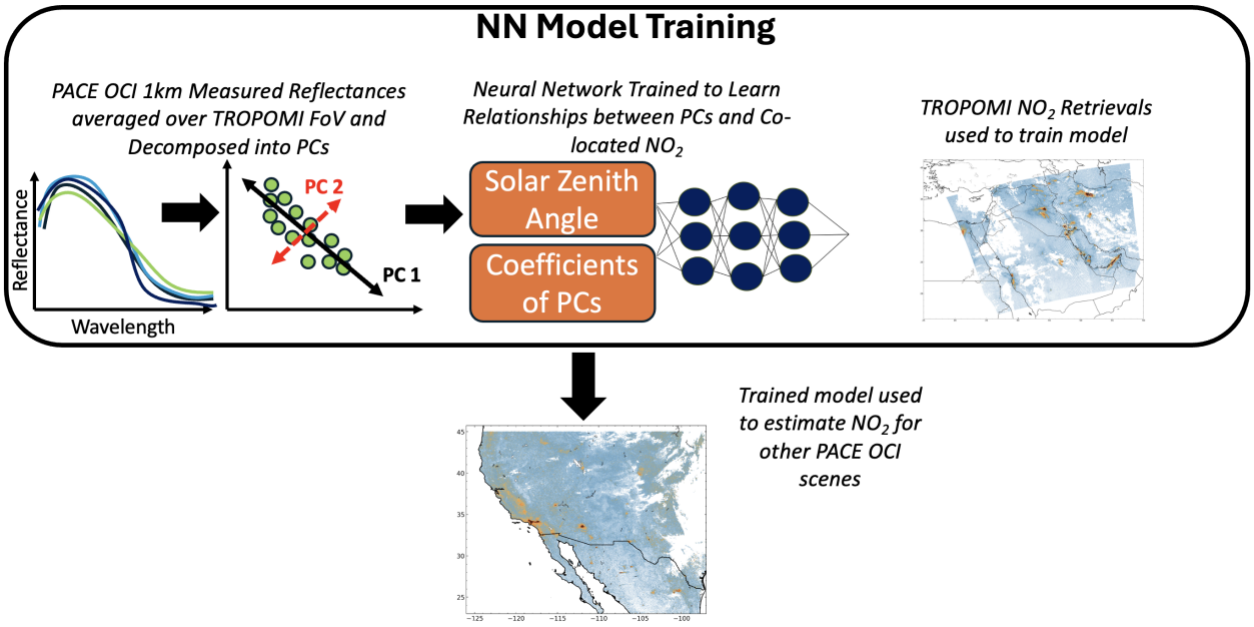


Figure 2. Flowchart showing machine learning approach to estimate $NO_2$ from PACE OCI

For water pixels, the situation is different. Over water, surface reflectance has a strong angular dependence, and additional spectral absorbers such as chlorophyll complicate the retrieval in the same UV/visible windows as $NO_2$. Because of these factors, direct estimation of the VCD is less reliable. Instead, the algorithm retrieves the SCD and AMF separately. Since the SCD is angular-dependent, to co-locate from TROPOMI, we need to utilize the geometric VCD (VCDgeo) which is defined as:

$$VCDgeo = \frac{SCD}{AMFgeo}$$

where

$$AMFgeo = \frac{1}{COS(SZA)} + \frac{1}{COS(VZA)}$$

Here, SZA is the solar zenith angle and VZA is the viewing zenith angle. The geometric VCD is purely geometric: it does not account for effects such as surface reflectivity, atmospheric scattering, or the vertical distribution of $NO_2$ (e.g., boundary-layer sensitivity).

To obtain the retrieved VCD, an effective AMF is needed. In this approach, the AMF is computed from the ratio:

$$AMF = \frac{VCDgeo}{VCD}$$

Here, the AMF is instead used to adjust the geometric VCD. The reason is that co-located TROPOMI SCDs are not directly suitable for PACE OCI, since the SCD depends on the specific viewing geometry of each instrument.

While this approach performs reasonably well, work is ongoing to develop a model that estimates the AMF directly from TROPOMI. This would allow PACE OCI to retrieve a slant column using a “true” AMF in line with traditional methods. Such an improvement is planned for a future release.

