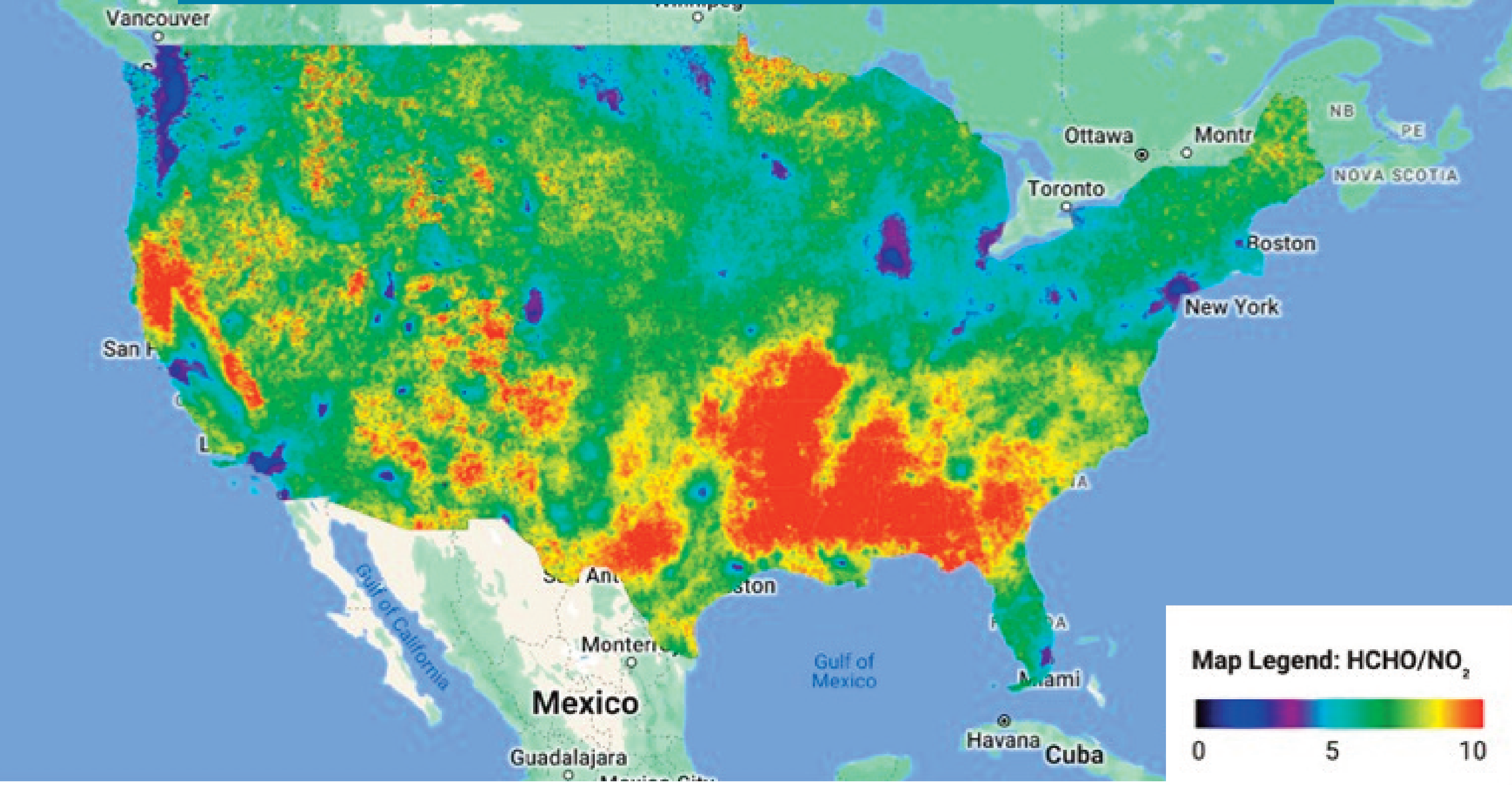

Satellite Data to Inform Ozone Sensitivity: A Practical Methodology Using Google Earth Engine

by Jennifer McGinnis, Tracey Holloway, Jenny Bratburd, Madankui Tao, and Arlene Fiore

Evaluating ozone sensitivity to emissions of its precursor gases with satellite data has evolved into a cutting-edge and increasingly popular application of remote sensing for health and air quality. Google Earth Engine offers a practical, user-friendly platform to support this analysis anywhere in the world using data from the TROPOMI instrument.

