JGR Atmospheres

RESEARCH ARTICLE
10.1029/2023JD038887

CLASP: CLustering of Atmospheric Satellite Products and Its Applications in Feature Detection of Atmospheric Trace Gases

Tabitha Lee[1] **and Yuxuan Wang**[1]

[1]Department of Earth and Atmospheric Sciences, University of Houston, Houston, TX, USA

Key Points:

- Density-based clustering provides a foundation when built upon can describe the spatiotemporal variability of atmospheric trace gases
- A clustering algorithm entitled CLustering of Atmospheric Satellite Products (CLASP) is presented to describe observations of trace gases
- CLASP is applied to TROPOspheric Monitoring Instrument NO_2 observations for event identification, plume variability, and source identification

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Y. Wang and T. Lee,
ywang246@Central.uh.edu;
tclee3@cougarnet.uh.edu

Citation:

Lee, T., & Wang, Y. (2023). CLASP: CLustering of Atmospheric Satellite Products and its applications in feature detection of atmospheric trace gases. *Journal of Geophysical Research: Atmospheres*, *128*, e2023JD038887. https://doi.org/10.1029/2023JD038887

Received 14 MAR 2023
Accepted 31 AUG 2023

Author Contributions:

Conceptualization: Tabitha Lee, Yuxuan Wang
Formal analysis: Tabitha Lee
Funding acquisition: Tabitha Lee, Yuxuan Wang
Investigation: Tabitha Lee
Methodology: Tabitha Lee
Project Administration: Yuxuan Wang
Software: Tabitha Lee
Supervision: Yuxuan Wang
Visualization: Tabitha Lee

Abstract Satellite instruments have the most potential of capturing trace gas variability as they continually observe the atmosphere and its composition over wide regions. Yet the increasingly large data size of satellite products poses a challenge for their use as traditional data processing methods (e.g., averaging) may not be effective to extract the spatiotemporal variability without prior knowledge of an emission source's spatial and temporal behavior, such as location, time, and plume shape. Here, an agile clustering algorithm entitled CLustering of Atmospheric Satellite Products (CLASP) is presented to identify the spatiotemporal variability of trace gases captured in satellite observations. We find the knowledge discovery method for large data sets, clustering, is suited for identifying the variability of trace gases in satellite observations, as such CLASP is rooted in density-based clustering methods. CLASP detects features from satellite observations and identifies their spatial, magnitude, and temporal axis leading to a better understanding of the spatiotemporal variability of atmospheric trace gases. To test the applicability of CLASP, the algorithm is applied to TROPOspheric Monitoring Instrument NO_2 observations illustrating some of its different capabilities. Implementing CLASP for event identification, capturing plume variability, and source detection, CLASP identified wildfires, observed disruptions from COVID-19 lockdown restrictions, and detected irregular emissions from oil and gas operations.

Plain Language Summary Satellite instruments capture the location and timing of trace gases as they have daily observations of the atmosphere over a wide spatial coverage. With increasing spatial and temporal resolutions of satellite products, traditional data-reducing methods such as averaging could be insufficient in analyzing increasingly large satellite observations. The user must have prior knowledge of the location and temporal behavior of emission sources in order to choose the suitable averaging intervals and such information is often lacking or unreliable for non-urban and point sources. To alleviate this limitation, a clustering algorithm titled CLustering of Atmospheric Satellite Products (CLASP) is presented in this work. CLASP identifies features in satellite observations by their location, amount, and time. CLASP was applied to the satellite observational data set, TROPOspheric Monitoring Instrument NO_2, to identify wildfires, observe COVID-19 lockdown changes, and spot irregular emissions from oil and gas operations.

1. Introduction

Satellite instruments detecting atmospheric composition observe influential atmospheric trace gases which are important for air quality, climate forcing, and the stratospheric ozone layer (Veefkind et al., 2012). Spaced-based remote sensing products provide routine and consistent observations which improve our understanding of the state of our atmosphere. Such observations help us understand processes controlling air quality, identify long- and short-term variability, track pollution plumes, constrain emissions and concentrations, and improve atmospheric models (National Academies of Sciences, Engineering, and Medicine, 2018; Sun et al., 2018; Veefkind et al., 2012). However, the ability of satellite observations to provide information on our atmosphere is limited as traditional data processing methods (e.g., averaging) may not be effective to extract the spatiotemporal variability without prior knowledge of an emission's source's spatial and temporal behavior, such as location, time, and plume shape. For reactive trace gases with short atmospheric lifetime, such as nitrogen dioxide (NO_2), sparse in situ sampling (e.g., surface networks), and coarse resolution observational products cannot always capture the spatial and temporal heterogeneity of NO_2 as it varies over short distances and time. These limitations impede efforts to further understand regional atmospheric composition, processes, trends, and source variability of such an important climate forcing species.

Writing – original draft: Tabitha Lee
Writing – review & editing: Tabitha Lee, Yuxuan Wang

NO_2 is one of many trace gases observed to improve our understanding of how natural processes and anthropogenic emissions influence air quality and climate (National Academies of Sciences, Engineering, and Medicine, 2018). NO_2 is a criterion pollutant and plays a key role in controlling tropospheric ozone (O_3). NO_2 can chemically react to form the nitrate radical (NO_3), nitric acid (HNO_3), and nitrate aerosol. Because it is a short-lived combustion tracer, NO_2 is widely used in detecting changes in fossil fuel emissions. Knowledge of NO_2 spatiotemporal variability is necessary for air quality and understanding the earth's atmosphere.

Satellite instruments have been used to continuously detect changes in atmospheric NO_2 and other trace gases since the 1990s, notably with polar-orbiting satellite missions such as Global Ozone Monitoring Experiment (GOME; Burrows et al., 1999), Global Ozone Monitoring Experiment 2 (GOME-2; Munro et al., 2016), SCanning Imaging Absorption spectroMeter for Atmospheric CartograpHY (SCIAMACHY; Bovensmann et al., 1999), Ozone Monitoring Instrument (OMI; Levelt et al., 2006), and TROPOspheric Monitoring Instrument (TROPOMI; Veefkind et al., 2012). These polar-orbiting instruments have been building off one another, improving aspects such as detection techniques and spatial resolution. Being one of the newest polar-orbiting instruments, TROPOMI carries an initial spatial resolution of 3.5×7 km^2, and since 6 August 2019, 3.5×5.6 km^2 at nadir, which is a great improvement in spatial resolution compared to that of GOME (40×320 km^2), GOME-2 (40×80 km^2), SCIAMACHY (30×60 km^2), and OMI (13×24 km^2). Improvements to the temporal resolution are made in geostationary satellite missions which observe air quality and composition at an hourly rate, such as the Tropospheric Emissions: Monitoring of Pollution (TEMPO; Zoogman et al., 2017) (2.1×4.7 km^2) mission over North America the Geostationary Environmental Monitoring Spectrometer (GEMS; Kim et al., 2020) (3.5×8 km^2) mission over Asia, and the future missions such as Sentinel-4 (Ingmann et al., 2012) (8×8 km^2) to be deployed in 2024 for observations over Europe.

Future difficulties in the application of satellite observations of trace gases are seen as the corresponding data size increases with each spatial and temporal resolution improvement. With the increased spatial resolution, TROPOMI has a data size of approximately 17.7 times that of OMI for 1 day's global pass. With the launch of TEMPO, the corresponding data size will be multiplied further due to the improvements in spatial and temporal resolution. The ever-growing observational record from satellite instruments limits its use as traditional data processing methods may not be robust enough to extract the spatiotemporal variability of trace gases leaving unanswered questions about our atmosphere.

Satellite observations of trace gases capture plumes or features which carry spatial, magnitude, and temporal properties. Features are seen when relatively high magnitude points are located near one another and create a "hotspot" effect. Existing processing practices to locate features are often performed retroactively where prior knowledge or supervision is required. Identifying features in observations, like known source points, city centers, wildfires (Griffin et al., 2021), oil and gas pipeline leaks (Pandey et al., 2019), or increased use of electrical generating units (van der A et al., 2020) require known information on the location and occurrence of the feature. Finch et al. (2022) reported that a supervised deep learning method, such as a convolutional neural network, can identify TROPOMI NO_2 plumes. However, this method requires pre-processing steps such as (a) creating normalized images of TROPOMI NO_2, and (b) training the model for accurate identification. A training data set was required to be created based only on the author's judgment of what is a plume. Finch et al. (2022) note a more precise approach in plume identification is needed as a high degree of subjectivity is included in their method.