2.1.3 Ozone Retrievals

Whereas $NO_2$ concentrations are largest near the surface, most $O_3$ resides in the upper troposphere and stratosphere**.** Consequently, a DOAS-style separation of slant column and air mass factor is not required for $O_3$ over water or land scenes. Instead, the neural network is trained to estimate the total vertical column $O_3$ directly, using co-located TROPOMI $O_3$ retrievals as reference.

PACE OCI observes two major $O_3$ absorption regions: the Huggins band (310–350 nm) and the Chappuis band (450–700 nm). We focus on the Huggins band, since the TROPOMI $O_3$ product used for training is derived from this spectral range. Although not exploited here, OCI's Chappuis band coverage offers unique potential for future tropospheric $O_3$ retrievals.

2.1.4 Scientific Theory Assumptions

The algorithm makes several important assumptions:

1. The TROPOMI $NO_2$ and $O_3$ data used for training are of high quality as uncertainty from the training data will be learned by the model
2. The PCA decomposition and training of the NN model contain a comprehensive set of samples that cover the distribution of possible retrievals

## 2.2 Algorithm Input Variables

**Table 2.** *Inputs to PACE Trace Gas Retrievals*

| Name | Long name | Unit |
|---|---|---|
| $rho_t$ | Top of atmosphere reflectance | $sr^{-1}$ |
| solar_zenith | Solar Zenith Angle at pixel location | degrees |
| sensor_zenith | Sensor Zenith Angle at pixel location | degrees |
| cldmask | PACE OCI cloud mask from operational L2 flags | unitless |

| | | |
|---|---|---|
| watermask | Watermask from PACE OCI L1b | unitless |

### 2.3 Algorithm Output Variables

**Table 3.** *Outputs from PACE Trace Gas Product*

| Name | Long name | Unit |
|---|---|---|
| no2 | Nitrogen Dioxide Vertical Column Density | molecules/cm$^2$ |
| ozone | Total Column Ozone | Dobson Units |

## 3 Algorithm Usage Constraints

# Known Issues and Limitations

- **Equatorial tilt region:**
  Unexpected retrieval features are present near the equator where the OCI instrument tilt changes. Users should avoid these data until improvements are implemented in future updates.
- **Over water scenes:**
  Caution is advised when using the PACE $NO_2$ product over water. Retrievals perform well for strong $NO_2$ signals but exhibit higher uncertainty for weaker signals. Work is underway to develop an improved algorithm tailored for ocean surfaces.
- **High sensor zenith angles (60°):**
  Because OCI and TROPOMI are co-located within a ±2-hour window, there is limited overlap at the edges of OCI granules where sensor zenith angles exceed 60°. Over water, where BRDF effects are stronger, retrieval errors increase at these large viewing angles. This issue will be mitigated in the next product release.

# Flagging

*The Level 2 product includes binary quality flags for $NO_2$ and $O_3$. Users are recommended to avoid retrievals where the quality flag is equal to 1.*

The Level 2 $NO_2$ Quality Flag is set to 1 for any of the following conditions:

- Scene is affected by clouds or snow/ice
- Solar zenith angle exceeds 70 degrees (retrievals accuracy decreases due to long atmospheric path length)
- Observed radiances are identified as low quality in L1B measurements

The Level 2 $O_3$ Quality Flag is set to 1 for any of the following conditions:

- Solar zenith angle exceeds 80 degrees (retrievals accuracy decreases due to long atmospheric path length)
- Observed radiances are identified as low quality in L1B measurements

*The Level 3 product is produced for PACE OCI pixels where the Level 2 quality flag is set to 0. Additionally, the Level 3 $O_3$ produce excludes pixels with a solar zenith angle greater than 70 degrees.*

## 4 Performance Assessment

### 4.1 Improvements from PACE OCI

One of the key advantages from PACE OCI is its spatial resolution, which is 5-10x higher than that of traditional trace gas instruments such as TROPOMI and TEMPO. Figure 4 in Fasnacht et al. (2025) illustrates this improved spatial resolution by comparing PACE OCI retrievals with those from with TROPOMI. These retrievals over eastern Saudia Arabia on April 26 and June 17, 2024, demonstrate PACE OCI's enhanced capability to resolve small-scale emissions and distinguish between multiple $NO_2$ plumes. Additionally, when averaged over the TROPOMI field of view, the PACE OCI $NO_2$ show good quantitative agreement with TROPOMI.