High levels of near-surface ozone damage human health and vegetation, leading to the regulation of ozone through the National Ambient Air Quality Standards (NAAQS). Most commonly, ozone is formed by volatile organic compounds (VOCs) and nitrogen oxides (NOx = NO_2 + NO) reacting in the presence of sunlight, especially on hot days. While ozone formation has historically been more sensitive to VOCs in urban areas, and more sensitive to NOx in non-urban areas, these patterns are changing as reductions in NOx emissions have moved many cities into NOx-sensitive ozone production regimes.

Air quality management for urban ozone requires characterizing the sensitivity of ozone to emitted VOCs versus NOx to determine efficient mitigation strategies. However, this analysis is severely limited by the lack of ground-based measurement data. Satellite data offer continuous spatial coverage of NO_2 and formaldehyde (HCHO, a VOC that is an indicator of ozone production). This is particularly valuable because it allows air quality managers to evaluate ozone sensitivity in areas without specific monitoring capabilities, such as areas away from monitors or where monitors do not measure the specific precursors of interest, namely HCHO.[1] This comprehensive coverage bridges the data gap and enables a more thorough evaluation of air quality and ozone-related dynamics across various locations.

The ratio of HCHO to NO_2, often referred to as formaldehyde nitrogen ratio (FNR), has been used to identify whether VOCs or NOx are the limiting factor in ozone production. Research supporting the application of satellite data to ozone sensitivity was first proposed by Dr. Randall Martin[2] and advanced by Dr. Bryan Duncan,[3] with more recent studies evolving the understanding of satellite-derived FNR for air quality applications.[4,5]

Despite the potential for FNR to inform ozone sensitivity, the extent of the relationship between the satellite retrieval of HCHO and ground level VOCs is a subject of ongoing research.[4,6] Other limitations include satellite retrieval error, complex chemistry, and variations in ozone sensitivity depending on the location of interest. Still, satellite-derived FNR can serve as an indicator given the scarcity of monitors for NO_2 and VOCs needed to calculate ozone indicator ratios with surface data.

Although satellite-derived FNR, NO_2, and HCHO have the potential to support air quality management, the use of these data products has been limited by technical hurdles of data access and manipulation. Typically, calculating FNR requires downloading the satellite data and calculating averages and ratios in a programming environment like Python. This level of data processing can be a barrier for resource-limited air quality organizations.

Applying Google Earth Engine

We have found that Google Earth Engine offers the most user-friendly approach for analyzing the best-available global satellite data for HCHO and NO_2 from an instrument called the TROPOspheric Monitoring Instrument (TROPOMI). While in orbit around Earth, TROPOMI records daily observations at approximately 1:30 pm local time. The data are available through Google Earth Engine from July 2018 to present day at a spatial resolution of

Table 1. Overview of ease of use (green/E: easy, yellow/M: medium, red/D: difficult; based on author judgement for users without programming experience), analysis capabilities and datasets available relevant to ozone sensitivity analysis including temporal averaging, division across variables, and inclusion of TROPOMI data (green/Y: yes, red/N: no) of five satellite data visualization tools: NASA Worldview, NASA Giovanni, NASA Panoply, Google Earth Engine, and flexible data analysis software including Python, IDL, R, Matlab, and other related environments.

	Ease of Use	*Allows time averaging (2+ days)*	*Allows division among variables*	*Includes TROPOMI*
NASA Worldview	**E**	**N**	**N**	**N**
NASA Giovanni	**E**	**Y**	**N**	**N**
Panoply	**E**	**N**	**Y**	**Y**
Google Earth Engine	**M**	**Y**	**Y**	**Y**
Python, IDL, R, Matlab, etc.	**D**	**Y**	**Y**	**Y**

3.5 x 7.5 km^2. Google Earth Engine allows for a range of analyses that can be useful in analyzing ozone sensitivities, including plotting averages of consecutive or nonconsecutive days, and calculating the product or ratio of variables. We apply these capabilities to highlight a key finding from remote sensing-based analysis: that higher-ozone days tend to be more NOx-limited, even in cities that are normally VOC-limited.[5]