Once identified, the preservation of the location and magnitude of known features such as city centers (Ialongo et al., 2020), large point sources such as power plants (Goldberg et al., 2019), and area sources (de Gouw et al., 2020) are often done with the basic technique of averaging. However, because averaging diminishes and smooths any variable features in the temporal, spatial, and magnitude dimension, it sometimes can hinder the discovery of time-varying emission and concentration patterns. This limitation motivated some fundamental questions on current practices using satellite observations of trace gases: (a) Can features be identified with no prior knowledge of any domain or time frame? (b) Can fine feature variability be preserved? (c) Can temporally frequent or infrequent emission sources be reproducibly identified?

Here, we will address these motivating questions to determine if trace gas variability seen in satellite observations can be resolved using existing or novel data processing techniques. To do so, we will first explore the data mining method of clustering and its application for identifying features in satellite observations of trace gases. Clustering is a knowledge discovery method for large data sets as it groups observations based on their similarity (Birant & Kut, 2007). To our best knowledge, clustering has not been applied for variability detection of atmospheric trace

gases but has great potential for translating this variability. Second, we will describe a new, agile clustering algorithm entitled CLustering of Atmospheric Satellite Products (CLASP), to identify and determine the frequency of features in satellite observations of trace gases. Third, we will show the application of CLASP on TROPOMI NO_2 observations for event identification, plume variability, and source identification. Finally, we will discuss the ability of CLASP to observe the spatiotemporal variability of other trace gases throughout the global data set of satellite observations of trace gases.

2. Evaluating Clustering Algorithms

Clustering algorithms are attractive for the task of identification of trace gas variability. They are an unsupervised learning method that does not need pre-processing, training, or prior knowledge of the data set. The two main branches of clustering techniques, centroid- and density-based clustering were explored for the identification of trace gas spatiotemporal variability. We chose not to use a common centroid-based method, *k*-means, in the development of CLASP as the algorithm requires a known number of features or clusters to search for. Given atmospheric gas variability due to meteorology, emissions, or lifetime, it is difficult to know the exact number of features that will appear in the satellite observation, and thus the number of clusters supplied to *k*-means cannot be known. Further, *k*-means can only be applied to one spatial or temporal dimension. Reducing or squeezing the data to one dimension can diminish the spatiotemporal variability of trace gases and dampen features of interest. A greater potential was found with density-based clustering methods as the number of clusters does not need to be known. Below we discuss the specifics of density-based clustering for its use with satellite observations of trace gases.

2.1. Density-Based Clustering

The most widely used density-based clustering algorithm, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), is designed to discover arbitrary-shaped clusters in any large database and simultaneously distinguish noise points when considering spatial properties (Ahmed & Razak, 2016). The algorithm creates clusters where the density of points is considerably higher than outside the cluster leaving the sparsely distributed points as noise (Ester et al., 1996). Identified shapes of clusters or features can be arbitrary as they are only determined by the distance between two points (Ester et al., 1996). These features of DBSCAN make it an attractive choice for clustering satellite observations for variability detection. However, DBSCAN's implementation of satellite observations detecting atmospheric trace gases is limited.

Few applications of DBSCAN extensions are seen in Franklin et al. (2019), Lu et al. (2021). A hierarchical version of DBSCAN called HDBSCAN was used to categorize flaring signals in VIIRS Nightfire observations by Franklin et al. (2019). Lu et al. (2021) used DBSCAN as an initial step in a Deep Neural Network model (DBSCAN-DNN) where the algorithm divided observations from the Suomi National Polar Partnership satellite (SNPP) and the AOD product of Himawari-8 into different pollution levels. With each extension of DBSCAN, there is still no consideration of the temporal element in satellite observations, only spatial. Spatial (S) and temporal (T), ST-DBSCAN from Birant and Kut (2007), builds off DBSCAN but applies three improvements for the clustering of spatial-temporal data (Birant & Kut, 2007). First, ST-DBSCAN can cluster data based on its spatial, temporal, and non-spatial-temporal attributes, such as magnitude; second, ST-DBSCAN can detect noise points when clusters of different densities exist; and third, it can ensure that cluster values are similar across the width of the cluster (Birant & Kut, 2007). ST-DBSCAN has been applied to a limited number of data sets carrying spatial and temporal properties and has found success (Birant & Kut, 2007). Yet, in each study, the density-based clustering algorithm was applied for the extraction of known attributes in the data set which requires prior knowledge of the domain and time frame.

DBSCAN and ST-DBSCAN were applied to TROPOMI NO_2 products for variability detection and found they did not lead to clear results (Figure S1 in Supporting Information S1). The reasons DBSCAN and ST-DBSCAN failed are (a) the algorithms cannot automatically extract features of interest, (b) there is difficulty finding correct input parameters, and (c) there is no or a poor consideration of temporal properties. To allow for trace gas variability a clustering algorithm must incorporate preprocessing methods that identify features, providing easy implementation with no prior knowledge of the data set, and consider the temporal properties of the observations.

It was discovered that density-based clustering provides a foundation that when built upon can describe the spatiotemporal variability of atmospheric trace gases. As such, high potential in DBSCAN's implementation is seen

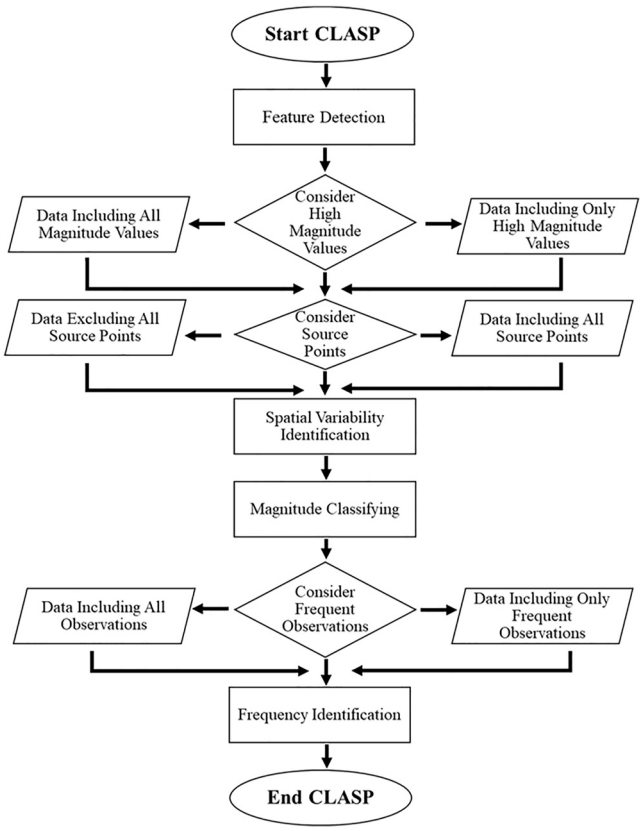

Figure 1. Flowchart of CLustering of Atmospheric Satellite Products. Ovals represent start and end terminators, rectangles indicate processes, diamonds are decisions, and rhombus represents the data which is then used in the algorithm.

for the capture of spatial variability when an appropriate preprocessing technique is applied but still requires temporal properties of observations to be considered (Figure S2 in Supporting Information S1). The potential identified with DBSCAN prompted the development of CLASP. In the next section, CLASP is described illustrating how clustering techniques can be used to identify spatiotemporal variability of trace gases in satellite observations.

3. CLASP Description

CLASP translates the spatiotemporal variability of trace gases in satellite observations into clusters that describe the feature's spatial location, magnitude variation, and temporal frequency. The algorithm is rooted in density-based clustering methods but builds on the techniques to find features with no prior knowledge of the data set and preserve feature variability. CLASP requires satellite observations on a level 3 grid (**O**) and consistent sampling frequency (e.g., 1 day, 1 hr, etc.) The algorithm asks for user inputs to define a sub-global region of interest (**Bounds**), describe the region (**MagInful**, **Source**), and control how clusters are formed (**EpsScale**, **DeltaM**, **MinPoints**, **FreqThreshold**, and **MinDates**). Details on input specifications are listed in Table S1 in Supporting Information S1. We note that CLASP requires observations to be screened using quality controls first before the user passes them into the algorithm. With inputs, CLASP moves through four steps: (a) feature detection, (b) spatial variability identification, (c) feature magnitude classification, and (d) frequency identification, which leaves each plume captured by CLASP described spatially, by magnitudes, and temporally. Figure 1 shows a flowchart of CLASP and Figure S3 in Supporting Information S1 displays the algorithm's structure. In the upcoming sections, CLASP is described in detail where a synthetic sample data set was used to illustrate and verify the methodology.