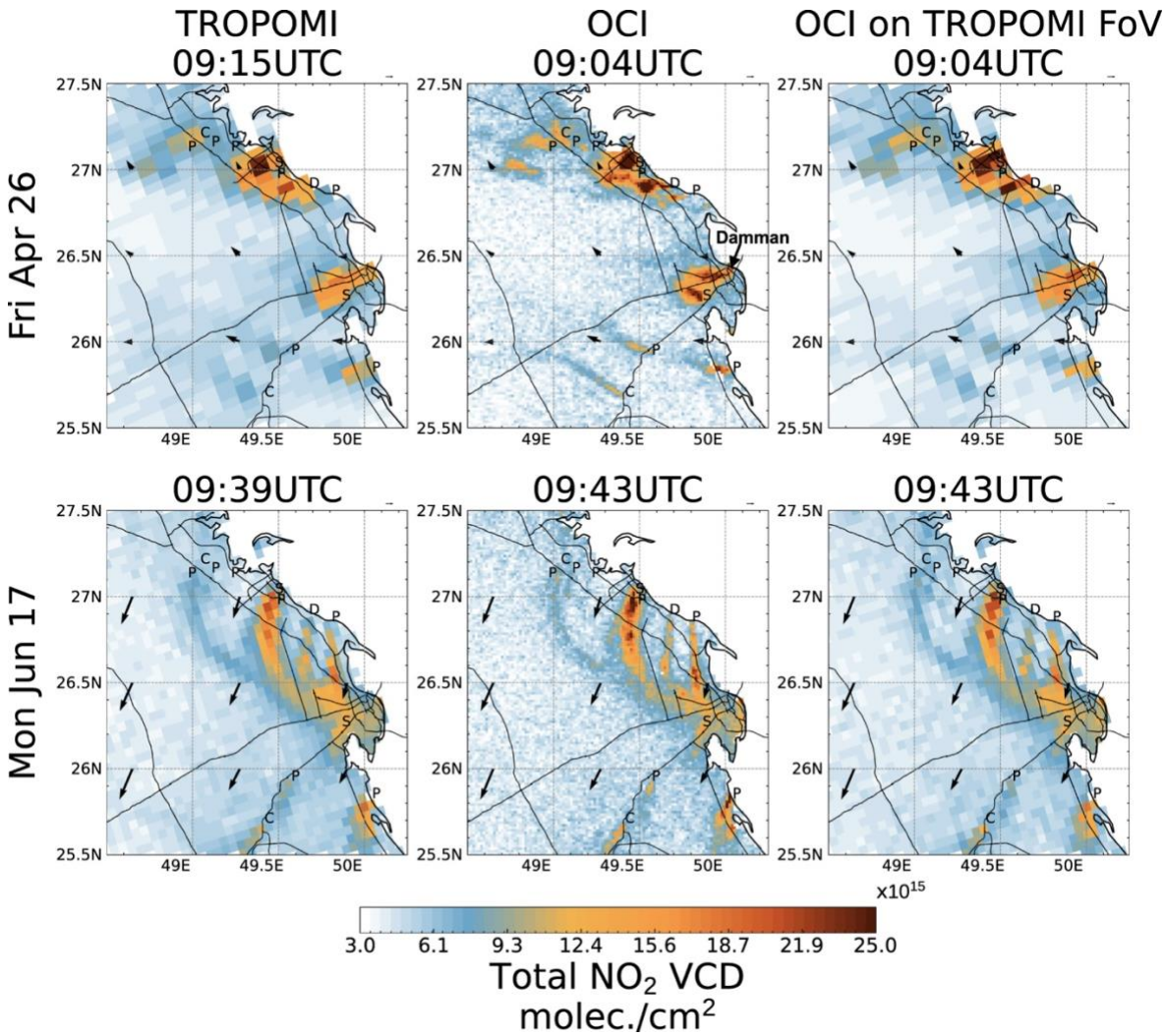


Figure 3. NO2 retrievals from PACE OCI and TROPOMI on April 26 and June 17 of 2024. This figure is from Fasnacht et al. 2025.