Google Earth Engine is the only web tool that includes TROPOMI and allows for temporal averaging and ratioing necessary for FNR analysis. Table 1 compares the capabilities of widely used platforms and tools used by the air quality community for satellite data analysis. These include NASA Worldview (https://worldview.earthdata.nasa.gov/), which provides near-real-time data and archived historical data for specific days but does not include TROPOMI or allow averaging across time or calculation of ratios for FNR; NASA Giovanni (https://giovanni.gsfc.nasa.gov/giovanni/), which allows for more complex analysis including averaging over time, but does not include TROPOMI nor does it allow calculation of ratios for FNR; and NASA Panoply (https://www.giss.nasa.gov/tools/panoply), which supports visualization, and includes TROPOMI, but has no robust temporal averaging capability. Python and other advanced data analysis tools (IDL, R, Matlab, etc.) are the most powerful and flexible tools for data analysis, but require technical expertise, specialized training, and/or dedicated staff time. Additionally, some analysis techniques follow a more selective filtering process that is best suited for Python or other data analysis tools (e.g., removing pixels over water).[5]

Google Earth Engine includes a web tool, the Code Editor, that allows users to manipulate datasets without downloading the data onto their local machine (see Figure 1), with free access for government research among other noncommercial applications purposes (see https://earthengine.google.com/noncommercial/).

Satellite Data for Ozone Indicators

The first step in calculating the FNR is to filter the satellite data based on data quality flags. Google Earth Engine automatically does this by removing the pixels with a lower quality than what is recommended in the product manuals (a quality assurance value less than 0.5 for HCHO and 0.75 for NO_2).[7,8] The next step is to select 10+ days for averaging, with the goal of smoothing out the "noise" in the HCHO data product (34+ is preferred, with averaging times for TROPOMI drawn from Vigouroux et al., 2020).[9] We consider here the New York City and Denver Metropolitan Statistical Areas (MSAs) to compare and expand upon previous work.[1,5] Figure 2 shows NO_2 and HCHO averaged over the summer ozone season for each of these two regions. Consistent with prior analyses of TROPOMI NO_2, the NO_2 vertical column densities (VCDs) over both New York City and Denver are centralized to the city center and decrease further away from the city (Figure 2a).[10] The abundance of HCHO (Figure 2b) does not show the same centralized pattern, due to VOC contributions from biogenic sources.[11]

Capturing FNR

To analyze lower-ozone and higher-ozone days, we used day-by-day characterizations of ambient ozone for 2022

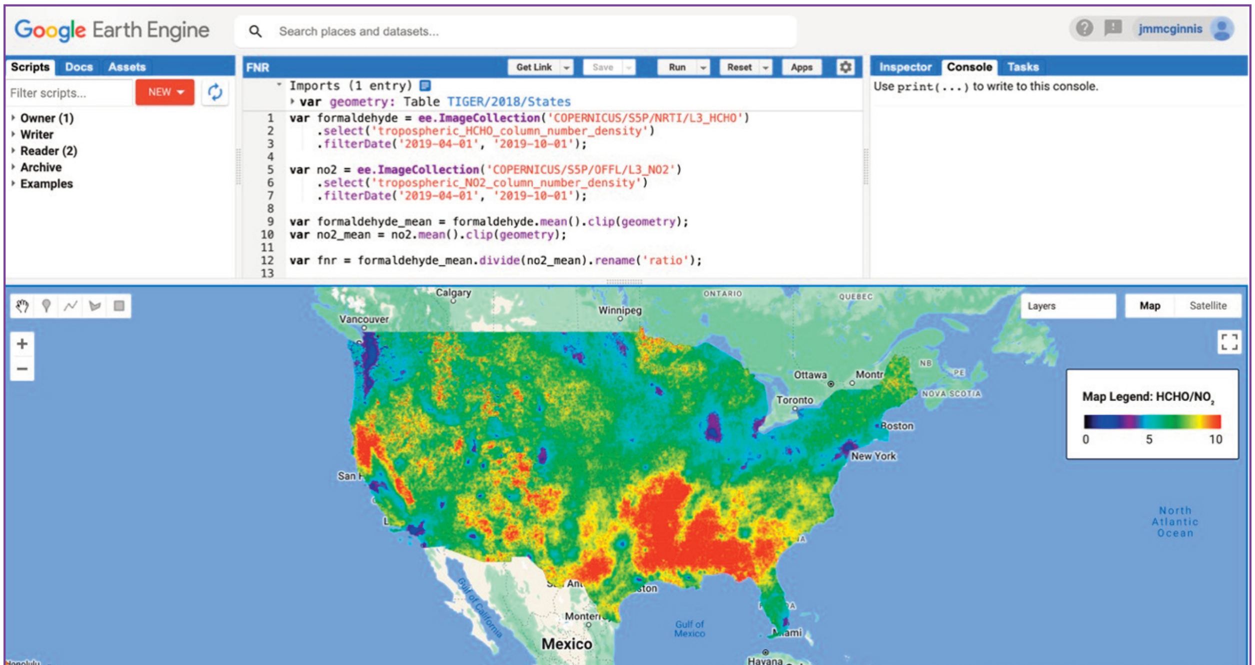

Figure 1. Screenshot of the Code Editor available in Google Earth Engine with example code and imagery.
Source: https://code.earthengine.google.com/