3.1. Feature Detection

The first step in CLASP is to determine features of interest. In the synthetic data set, 5 days of values are considered, where each day has three randomly generated circular-shaped features with noise. Magnitude is seen in the data set where values range from 0 to 21 and represent a generic quantitative measure. Figure 2a displays a single day in the data set for the defined region of interest, and Figure S4 in Supporting Information S1 displays results for all days. The magnitude of atmospheric trace gases varies on the spatial location and temporal interval of their observational data set, as such the features of interest can change on the same variables. Relative methods are utilized to capture the variability of atmospheric trace gas features. CLASP employs the region's magnitude values to determine features relatively for each time step and region inputted by the user. Specifically, the relative cumulative frequency of the magnitude is used to assess the region and reveal where high magnitude values encompassing a pollutant hotspot are found. Following (Lee et al., 2022; Matschullat et al., 2000) the relative cumulative frequency of the magnitude is used to determine the threshold below which data is excluded and the main pollutant burden is revealed. The relative cumulative frequency curve is fitted with a piecewise linear function which produces "turning points" to be used as threshold values. As suggested by Matschullat et al. (2000) threshold values can be interpreted as where different processes influence the region's distribution of atmospheric trace gases. The threshold points can be used to define the characteristics of the regions. Partitioning the data set at threshold points can identify the relative "background" values or samples that may not be influenced by human impacts (Matschullat et al., 2000). Conversely, partitioning can identify the samples that carry a higher magnitude, such as those influenced by anthropogenic activities, biomass-burning events, or other high-magnitude-producing processes. CLASP takes advantage of this partitioning when selecting the appropriate threshold value for the region analyzed. The input variable, **MagInflu** (**yes** or **no**), allows the user to define which characteristics of the

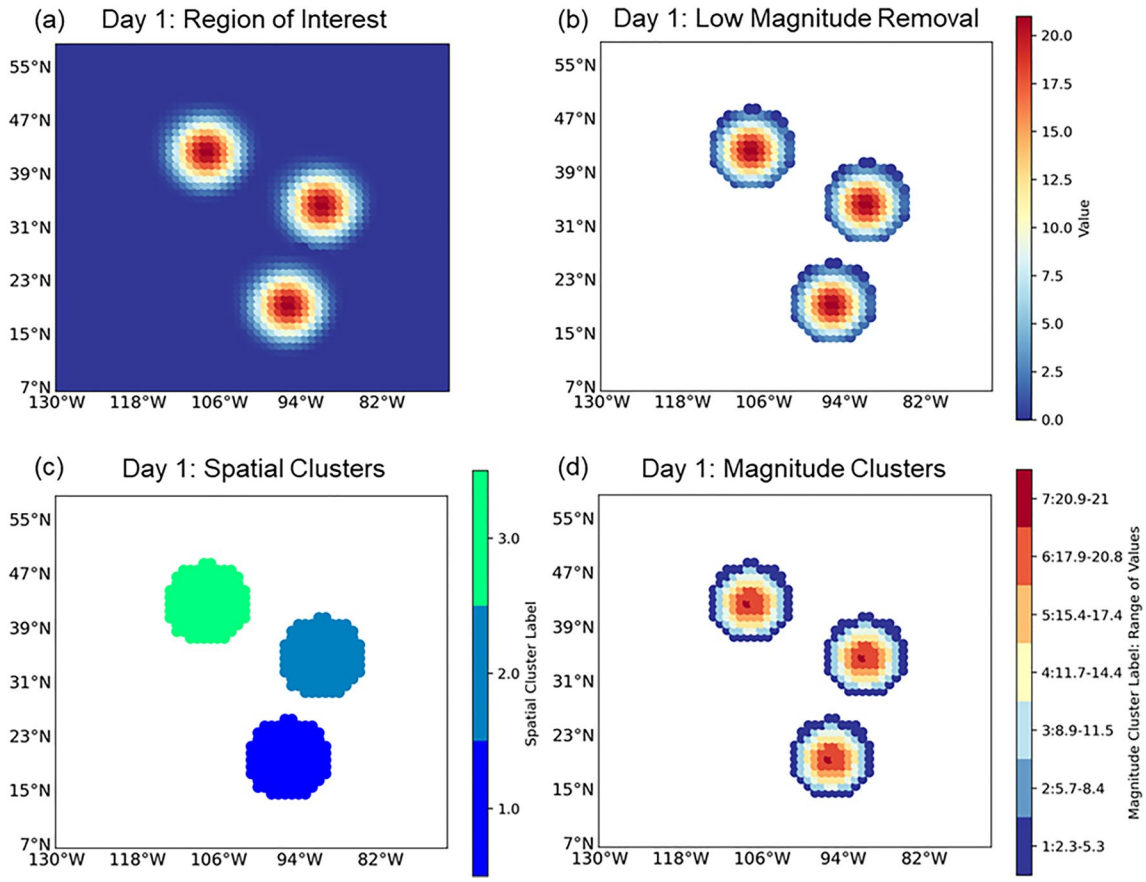

Figure 2. (a) First step in CLustering of Atmospheric Satellite Products, with inputs MagInflu = no, Source = yes, EpsScale = 1, MinPoints = 5, DeltaM = 1, freqthreshold = no, to define a region of interest, (b) second step to remove low magnitude values, (c) third step defines spatial clusters where each distinct color represents a different cluster, and (d) fourth step creates magnitude clusters where each distinct color represents a different cluster and magnitude range as indicated in the color bar. To find each day's spatial and magnitude clusters for the synthetic data set was less than 5 s.

region or its relative magnitude range it wants to observe and classify. Once **MagInflu** is applied, the region's features of interest are revealed (Figure 2b).

The input variable **Source** (**yes** or **no**) can be used to further describe which features are of interest to the user. Source determines if features from known sources are to be identified by CLASP, as shown in Figure S5 in Supporting Information S1. Sources could be emission source locations from publicly available data sets, such as EPA's Continuous Emission Monitoring Systems (CEMS), EPA's National Emission Inventory (NEI), offshore platform locations (BOEM), the U.S. Census for significant city centers, fire locations (FINN, GDEF4), or recorded lightning strikes (Geostationary Lightning Mapper (GLM)). Source location information is controlled by the user and can be tailored to specific analyses. When **Source** is **no** CLASP will exclude features that include a source point, and when **Source** is **yes** CLASP will include all features detected in the region (Figure S6 and S8 in Supporting Information S1). CLASP allows for the detection of all features in a region, controlled by the user inputs on MagInflu and Source.

3.2. Spatial Variability Identification

With features detected, CLASP moves to its second step where the spatial variability of individual features is identified resulting in spatial clusters. Atmospheric trace gases are dependent on meteorological, chemical, and human variables that control the location of the detected feature. It is important to identify the spatial variability of each feature to understand its source origins and the pollutants' impact on the region's air quality. Therefore, the latitude and longitude points of the detected features are used to estimate the spatial similarity between the features. Methods were adapted from DBSCAN for identifying spatial variability. Traditional applications of

DBSCAN see clustering along the magnitude axis to determine if points are of similar density to where a cluster can be formed. In CLASP, DBSCAN is used to cluster along the spatial axis. It is assumed that the latitude and longitude pairs of the observational data set are found on a 2-D surface. The assumption of a 2-D surface is made as it is the *similarity* between points that are being calculated. It is the similarity distance between points that are used to infer relative spatial positions. The Euclidean distance ($d(p,q) = \sqrt{\sum_i^n (q_i - p_i)^2}$) is calculated for each latitude and longitude pair, providing a similarity metric between each pair. Objects whose similarity metric is less than double the minimum spatial scale of the data set when the minimum number of points is met form a spatial cluster shown in Figure 2c.

3.3. Feature Magnitude Classification

The third step in CLASP is to classify a feature's magnitude found in each spatial cluster. The magnitude is the quantity that a satellite retrieves and describes the distribution of the severity of the pollutant in the region. For TROPOMI NO_2, the magnitude is tropospheric vertical column density. It is necessary to classify the distribution to understand the different levels of the pollutant experienced in a region and its perceived impact. Magnitude clusters are characterized by the local magnitude distribution within each spatial cluster (Wu et al., 2004). To observe the distribution of magnitudes captured in the region, values are divided into intervals using Sturge's rule ($b = [log_2 n + 1]$). It is assumed that data points close to one another have similar magnitudes or values, thus clusters are formed by identifying nearest neighbors. The input variables **EpsScale**, **DeltaM**, and **MinPoints** are used to describe the characteristics of the magnitude clusters. **EpsScale** controls the spatial size of each magnitude cluster identified, **DeltaM** allows the user to define the minimum difference magnitude clusters can possess, and **Minpoints** ensures each magnitude cluster contains a minimum number of points. Input parameters help describe the magnitude observed in each spatial cluster and form magnitude clusters that easily describe the magnitude distribution in a region as shown in Figure 2d.