### 4.2 Validation Methods

Validation of the $NO_2$ retrievals is ongoing, but early results were reported in Fasnacht et al. (2025). In this paper, comparisons with ground-based Pandora instruments indicated that the quality of the OCI $NO_2$ retrievals is comparable to TROPOMI. This is expected, as a machine learning model can only perform as well as the data for which it was trained. Since that publication, the model has been further refined. Figure 4 shows updated comparisons between OCI and ground-based Pandora measurements, alongside corresponding comparisons from TROPOMI. The Pandora instruments have been shown to be accurate to within 0.1 DU or 2.69e15 molecules/$cm^{-2}$. (Herman et al. 2009). For this analysis, measurements were averaged hourly across all pixels within 5 km of each Pandora site. As previously reported, the PACE OCI

$NO_2$ retrievals show agreement with Pandora similar to that of TROPOMI. There are generally more TROPOMI pixels shown in this comparison because TROPOMI retrievals under partially cloudy skies, while the initial PACE OCI retrievals are limited to cloud-free scenes. Figure 5 shows a similar comparison, but for satellite pixels within 1 km of the Pandora station. Since PACE OCI has a higher spatial resolution, it has 4x as many match ups within 1km of the Pandora stations compared to TROPOMI. There is also some improvement, as expected, when restricting the comparison to PACE OCI pixels within 1km of the Pandora stations. Further validation efforts include more extensive comparisons with ground-based measurements and plans to incorporate airborne sensors to assess higher resolution spatial patterns.

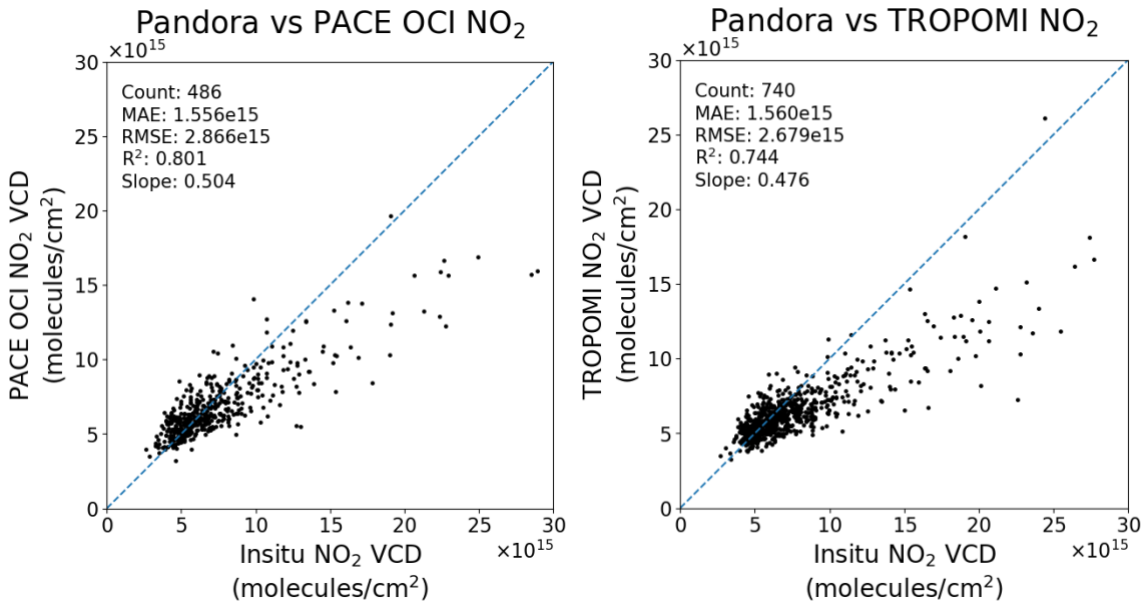


Figure 4. Satellite retrieved $NO_2$ vertical column density comparisons with ground-based Pandora measurements during May 2024. The measurements are averaged hourly for any satellite pixels within 5km of the Pandora site.