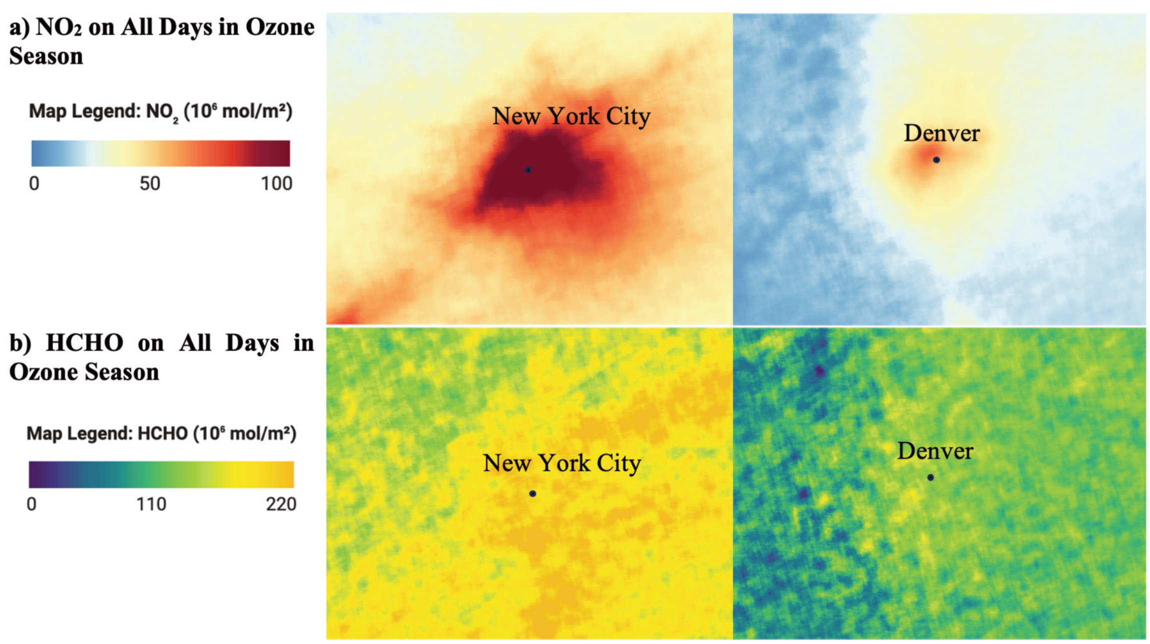

Figure 2. The ozone season is 153 days long, but data were not available at every pixel for every day because quality levels vary across the domain and retrieval of data: (a) NO_2 average over New York City (left) (mean of 112 days available per pixel) and Denver (right) (mean of 105 days available per pixel) of all days during the ozone season (May 1–September 30); (b) HCHO average over New York City (left) (mean of 127 days available per pixel) and Denver (right) of all days during the ozone season (May 1–September 30) (mean of 109 days available per pixel). The exact number of days plotted per pixel can be extracted from Google Earth Engine if desired. The units of mol/m^2 are used rather than the commonly used molecules/cm^2 because The European Space Agency reports TROPOMI data products to comply with the International System of Units. This can be converted to molecules/cm^2 by multiplying the product by 6.02214×10^{19}.