3.4. Frequency Identification

After trace gas features are identified spatially and by their magnitude (shown in Figure 2), the collection of classified features is then described temporally. CLASP tracts day-to-day plume variability by creating temporal clusters, shown in Figure 3 for the sample synthetic data set. Temporal clusters are formed when the period of analysis is longer than the minimum time step of the observational data set (e.g., 1 day for TROPOMI NO_2). Identifying the frequency of features informs the user if such a feature is a repeated event providing a better understanding of the temporal variability of the pollutant in the region of interest. With each identified spatial cluster in the temporal period, the frequency of the latitude and longitude points is calculated. The unique frequency of points within each spatial cluster is used to create temporal clusters. Two input variables control temporal cluster formation, **FreqThreshold**, and **MinDates**. **FreqThreshold** allows for infrequent points to be removed revealing only regularly occurring features in the region, and **MinDates** ensures that a specified number of dates are considered. It is assumed that data points near one another will be of similar frequency, allowing for the creation of temporal clusters shown in Figure 3.

After temporal clusters are formed, CLASP outputs a data set that carries the set of observations and spatial, magnitude, and temporal cluster label descriptions for each identified feature. CLASP took approximately 10 s to identify and categorize the three sources observed (Figure S13 in Supporting Information S1). The number of sources does not linearly affect the computational time of CLASP, however, the resolution does (Figure S13 in Supporting Information S1). With each clustering output, CLASP allows the user to easily select what type of features it wants to analyze, from the location of a feature, its magnitude range, or its frequency. CLASP moves beyond popular clustering methods and allows for satellite observations of trace gases to be described more fully providing a better understanding of the spatiotemporal variability of atmospheric trace gases. CLASP relatively assesses a region and extracts important features based on the magnitude distribution. The detected features are then placed in spatial and magnitude clusters where the temporal frequency and pollutant burden of the feature can be determined. With CLASP, we see that the user does not have to have prior knowledge of a domain or time frame as CLASP can search any given region or period and detect the relatively important features. Feature variability is preserved with CLASP, as it identifies each feature and assesses its frequency. With the algorithm structure presented here, one can better understand the spatiotemporal variability of atmospheric trace gases and our atmosphere.

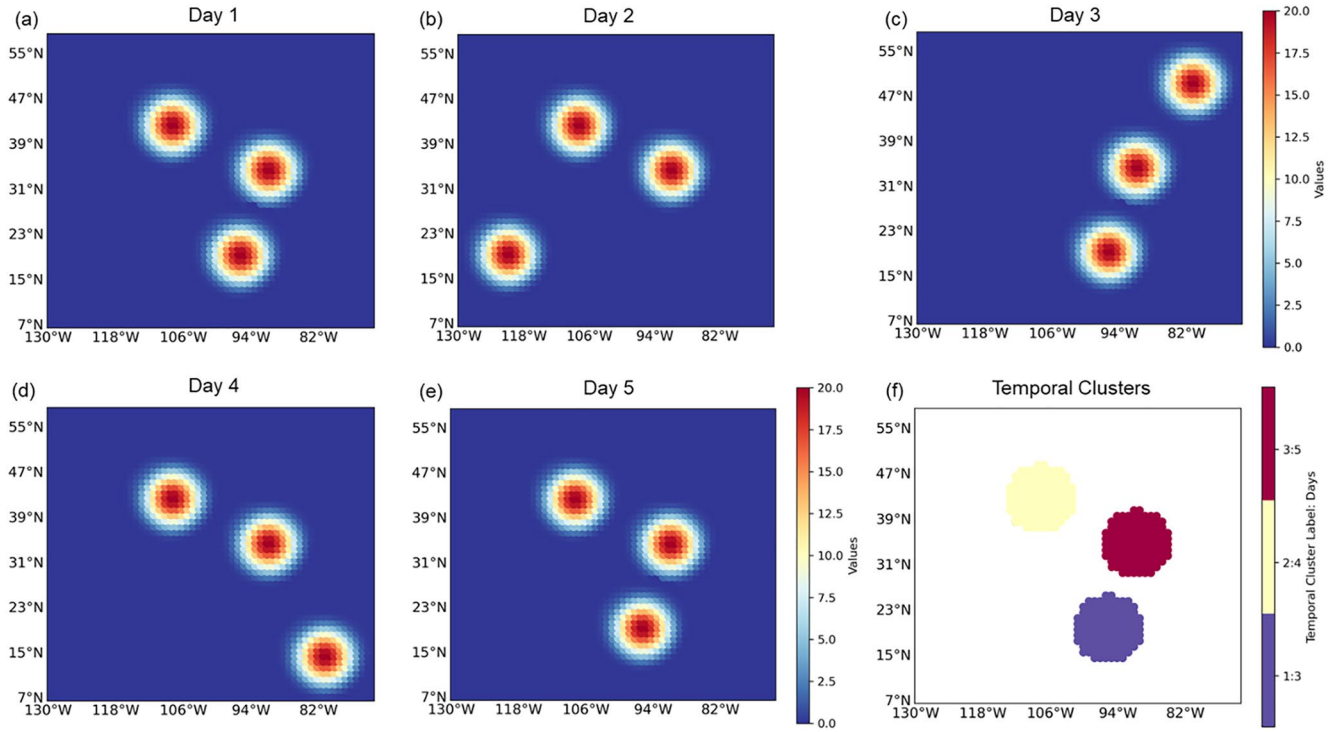

Figure 3. (a–e) The varying plume locations within the synthetic data set. (f) The last step in CLustering of Atmospheric Satellite Products (CLASP) is to construct the temporal clusters with MinDates = 2, capturing how CLASP tracks day-to-day plume variability.

4. Application With TROPOMI NO_2

To test the applicability of CLASP, CLASP was applied to TROPOMI NO_2 observations. TROPOMI is aboard the Sentinel 5 Precursor (S5P). The sun-synchronous instrument uses passive remote sensing to observe gases down into the troposphere and gives near-global coverage in 1 day with an equator crossing time near 13:30 local solar time. TROPOMI operational nitrogen dioxide (NO_2) data processor version 1.2.2–1.4.0 (KNMI, 2018, https://doi.org/10.5270/S5P-s4ljg54) and processor version 2.2.0–2.4.0 (KNMI, 2021, https://doi.org/10.5270/S5P-9bnp8q8), which carries an improved spatial resolution of 5.5 × 3.5 km at nadir are used to illustrate the capabilities of CLASP. Specific processor version information for each year used in the search space is listed in Text S1 in Supporting Information S1. Level-2 TROPOMI products where the quality assurance (qa) value is higher than 0.75, solar zenith angle less than 75° and a cloud radiance fraction below 0.3 were oversampled to 0.01° × 0.01° using a physics-based oversampling process (Sun et al., 2018). For each search space, at least 85% of quality pixels were required and dates employed had no separation of weekdays and weekends. Any exclusion of dates was due to the inability to pass the criteria employed for overpass selection for reasons such as high cloud obstructions or inability to meet quality controls. Processing a 1° × 1° grid of TROPOMI NO_2 observations oversampled to 0.01° × 0.01° using 4 CPU cores, for 1 day, 1 month, and 6 months CLASP takes ∼3, ∼90, and ∼180 s, respectively. Increasing the processing cores will increase the speed of the algorithm for larger search spaces.

The state of Texas was chosen as a search location due to its diverse landscape of industrial and commercial processes. TROPOMI NO_2 observations during January 2019–October 2022 over the state of Texas were tested and first divided into 1° × 1° grids (Figure S7 in Supporting Information S1) to reduce the size of the data being used in CLASP. CLASP searched the grid boxes including source points to identify the spatiotemporal variability of atmospheric NO_2. With outputs delivered by CLASP, three case studies were selected to show a few different capabilities of CLASP: (a) event identification, (b) spatiotemporal variability preservation, and (c) irregular source identification.