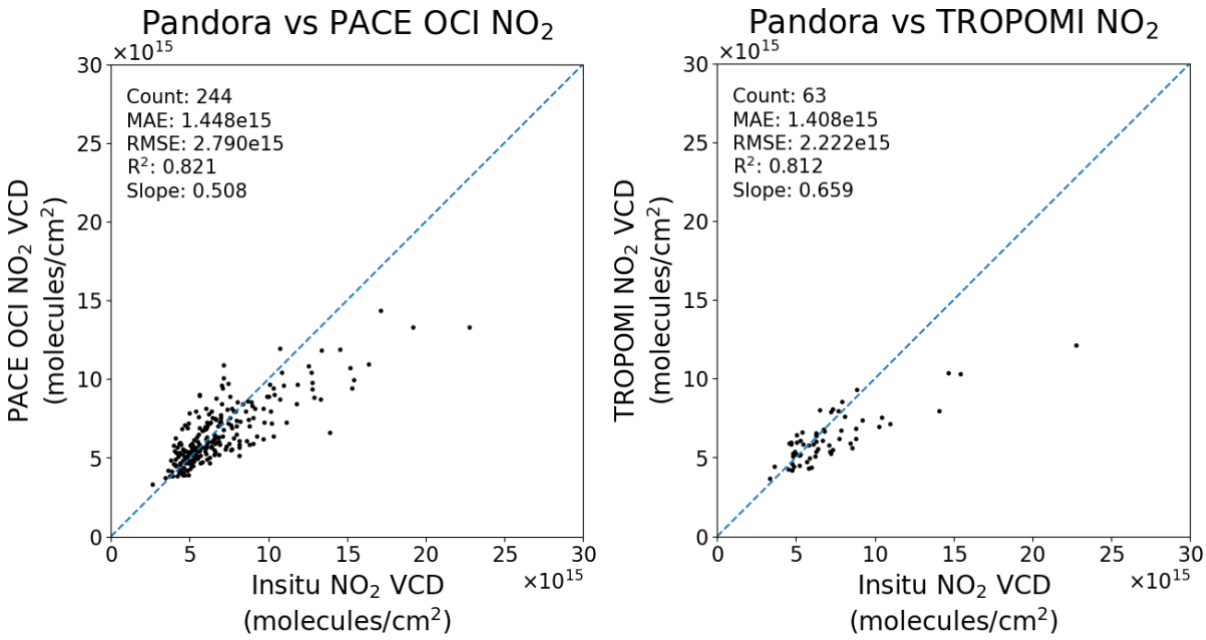


Figure 5. Similar to figure 4, but satellite retrievals are averaged within 1 km of Pandora stations.

Initial validation of the $O_3$ product is also ongoing. Early comparisons with a limited set of ground-based Dobson-Brewer instruments (Bais 1997) and Pandoras (Herman et al. 2009) from May 2024 show a root mean square error of approximately 7-12 DU (Figures 5 and 6). Similar to the $NO_2$ comparisons, these results are comparable to those between TROPOMI and the ground based instruments. The algorithm is being further refined to reduce discrepancies with the ground-based measurements and a future publication will provide more comprehensive evaluation of the $O_3$ product.