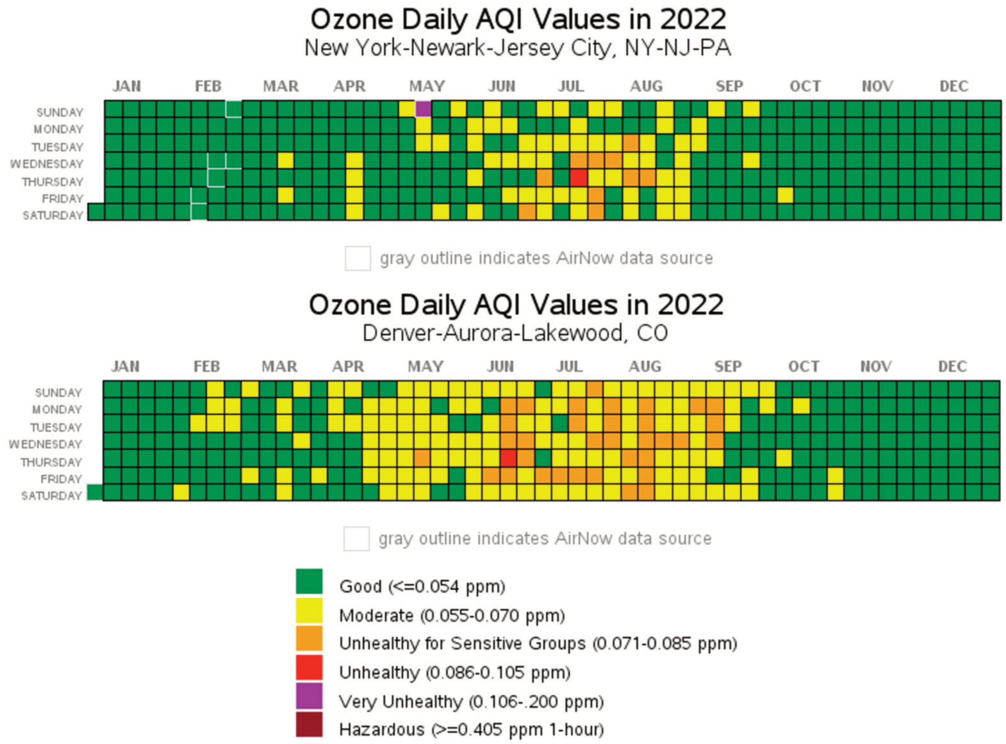

Figure 3. Daily ozone AQI values for 2022 from EPA Outdoor Air Quality Data Viz Tools, Single Year Tile Plot, for the New York City MSA (New York, NY; Newark, NJ; and Jersey City, PA; top) and the Denver MSA (Denver, Aurora, and Lakewood, CO; bottom). For this analysis, green/Good and yellow/Moderate are considered lower-ozone days; all day orange/Unhealthy for Sensitive Groups or higher (0.071 ppm or higher) are considered higher-ozone days.
Source: https://www.epa.gov/outdoor-air-quality-data/air-data-tile-plot

from the U.S. Environmental Protection Agency (EPA; see https://www.epa.gov/outdoor-air-quality-data/air-data-tile-plot), as shown in Figure 3. The New York City MSA had 12 higher-ozone days in 2022; the Denver MSA, 35 higher-ozone days (not all days have valid TROPOMI data, so we also applied a mask to regions with less than 10 days included in the averaged pixels). For each urban area, temporally averaged HCHO data were divided by the temporally averaged NO_2 data to calculate the FNR for the entire ozone season, lower-ozone days, and higher-ozone days in Google Earth Engine (see Figure 4).

Based on Jin et al. (2020), we define FNR values indicative of VOC-limited, NOx-limited, or transitional (affected by both) regimes. Table 2 reports the FNR thresholds from Jin et al. (2020) for seven cities, including New York City, so city-specific values are used where available.[4] The multi-city average value is used for Denver.[12]

Ozone Sensitivity in Two Cities

Averaged over the full ozone season, the FNR for the downtown area of both cities appear VOC-limited (Figure 4a). This is expected as city centers tend to have a dense population with many sources of anthropogenic NOx, making VOCs the limiting component in ozone production. Further away from the city center, the region becomes more NOx-limited. This agrees with previous analysis showing more NOx-limited regimes in the surrounding suburban areas, most likely due to fewer NOx sources further from the city center.[3,5,10] This pattern is consistent during lower-ozone days in both cities where the city center is VOC-limited, and the surrounding suburban areas are more sensitive to NOx (Figure 4b).

For higher-ozone days in New York City, satellite-retrieved FNR values indicate a shift toward NOx-limited regimes (Figure 4c), this trend is consistent with Tao et al.'s (2022) analysis for summer 2018.[5] We find that in both cities more of the region is NOx-limited on higher-ozone days than lower-ozone days, implying a transition to NOx-limited regimes on the most polluted days. These results suggest that reductions in NOx emissions in New York City and Denver may decrease ozone production on days above the ozone NAAQS.