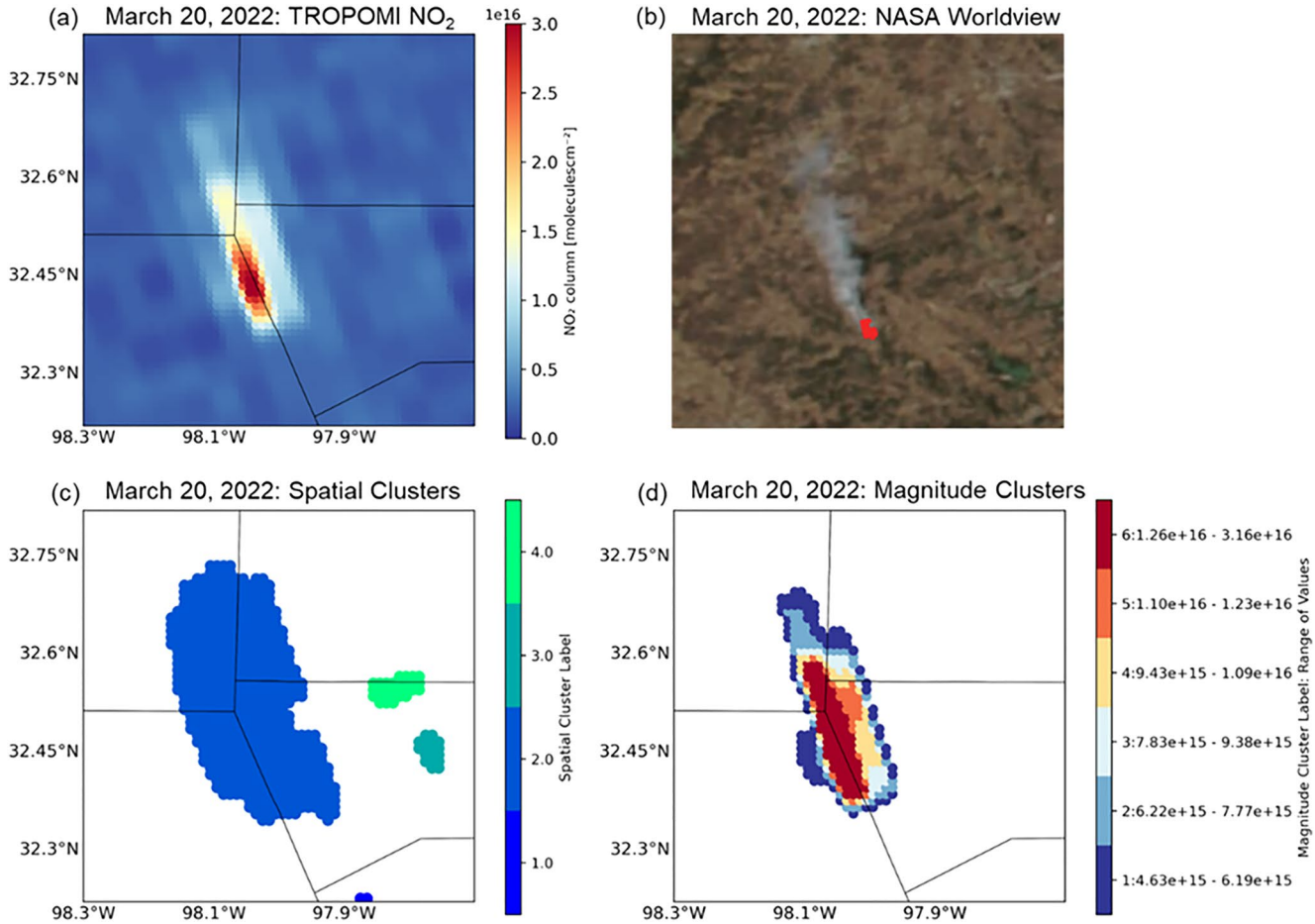

Figure 4. CLustering of Atmospheric Satellite Products (CLASP) used for wildfire identification with inputs MagInflu = yes, Source = yes, EpsScale = 5, MinPoints = 10, DeltaM = 5. (a) TROPOspheric Monitoring Instrument NO_2 observation on 20 March 2022, (b) MODIS true-color image with VIIRS fire hotspots (c) spatial clusters identified by CLASP where each distinct color represents a different cluster and (d) magnitude clusters identified by CLASP, where each distinct color represents a different cluster and magnitude range as indicated in the color bar.

4.1. Event Identification: Wildfires

One key strength of CLASP is event identification. CLASP individually searches the distribution of trace gases at the timestep of the provided data allowing for unique features to be identified. As an example, from the user-defined search space described above, CLASP identified a single emission event on 20 March 2022 located in Central Texas (USA) west of the Dallas-Fort Worth city center (Figure 4a). With CLASP, it was observed this event only occurred one time in this location and was an anomalous event for this region. Current emission inventories employed in the search did not associate this identified feature with any known source point. The joint MODIS and VIIRS (Schroeder et al., 2014) true-color image at approximately the same time as the TROPOMI overpass (Figure 4b) (obtained from NASA Worldview: https://worldview.earthdata.nasa.gov/, last access: 3 November 2022) and news reports (The New York Times, 2023, https://www.nytimes.com/2022/03/20/us/wildfires-texas-eastland-complex.html; Single Incident Information, 2023, https://inciweb.nwcg.gov/incident-information/txtxs-big-l-fire) were used together for the determination of the predominant emitter for this event. It was discovered that this event was associated with wildfire activity.

Figures 4c and 4d displays the spatial and magnitude clusters identified by CLASP. Spatial clusters reveal the location and extent of the NO_2 plumes observed in the region while the magnitude clusters display the range of values emitted by the event. Spatial clusters whose magnitude did not carry similar values to that which dominated the region were removed leaving the relative NO_2 burden for the region. The identification of this wildfire event from the user-defined grid box search of TROPOMI observations over Texas illustrates the combined effect of TROPOMI and CLASP. Since wildfire activity can be identified using other databases and methods, this is

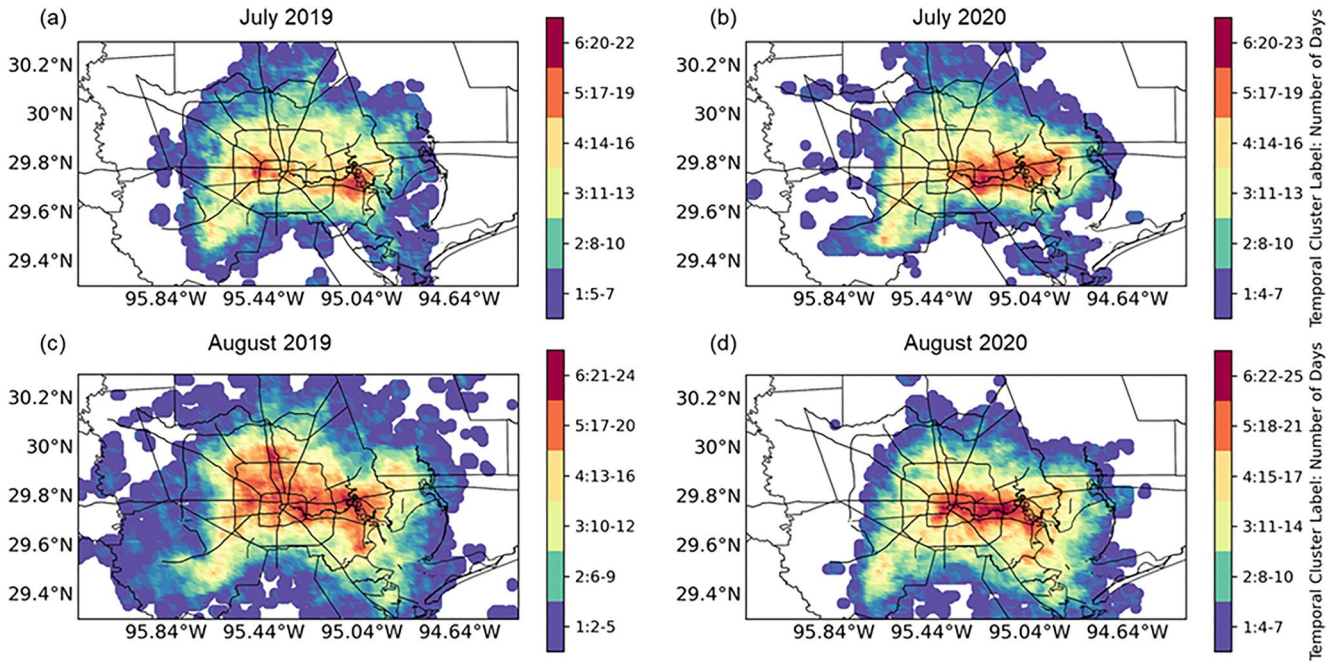

Figure 5. CLustering of Atmospheric Satellite Products (CLASP) temporal clusters in the Houston, TX area for pre-lockdown (a) July 2019 and (c) August 2019, and lockdown (b) July 2020 and (d) August 2020 conditions, with inputs MagInflu = yes, Source = y, EpsScale = 5, MinPoints = 10, DeltaM = 5, freqthreshold = no, MinDates = 2. Each color represents a different cluster corresponding to the number of dates a plume was identified by CLASP, as indicated in the color bar where the cluster number is shown with the number of dates.

not the motivator of CLASP development. However, the correct identification of the wildfire hotspot by CLASP, as supported by Figure 4b, illustrates the algorithm's validity; that is, single emission events or distinct features occurring at the sampling frequency can be captured by CLASP without knowledge of where or when the event has taken place. With high-resolution satellite observations and CLASP, trace gas distributions can be identified at the rate at which satellite observations are available.