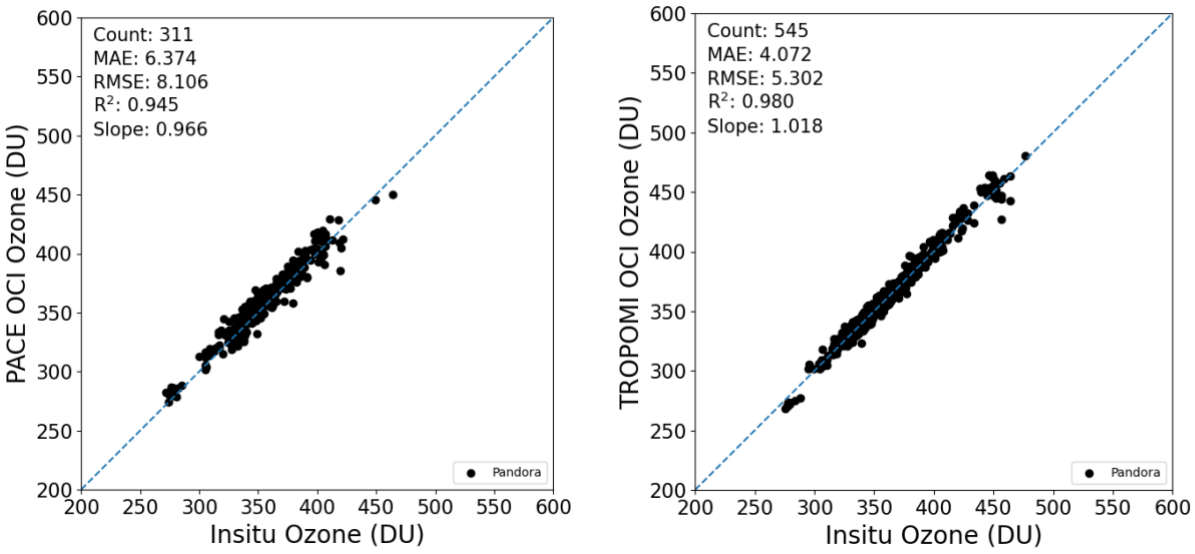


Figure 6. Satellite retrieved $O_3$ comparisons with ground-based Pandora measurements during May 2024. The measurements are averaged hourly for any satellite pixels within 5km of the Pandora site.

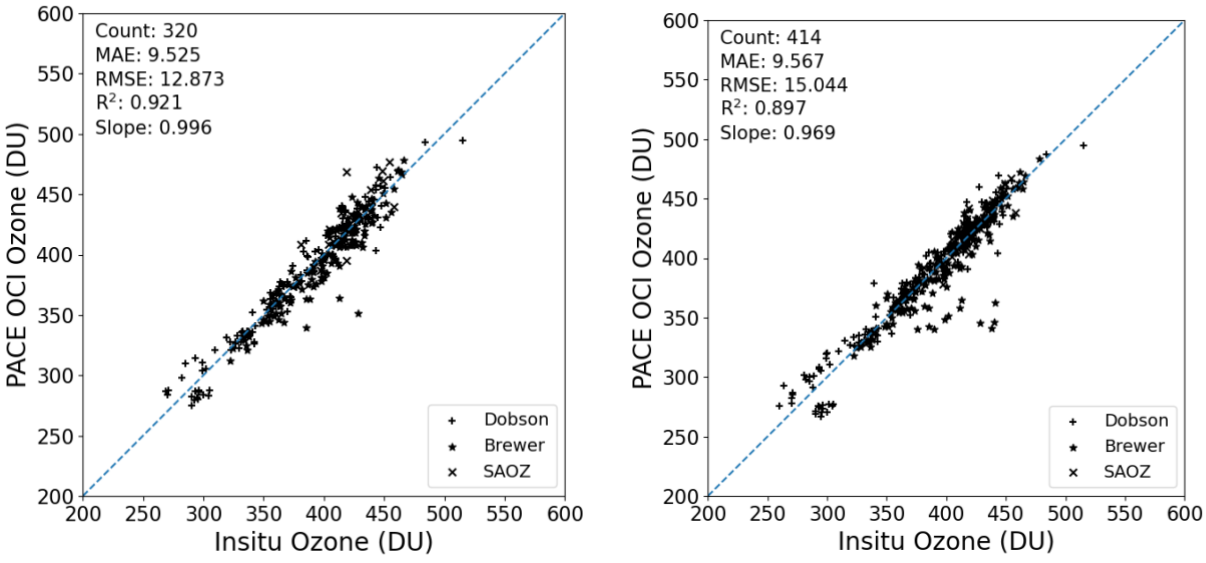


Figure 7. Satellite retrieved $O_3$ comparisons with ground-based Dobson and Brewer measurements during May 2024. The measurements are averaged daily for any satellite pixels within 5 km of the Pandora site.

## 5. Data Access

### 5.1 Input Data Access

Entry #1

URL

https://oceancolor.gsfc.nasa.gov/data/download_methods/

DESCRIPTION

PACE OCI L1b files can be downloaded through the various methods described at the link above.

Entry #2

URL

https://search.earthdata.nasa.gov/search?q=S5P_L2__NO2____1

DESCRIPTION

The $NO_2$ data utilized to train the model can be downloaded from the NASA Earthdata link provided above.