Conclusion

Satellite data can inform ozone management strategies through the use of the FNR, especially analyzed on higher-ozone days. Additionally, plotting HCHO and NO_2 satellite data provides a visual representation of the spatial

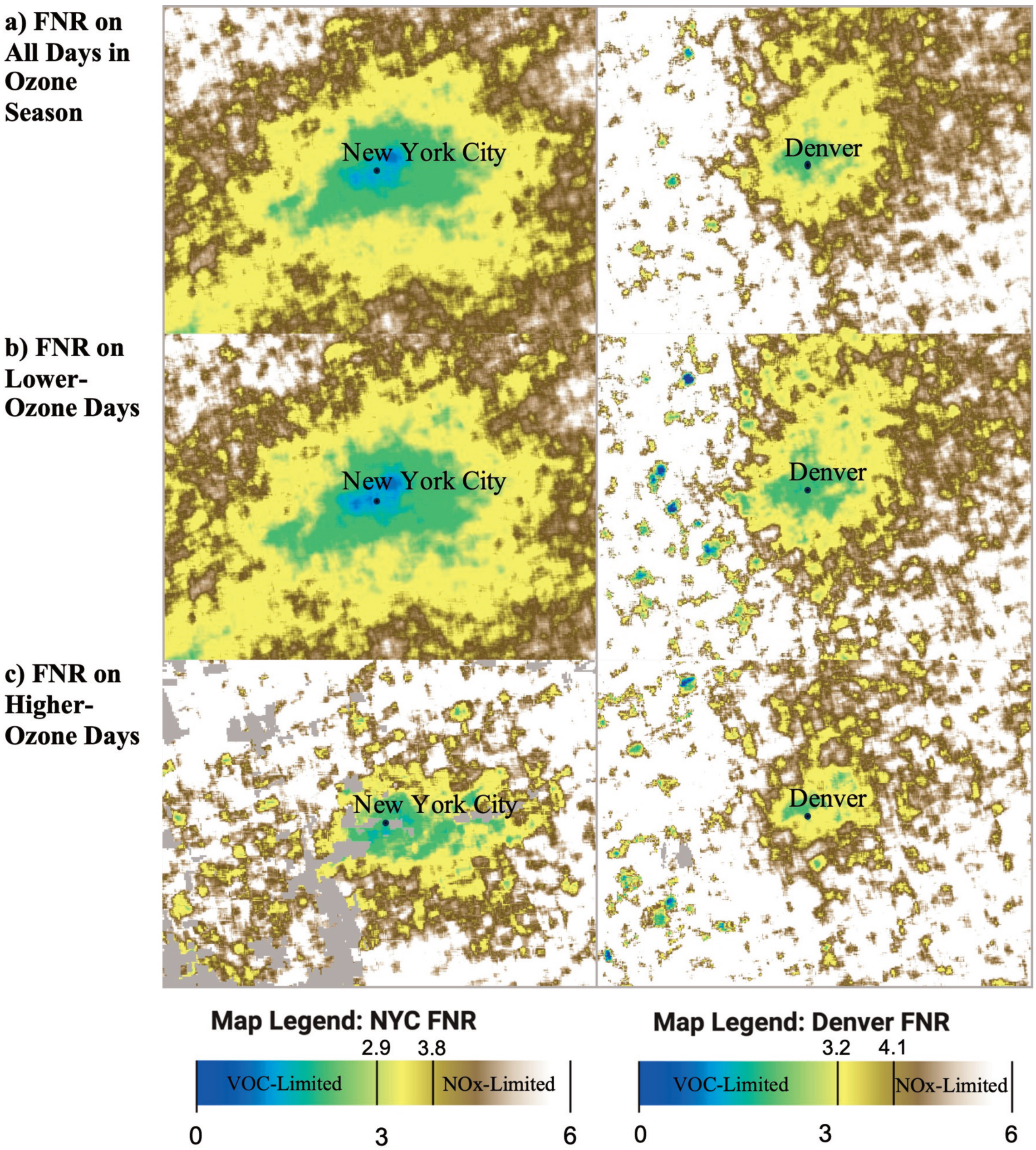

Figure 4. Because the quality assurance of NO_2 is more selective than that of HCHO, in some areas there are a different number of available days of NO_2 in comparison to HCHO. However, a mask was applied to ensure at least 10 pixels of NO_2 data and HCHO data were temporally averaged before calculating the ratio: (a) FNR average of all days during the ozone season. Over New York City (left), 153 days were plotted (mean of 112 days of NO_2 and 127 days of HCHO available per pixel). Over Denver (right), 153 days were plotted (mean of 105 days of NO_2 and 109 days of HCHO available per pixel). (b) FNR on days categorized as lower-ozone days during the ozone season. Over New York City, 141 days were plotted (mean of 100 days of NO_2 and 114 days of HCHO available per pixel). Over Denver, 118 days were plotted (mean of 72 days of NO_2 and 76 days of HCHO available per pixel). (c) FNR on days categorized as higher-ozone days during the ozone season. Over New York City, 12 days were plotted (mean of 12 days of NO_2 and 12 days of HCHO available per pixel). Over Denver, 35 days were plotted (mean of 32 days of NO_2 and 32 days of HCHO available per pixel).