4.2. Trace Gas Spatiotemporal Variability Preservation: COVID-19 Lockdown

CLASP captured the variations of NO_2 within a search period. To demonstrate the encapsulation of variability and its preservation, CLASP outputs during the COVID-19 pandemic shutdown over the Houston metropolitan area in Texas were analyzed. The Texas government declared a public health disaster in March 2020 which changed emission patterns across the state. Shutdown orders drove plume variability due to normality being disrupted and this disruption was captured by TROPOMI observations. Houston has a diverse emission profile where NO_x is readily found as the region is home to heavy road traffic as well as many industrial facilities. With TROPOMI observations, CLASP's temporal clusters were able to identify and preserve the variations in NO_2 emissions patterns experienced in Houston during the COVID-19 lockdown period.

Figure 5 displays CLASP temporal clusters for pre-lockdown (2019) and lockdown (2020) conditions during July and August. The months of July and August were chosen for analysis as quality retrievals from TROPOMI NO_2 following selection criteria were found across both periods and were included during the shutdown period (March 2020–March 2021, Houston Public Media, 2021). County lines and primary roads (TIGER/Line Shapefiles: https://www.census.gov/cgi-bin/geo/shapefiles/index.php) are overlaid on the temporal clusters. For each period, the overall emission pattern is similar but specific variations in the location of frequent plume occurrences are seen. In pre-lockdown conditions, Figures 5a and 5c, temporal clusters of NO_2 are found throughout the region, near major highways and industrial areas with similar frequency. In lockdown conditions, Figures 5b and 5d, there is a decrease in the daily frequency of temporal clusters around major highways and an increase in the frequency of NO_2 clusters near an industrial center, commonly known as the Houston Ship Channel. Spatiotemporal variations are due to shutdown orders causing a decline in road traffic emissions when many industrial facilities were deemed essential services and continued operations. Archer et al. (2020), Goldberg et al. (2020), and Fioletov et al. (2022)

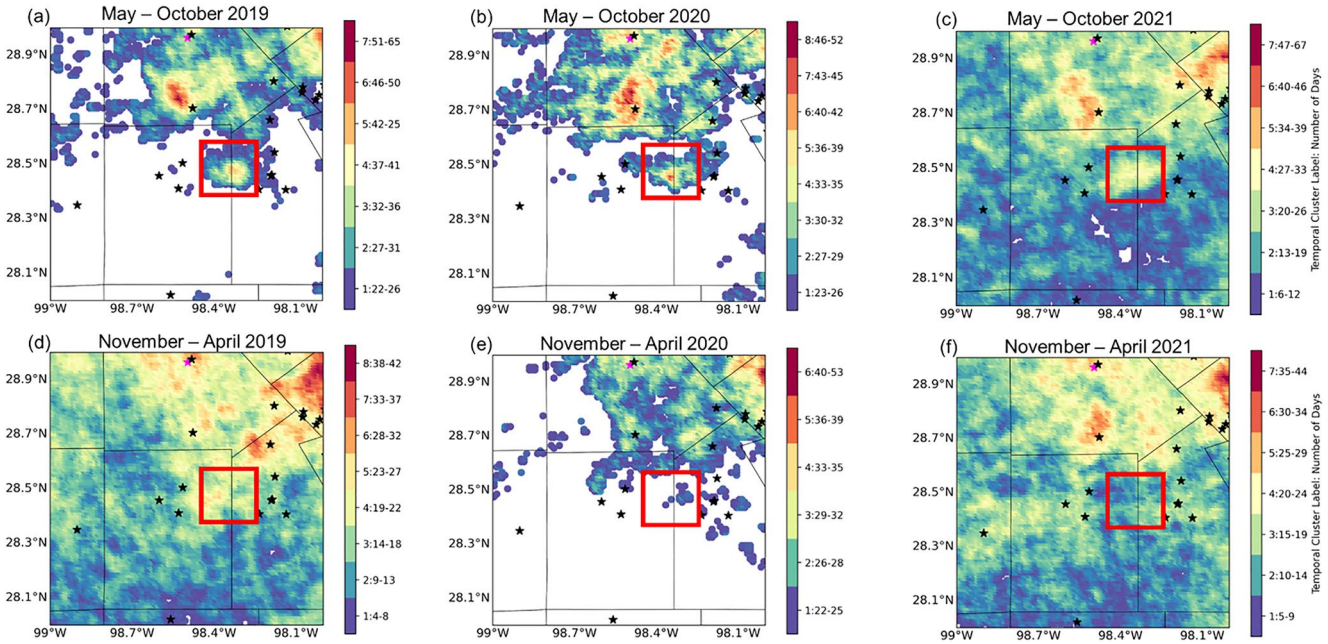

Figure 6. CLustering of Atmospheric Satellite Products (CLASP) temporal clusters identifying an irregular source delineated with a red box near San Antonio, TX strongly observed in May-October (a) 2019, (b) 2020, (c) 2021, and not strongly observed in November–April (d) 2019, (e) 2020, and (f) 2021, with inputs MagInflu = yes, Source = yes, EpsScale = 5, MinPoints = 10, DeltaM = 5, freqthreshold = no, MinDates = 4. Each color represents a different cluster corresponding to the number of dates a plume was identified by CLASP, as indicated in the color bar where the cluster number is shown with the number of dates. Black stars identify source point locations from EPA's 2017 National Emissions Inventory. Pink stars identify cities with a population greater than 10,000 people as reported by the 2020 United States census.

observed COIVD-19 driven changes in NO_2 seen from TROPOMI and OMI, where column density declines for the lockdown period of April 2–20, 2020 in Houston range from ∼15% to 20%. Specific column density variations in identified plumes during July and August 2019-2020 are further reflected in CLASP's magnitude clusters. There is a 9.2% and 6.7% decrease pre (2019) and lockdown (2020) conditions for July and August, respectively.

Traditional data processing methods such as averaging shown in Figure S10 in Supporting Information S1, do not capture the plume variations during this period and dampen the location of the observed plume differences driven by COVID-19 changes. By contrast, the variability of plume location, shape, and frequency is preserved when using CLASP. The temporal clusters of CLASP allow for the observation of the cumulative impact of a trace gas, where averaging displays the mean plume shape and no information about the temporal attributes of the data set. CLASP's temporal clusters go beyond averaging and allow the user to be informed of where and the frequency of trace gas emissions captured by satellite observations.

4.3. Irregular Sources: Oil and Gas Operations

During the course of the search, CLASP identified irregular emission patterns, an example of which is located south of San Antonio, TX, as shown in Figure 6 delineated with a red box. This emission source occurs at intermittent intervals throughout the search period with varying frequencies. The identified infrequent source is in the Eagle Ford basin, a prominent oil and gas-producing region home to numerous active gas and oil well sites (Figure S11 in Supporting Information S1). It is expected that this infrequent source is associated with oil and gas production. In oil and gas operations, NO_2 is a pollutant that is released by equipment (internal combustion engines) used to run operations including drilling and gas flaring (de Gouw et al., 2020; Dix et al., 2020). NO_x emissions are not linearly emitted in oil and natural gas operations (Dix et al., 2020) due to the varied use of equipment, flaring practices, and vehicular emissions observed during good production. The varied NO_x emissions from oil and gas operations create many irregular NO_x sources across the state of Texas which may go unnoticed using traditional data processing methods (Figure S12 in Supporting Information S1) or to the lack of knowledge of activities. Figure 6 displays temporal clusters for 2019, 2020, and 2021 separated into two 6-month periods within each year (May-October, November-April). From May-October for

each year, there is an increase in the number of NO_2 plumes identified by CLASP correlating to a frequent NO_2 source during these months. When each year's first 6-month period (May-October) is compared to the second 6-month period (November-April), there is a clear decrease in the number of occurrences that NO_2 clusters identified by CLASP. Seasonality is observed in oil and gas production emissions but is largely dependent on the stages of production (Dix et al., 2020). With CLASP, the seasonality of specific oil and gas wells can be identified and observed as well as the frequency and location of their emissions. Traditional averaging of the yearly six-month periods (Figure S12 in Supporting Information S1) diminished the prominence of the CLASP-identified features. As a result, the determination of its importance cannot be concluded from the averaging results alone. With CLASP, one can now observe the seasonality of the emissions and know that the intermittent feature has a similar temporal frequency to a nearby known source point. Further analysis of this NO_2 hotspot can be performed to understand its impact on regional air quality during high emitting periods. With CLASP's temporal clusters, irregular signals captured by satellite observations can be identified.