### 5.2 Output Data Access

URL

https://search.earthdata.nasa.gov/search?q=*trgas*&ac=true

DESCRIPTION

The output products can be downloaded from the NASA Earthdata link provided above.

## Contact Details

Zachary Fasnacht

Roles : Methodology, Writing, Corresponding Author & Investigation

Affiliations: NASA Goddard Space Flight Center/SSAI

Email: zachary.t.fasnacht@nasa.gov

## References


Bais, A. F. (1997). Absolute spectral measurements of direct solar ultraviolet irradiance with a Brewer spectrophotometer. *Applied Optics, 36*(21), 5199–5204. https://doi.org/10.1364/AO.36.005199

Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstätter-Weißenmayer, A., Richter, A., De Beek, R., Hoogen, R., Bramstedt, K., Eichmann, K.-U., Eisinger, M., & Perner, D. (1999). The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results. *Journal of the Atmospheric Sciences, 56*(2), 151–175. https://doi.org/10.1175/1520-0469(1999)056

Fasnacht, Z., Joiner, J., Bucsela, E., Bandel, M., Liu, F., Lamsal, L., & Krotkov, N. (2025). Utilizing machine learning for high-resolution $NO_2$ total columns from PACE OCI. *Environmental Research Letters, 20*(7), 074052. https://doi.org/10.1088/1748-9326/addfef

Herman, J., Cede, A., Spinei, E., Mount, G., Tzortziou, M., & Abuhassan, N. (2009). $NO_2$ column amounts from ground-based Pandora and MFDOAS spectrometers using the direct-sun DOAS technique: Intercomparisons and application to OMI validation. *Journal of Geophysical Research: Atmospheres, 114*(D13). https://doi.org/10.1029/2009JD011848

Joiner, J., Marchenko, S., Fasnacht, Z., Lamsal, L., Li, C., Vasilkov, A., & Krotkov, N. (2023). Use of machine learning and principal component analysis to retrieve nitrogen dioxide ($NO_2$) with hyperspectral imagers and reduce noise in spectral fitting. *Atmospheric Measurement Techniques, 16*, 481–500. https://doi.org/10.5194/amt-16-481-2023

Krueger, A. J. (1989). TOMS: A satellite instrument for monitoring the Earth's ozone. *Journal of Atmospheric and Terrestrial Physics, 51*(5), 423–438. https://doi.org/10.1016/0021-9169(89)90044-3

Mateer, C. L., & Hudson, R. D. (1973). Use of the backscattered ultraviolet technique to measure total ozone. *Journal of the Atmospheric Sciences, 30*(4), 852–857. https://doi.org/10.1175/1520-0469(1973)030<0852:UOTBUT>2.0.CO;2

McPeters, R. D., Bhartia, P. K., Krueger, A. J., Herman, J. R., Wellemeyer, C. G., Seftor, C. J., Jaross, G., Taylor, S. L., Swissler, T., Torres, O., Labow, G., Byerly, W., & Cebula, R. P. (1996). *Earth Probe Total Ozone Mapping Spectrometer (TOMS) Data Products User's Guide* (NASA Technical Publication 1996-206895). NASA Goddard Space Flight Center.

Richter, A., & Burrows, J. P. (2002). Tropospheric $NO_2$ from GOME measurements. *Advances in Space Research, 29*(11), 1673–1683. https://doi.org/10.1016/S0273-1177(02)00100-X

van Geffen, J., Boersma, K. F., Eskes, H., Sneep, M., ter Linden, M., Zara, M., & Veefkind, J. P. (2020). S5P TROPOMI $NO_2$ slant column retrieval: Method, stability, uncertainties and

comparisons with OMI. *Atmospheric Measurement Techniques, 13*, 1315–1335. https://doi.org/10.5194/amt-13-1315-2020

van Geffen, J., Eskes, H., Compernolle, S., Pinardi, G., Verhoelst, T., Lambert, J.-C., Sneep, M., ter Linden, M., Ludewig, A., Boersma, K. F., & Veefkind, J. P. (2022). Sentinel-5P TROPOMI $NO_2$ retrieval: Impact of version v2.2 improvements and comparisons with OMI and ground-based data. *Atmospheric Measurement Techniques, 15*, 2037–2060. https://doi.org/10.5194/amt-15-2037-2022