Table 2. Thresholds for FNR analyses from the seven cities and recommended overall value discussed from Jin et al. (2020). Uncertainty is expressed as two standard deviations (2σ). These thresholds were derived using an empirical observation-based approach to identify the FNR value associated with the highest probability that surface ozone exceeded 70 ppb.

	FNR Thresholds
Los Angeles, CA	[4.1-5.0] 2σ=1.8
New York City, NY	[2.9-3.8] 2σ=0.4
Chicago, IL	[3.2-4.1] 2σ=0.6
Washington D.C.	[3.2-4.1] 2σ=1.2
Pittsburg, PA	[4.3-4.3] 2σ=0.6
Atlanta, GA	[3.2-4.1] 2σ=2.8
Houston, TX	[2.6-3.5] 2σ=2.4
Overall	[3.2-4.1] 2σ=0.2

distribution of these pollutants, allowing air quality managers to identify localized areas of concern and prioritize mitigation efforts accordingly. We present here an approach using web tools from EPA and Google Earth Engine to support satellite data analysis over any region of the United States and any year since 2018 (the first year for which TROPOMI is available). Our goal is to support air quality managers utilization of satellite data to support ozone assessment for public health, State Implementation Plans, and other applications. This integrated approach facilitates a more holistic understanding of ozone dynamics and supports the implementation of targeted measures to improve air quality. **em**

Jennifer McGinnis completed this work in the Nelson Institute Center for Sustainability and the Global Environment (SAGE) at the University of Wisconsin-Madison (UW-Madison). She began her graduate studies at Colorado State University in Fall 2023. **Dr. Tracey Holloway** is the Jeff Rudd and Jeanne Bissell Professor of Energy Analysis and Policy at UW-Madison based in SAGE, and she leads the NASA Health and Air Quality Applied Sciences Team (HAQAST). **Dr. Jenny Bratburd** is the Outreach Manager for HAQAST, based in SAGE at UW-Madison. **Madankui Tao** is a graduate student at the Department of Earth and Environmental Sciences at Columbia University. **Dr. Arlene Fiore** is a member of HAQAST and is a professor in the Department of Earth, Atmospheric and Planetary Sciences at the Massachusetts Institute of Technology. Email: **taholloway@wisc.edu**.

Acknowledgment: Work for this report was funded by the NASA Health and Air Quality Applied Sciences Team (HAQAST). We would like to thank all the groups that made the data and tools used in this report publicly available for air quality and health analysis, including the EPA, NASA, The European Space Agency, and Google Earth Engine. For tutorials on how to conduct this analysis in Google Earth Engine, go to **https://hollowaygroup.org/tutorials**.