Above are a few examples of applications where CLASP can be implemented: identification of single emission events, plume variability, and irregular emission sources. With each example, prior spatial and temporal knowledge of an event is not known, yet CLASP can identify and preserve NO_2 spatiotemporal variability. However, it should be noted that the uses of CLASP are not limited to the examples shared here. Numerous applications can be seen with the combined use of satellite observations of trace gases and CLASP. CLASP can be applied to large data sets of satellite observations and locate important features observed in a region where the air quality can then be assessed.

5. Conclusions and Future Work

This work developed a clustering algorithm, CLASP to identify and preserve gaseous variability in satellite observations of trace gases. The method can be automated and rooted in density-based clustering techniques but goes further than traditional data processing, supervised deep learning, or common clustering methods to adequately capture the variability of short-lived reactive gases. Long-term averaging can dampen gaseous variability unintentionally which can lead to an incomplete understanding of trace gas distribution. CLASP can identify and describe time-varying trace gases and sources where traditional data processing methods cannot. With a few inputs CLASP can reproducibly identify and describe a region's observed signals (a) spatially, (b) by magnitude, and (c) temporally. As an example of application, CLASP was applied to TROPOMI NO_2 observations across the state of Texas for the identification of single emission events like wildfires, plume variability observed with COVID-19 lockdown restrictions, and identifying irregular sources from oil and gas operations. Such information obtained in applications can be used to gain a better understanding of a region's air quality.

With the continued missions of satellite instruments detecting atmospheric composition (GOME, OMI, SCIAMACHY, TROPOMI) and with the recent launch of missions such as TEMPO (launch date 7 April 2023), the data size of satellite observations of trace gases is reaching an exponential amount. The data size of observations in conjunction with traditional data processing methods creates difficulties in interpretation and can hinder the use of such quality observations, as features may be dampened or missed. When CLASP and satellite observations of trace gases are combined, understanding of the spatiotemporal variability of trace gases and our atmosphere is increased. With the use of CLASP, the spatial and temporal information on identified features can be used in studies to better understand processes and allow for quantification using top-down methods to take place. As satellite observations are spatially and temporally extensive, overlooked locations, where ground-based measurements are lacking, can have their air quality assessed using CLASP. Further, the distribution of the pollutant burden in a region can be determined to where disparities can be identified with CLASP. Information obtained by CLASP can be used in the development of policies from decision-makers as they look to understand processes controlling air quality, identify long and short-term trends and emission sources, as well as understand the distribution of a pollutant in a region that could aid in the development of plans for mitigation strategies. Finally, CLASP can be employed for data sets of other trace gases (i.e., methane, sulfur dioxide, ozone, etc.) and other identification scenarios. CLASP has a wide range of pertinency and is not limited to the trace gas of NO_2 or TROPOMI observations. With the newly launched geostationary satellite instrument over North America, TEMPO, providing hourly column observations of up to 2 km × 4.5 km at nadir and CLASP our understanding of trace gas variability will be greatly increased to where we can gain a better understanding of our atmosphere and its processes.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The source code in python and synthetic data used for CLASP and its examples are available at https://github.com/tabitha-lee/CLASP via https://doi.org/10.5281/zenodo.7709613 (Lee, 2023a). A Jupyter Notebook demonstrating how to use CLASP and create Figure 1 is available at https://github.com/tabitha-lee/CLASP/blob/main/How%20to%20use%20CLASP.ipynb, is hosted at GitHub and is preserved at https://doi.org/10.5281/zenodo.7709613 (Lee, 2023b).

Acknowledgments
This work is supported by NASA ROSES-20 Future Investigators in NASA Earth and Space Science and Technology (FINESST) Program (80NSSC21K1616). The authors acknowledge the Tropospheric Monitoring Instrument (TROPOMI) group for providing offline total NO_2 column measurements (available https://s5phub.copernicus.eu/dhus/#/home or https://earthdata.nasa.gov/). The authors acknowledge the United States Environmental Protection Agency for providing source point emission data (available at https://www.epa.gov/air-emissions-inventories/air-pollutant-emissions-trends-data), Continuous Emission Monitoring System (CEMS) data from power plants (available at https://ampd.epa.gov/ampd/), Bureau of Ocean Energy Management (BOEM) for offshore platform locations (available at https://www.boem.gov/environment/environmental-studies/ocs-emissions-inventories), the U.S. Census for significant city centers (available at https://www.census.gov/), Fire INventory from NCAR (FINN) for fire locations (available at https://www2.acom.ucar.edu/modeling/finn-fire-inventory-ncar), Geostationary Lightning Mapper (GLM) for recorded lightning strikes (available at www.class.noaa.gov), and the Railroad Commission of Texas for oil and gas well locations in Texas (available at https://www.rrc.texas.gov/resource-center/research/data-sets-available-for-download/). The authors acknowledge the use of imagery from the NASA Worldview application (https://worldview.earthdata.nasa.gov), part of the NASA Earth Observing System Data and Information System (EOSDIS). The authors thank Dr. Kang Sun for providing the physics-based oversampling algorithm, whose source code is available online (https://github.com/Kang-Sun-CfA/Oversampling_matlab/blob/master/popy.py).

References

Ahmed, K. N., & Razak, T. A. (2016). IJARCCE an overview of various improvements of DBSCAN algorithm in clustering spatial databases. *International Journal of Advanced Research in Computer and Communication Engineering*, *5*(2), 360. https://doi.org/10.17148/IJARCCE.2016.5277

Archer, C. L., Cervone, G., Golbazi, M., Al Fahel, N., & Hultquist, C. (2020). Changes in air quality and human mobility in the USA during the COVID-19 pandemic. *Bulletin of Atmospheric Science & Technology*, *1*(3–4), 491–514. https://doi.org/10.1007/s42865-020-00019-0

Birant, D., & Kut, A. (2007). ST-DBSCAN: An algorithm for clustering spatial-temporal data. *Data & Knowledge Engineering*, *60*(1), 208–221. https://doi.org/10.1016/j.datak.2006.01.013

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noel, S., Rozanov, V. V., et al. (1999). SCIAMACHY: Mission objectives and measurement modes. *Journal of the Atmospheric Sciences*, *56*(2), 127–150. https://doi.org/10.1175/1520-0469(1999)056<0127:smoamm>2.0.co;2

Burrows, J. P., Weber, M., Buchwitz, M., Rozanov, V., Ladstatter-Weibenmayer, A., Richter, A., et al. (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)0562.0.CO;2

Copernicus Sentinel Data Processed by ESA, Koninklijk Nederlands Meteorologisch Instituut (KNMI). (2018). *Sentinel-5P TROPOMI tropospheric NO_2 1-Orbit L2 7 km × 3.5 km*. Goddard Earth Sciences Data and Information Services Center (GES DISC). https://doi.org/10.5270/S5P-s4ljg54

Copernicus Sentinel Data Processed by ESA, Koninklijk Nederlands Meteorologisch Instituut (KNMI). (2021). *Sentinel-5P TROPOMI tropospheric NO_2 1-Orbit L2 5.5 km × 3.5 km*. Goddard Earth Sciences Data and Information Services Center (GES DISC). https://doi.org/10.5270/S5P-9bnp8q8

de Gouw, J. A., Veefkind, J. P., Roosenbrand, E., Dix, B., Lin, J. C., Landgraf, J., & Levelt, P. F. (2020). Daily satellite observations of methane from oil and gas production regions in the United States. *Scientific Reports*, *10*(1), 1379. https://doi.org/10.1038/s41598-020-57678-4