References

1. Witman, S.; Holloway, T.; Reddy, P. Integrating Satellite Data into Air Quality Management: Experience from Colorado; *EM Magazine*, February 2014, 34-38.
2. Martin, R.V.; Fiore, A.M.; Van Donkelaar, A. Space-Based Diagnosis of Surface Ozone Sensitivity to Anthropogenic Emissions; *Geophys. Res. Letts*. 2004, *31* (6); https://doi.org/10.1029/2004GL019416.
3. Duncan, B.N.; Yoshida, Y.; Olson, J.R.; Sillman, S.; Martin, R.V.; Lamsal, L.; Hu, Y.; Pickering, K.E.; Retscher, C.; Allen, D.J.; Crawford, J.H. Application of OMI Observations to a Space-Based Indicator of NOx and VOC Controls on Surface Ozone Formation; *Atmos. Environ*.2010, *44* (18), 2213-2223; https://doi.org/10.1016/j.atmosenv.2010.03.010.
4. Jin, X.; Fiore, A.; Boersma, K.F.; Smedt, I.D.; Valin, L. Inferring Changes in Summertime Surface Ozone–NOx–VOC Chemistry over U.S. Urban Areas from Two Decades of Satellite and Ground-Based Observations; *Environ. Sci. Technol.* 2020, *54* (11), 6518-6529. https://doi.org/10.1021/acs.est.9b07785.
5. Tao, M.; Fiore, A. M.; Jin, X.; Schiferl, L.D.; Commane, R.; Judd, L.M.; Janz, S.; Sullivan, J.T.; Miller, P.J.; Karambelas, A.; Davis, S.; Tzortziou, M.; Valin, L.; Whitehill, A.; Civerolo, K.; Tian, Y. Investigating Changes in Ozone Formation Chemistry during Summertime Pollution Events over the Northeastern United States; *Environ. Sci. Technol.* 2022, *56* (22), 15312-15327; https://doi.org/10.1021/acs.est.2c02972.
6. Souri, A.H.; Johnson, M.S.; Wolfe, G.M.; Crawford, J.H.; Fried, A.; Wisthaler, A.; Brune, W.H.; Blake, D.R.; Weinheimer, A.J.; Verhoelst, T.; Compernolle, S.; Pinardi, G.; Vigouroux, C.; Langerock, B.; Choi, S.; Lamsal, L.; Zhu, L.; Sun, S.; Cohen, R.C.; Min, K.-E.; Cho, C.; Philip, S.; Liu, X.; Chance, K. Characterization of Errors in Satellite-Based HCHO / NO_2 Tropospheric Column Ratios with Respect to Chemistry, Column-to-PBL Translation, Spatial Representation, and Retrieval Uncertainties; *Atmos. Chem. Phys.* 2023, *23* (3), 1963-1986; https://doi.org/10.5194/acp-23-1963-2023.
7. Romahn, F.; Pedergnana, M.; Loyola, D.; Apituley, A.; Sneep, J.; Veefkind, J.P. Sentinel-5 Precursor/TROPOMI Level 2 Product User Manual Formaldehyde HCHO; *Remote Sens. Environ*. 2022, *120*, 70-83.
8. Eskes, H.; Geffen, J. van; Boersma, F.; Eichmann, K.-U.; Apituley, A.; Pedergnana, M.; Sneep, M.; Veefkind, J. P.; Loyola. Sentinel-5 Precursor/TROPOMI Level 2 Product User Manual Nitrogen Dioxide, 2019, No. 3.
9. Vigouroux, C.; Langerock, B.; Bauer Aquino, C.A.; Blumenstock, T.; Cheng, Z.; De Mazière, M.; De Smedt, I.; Grutter, M.; Hannigan, J.W.; Jones, N.; Kivi, R.; Loyola, D.; Lutsch, E.; Mahieu, E.; Makarova, M.; Metzger, J.-M.; Morino, I.; Murata, I.; Nagahama, T.; Notholt, J.; Ortega, I.; Palm, M.; Pinardi, G.; Röhling, A.; Smale, D.; Stremme, W.; Strong, K.; Sussmann, R.; Té, Y.; van Roozendael, M.; Wang, P.; Winkler, H. TROPOMI–Sentinel-5 Precursor Formaldehyde Validation Using an Extensive Network of Ground-Based Fourier-Transform Infrared Stations; *Atmos. Meas. Tech*. 2020, *13* (7), 3751-3767; https://doi.org/10.5194/amt-13-3751-2020.
10. Goldberg, D.L.; Harkey, M.; de Foy, B.; Judd, L.; Johnson, J.; Yarwood, G.; Holloway, T. Evaluating NOx Emissions and Their Effect on O_3 Production in Texas Using TROPOMI NO_2 and HCHO; *Atmos. Chem. Phys*. 2022, *22* (16), 10875-10900; https://doi.org/10.5194/acp-22-10875-2022.
11. Zhu, L.; Jacob, D.J.; Keutsch, F.N.; Mickley, L.J.; Scheffe, R.; Strum, M.; González Abad, G.; Chance, K.; Yang, K.; Rappenglück, B.; Millet, D.B.; Baasandorj, M.; Jaeglé, L.; Shah, V. Formaldehyde (HCHO) As a Hazardous Air Pollutant: Mapping Surface Air Concentrations from Satellite and Inferring Cancer Risks in the United States; *Environ. Sci. Technol.* 2017, *51* (10), 5650-5657; https://doi.org/10.1021/acs.est.7b01356.
12. Jin, X.; Fiore, A.M.; Geigert, M. Using Satellite Observed Formaldehyde (HCHO) and Nitrogen Dioxide (NO_2) as an Indicator of Ozone Sensitivity in a SIP; Technical Guidance Document, June 22, 2018; https://doi.org/10.7916/D8M34C7V.