Dix, B., de Bruin, J., Roosenbrand, E., Vlemmix, T., Francoeur, C., Gorchov-Negron, A., et al. (2020). Nitrogen oxide emissions from U.S. Oil and gas production: Recent trends and source attribution. *Geophysical Research Letters*, *47*(1), 1–9. https://doi.org/10.1029/2019GL085866

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. Retrieved from www.aaai.org

Finch, D. P., Palmer, P. I., & Zhang, T. (2022). Automated detection of atmospheric NO_2 plumes from satellite data: A tool to help infer anthropogenic combustion emissions. *Atmospheric Measurement Techniques*, *15*(3), 721–733. https://doi.org/10.5194/amt-15-721-2022

Fioletov, V., McLinden, C. A., Griffin, D., Krotkov, N., Liu, F., & Eskes, H. (2022). Quantifying urban, industrial, and background changes in NO_2 during the COVID-19 lockdown period based on TROPOMI satellite observations. *Atmospheric Chemistry and Physics*, *22*(6), 4201–4236. https://doi.org/10.5194/acp-22-4201-2022

Franklin, M., Chau, K., Cushing, L. J., & Johnston, J. E. (2019). Characterizing flaring from unconventional oil and gas operations in south Texas using satellite observations. *Environmental Science and Technology*, *53*(4), 2220–2228. https://doi.org/10.1021/acs.est.8b05355

Goldberg, D. L., Anenberg, S. C., Griffin, D., McLinden, C. A., Lu, Z., & Streets, D. G. (2020). Disentangling the impact of the COVID-19 lockdowns on urban NO_2 from natural variability. *Geophysical Research Letters*, *47*(17), e2020GL087978. https://doi.org/10.1029/2020GL089269

Goldberg, D. L., Lu, Z., Streets, D. G., de Foy, B., Griffin, D., Mclinden, C. A., et al. (2019). Enhanced capabilities of TROPOMI NO_2: Estimating NO_x from North American cities and power plants. *Environmental Science and Technology*, *53*(21), 12594–12601. https://doi.org/10.1021/acs.est.9b04488

Griffin, D., McLinden, C. A., Dammers, E., Adams, C., Stockwell, C. E., Warneke, C., et al. (2021). Biomass burning nitrogen dioxide emissions derived from space with TROPOMI: Methodology and validation. *Atmospheric Measurement Techniques*, *14*(12), 7929–7957. https://doi.org/10.5194/amt-14-7929-2021

Houston Public Media. (2021). Newsroom, Texas. "How We got here: A timeline of gov. Greg abbott's covid policies". Retrieved from https://www.houstonpublicmedia.org/articles/news/politics/2021/08/20/406474/how-we-got-here-a-timeline-of-gov-greg-abbotts-covid-policies/

Ialongo, I., Virta, H., Eskes, H., Hovila, J., & Douros, J. (2020). Comparison of TROPOMI/Sentinel-5 Precursor NO_2 observations with ground-based measurements in Helsinki. *Atmospheric Measurement Techniques*, *13*(1), 205–218. https://doi.org/10.5194/amt-13-205-2020

Ingmann, P., Veihelmann, B., Langen, J., Lamarre, D., Stark, H., & Courrèges-Lacoste, G. B. (2012). Requirements for the GMES atmosphere service and ESA's implementation concept: Sentinels-4/-5 and-5p. *Remote Sensing of Environment*, *120*, 58–69. https://doi.org/10.1016/j.rse.2012.01.023

Kim, J., Jeong, U., Ahn, M.-H., Kim, J. H., Park, R. J., Lee, H., et al. (2020). New Era of air quality monitoring from space: Geostationary environment monitoring spectrometer (GEMS). *Bulletin of the American Meteorological Society*, *101*(1), E1–E22. https://doi.org/10.1175/bams-d-18-0013.1

Lee, T. (2023a). CLASP (version 1.0) [Software]. Zendo. https://doi.org/10.5281/zenodo.7709613

Lee, T. (2023b). Jupyter Notebook how to use CLASP (version 1.0) [Computational Notebook]. Zendo. https://doi.org/10.5281/zenodo.7709613

Lee, T., Wang, Y., & Sun, K. (2022). Impact of Hurricane Ida on nitrogen oxide emissions in southwestern Louisiana detected from space. *Environmental Science and Technology Letters*, *9*(10), 808–814. https://doi.org/10.1021/acs.estlett.2c00414

Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., Malkki, A., Huib Visser, A., Johan de Vries, H., et al. (2006). The ozone monitoring instrument. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(5), 1093–1101. https://doi.org/10.1109/TGRS.2006.872333

Lu, X., Wang, J., Yan, Y., Zhou, L., & Ma, W. (2021). Estimating hourly PM2.5 concentrations using Himawari-8 AOD and a DBSCAN-modified deep learning model over the YRDUA, China. *Atmospheric Pollution Research*, *12*(2), 183–192. https://doi.org/10.1016/j.apr.2020.10.020

Matschullat, J., Ottenstein, R., & Reimann, C. (2000). Geochemical background—Can we calculate it? *Environmental Geology*, *39*(9), 990–1000. https://doi.org/10.1007/s002549900084

Munro, R., Lang, R., Klaes, D., Poli, G., Retscher, C., Lindstrot, R., et al. (2016). The GOME-2 instrument on the metop series of satellites: Instrument design, calibration, and level 1 data processing—An overview. *Atmospheric Measurement Techniques*, *9*(3), 1279–1301. https://doi.org/10.5194/amt-9-1279-2016

National Academies of Sciences, Engineering, and Medicine. (2018). *Thriving on our changing planet: A decadal strategy for Earth observation from space*. The National Academies Press. https://doi.org/10.17226/24938

Pandey, S., Gautam, R., Houweling, S., Denier Van Der Gon, H., Sadavarte, P., Borsdorff, T., et al. (2019). Satellite observations reveal extreme methane leakage from a natural gas well blowout. *Proceedings of the National Academy of Sciences*, *116*(52), 26376–26381. https://doi.org/10.1073/pnas.1908712116/-/DCSupplemental

Schroeder, W., Oliva, P., Giglio, L., & Csiszar, I. A. (2014). The new VIIRS 375 m active fire detection data product: Algorithm description and initial assessment. *Remote Sensing of Environment*, *143*, 85–96. https://doi.org/10.1016/j.rse.2013.12.008

Single Incident Information. (2023). Tx big L fire information. Retrieved from https://inciweb.nwcg.gov/incident-information/txtxs-big-l-fire

Sun, K., Zhu, L., Cady-Pereira, K., Chan Miller, C., Chance, K., Clarisse, L., et al. (2018). A physics-based approach to oversample multi-satellite, multispecies observations to a common grid. *Atmospheric Measurement Techniques*, *11*(12), 6679–6701. https://doi.org/10.5194/amt-11-6679-2018

The New York Times. (2023). Weather conditions continue to fuel Texas fires. Retrieved from https://www.nytimes.com/2022/03/20/us/wildfires-texas-eastland-complex.html

van der A, R. J., de Laat, A. T. J., Ding, J., & Eskes, H. J. (2020). Connecting the dots: NO_x emissions along a West siberian natural gas pipeline. *Npj Climate and Atmospheric Science*, *3*(1), 16. https://doi.org/10.1038/s41612-020-0119-z

Veefkind, J. P., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., et al. (2012). TROPOMI on the ESA sentinel-5 precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sensing of Environment*, *120*, 70–83. https://doi.org/10.1016/j.rse.2011.09.027

Wu, X., Chen, Y., Brooks, B. R., & Su, Y. A. (2004). The local maximum clustering method and its application in microarray gene expression data analysis. In *EURASIP journal on applied signal processing (issue 1)*.

Zoogman, P., Liu, X., Suleiman, R. M., Pennington, W. F., Flittner, D. E., Al-Saadi, J. A., et al. (2017). Tropospheric emissions: Monitoring of pollution (TEMPO). *Journal of Quantitative Spectroscopy and Radiative Transfer*, *186*, 17–39. https://doi.org/10.1016/j.jqsrt.2016.05.008

References From the Supporting Information

Eskes, H. J., van Geffen, J., Boersma, F., Eichmann, K.-U., Apituley, A., Pedergnana, M., et al. (2023). Sentinel-5 precursor/TROPOMI level 2 product user manual nitrogen dioxide. Retrieved from https://sentinel.esa.int/documents/247904/3541451/Sentinel-5P-Nitrogen-Dioxide-Level-2-Product-Readme-File.%20pdf