2021

A METHODOLOGY FOR THE CREATION OF VOLCANIC GAS HAZARD MAPS USING SATELLITE-DERIVED SULFUR DIOXIDE

Sanna J. Mairet
Michigan Technological University, sjmairet@mtu.edu

A METHODOLOGY FOR THE CREATION OF VOLCANIC GAS HAZARD MAPS USING SATELLITE-DERIVED SULFUR DIOXIDE

By

Sanna J. Mairet

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Geology

MICHIGAN TECHNOLOGICAL UNIVERSITY

2021

This thesis has been approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE in Geology.

Department of Geological and Mining Engineering and Sciences

Thesis Advisor:	*Dr. Simon A. Carn*
Committee Member:	*Dr. Ann Maclean*
Committee Member:	*Dr. William Rose*
Department Chair:	*Dr. Aleksey Smirnov*

Table of Contents

List of Figures

List of Tables

Acknowledgements

This research was funded in part thanks to the Michigan Space Grant Consortium. I would like to express my sincere gratitude to my advisors, Dr. Simon Carn and Dr. Ann Maclean, for their support and guidance. Thanks to Dr. William Rose and all my professors at Michigan Tech for their insightful comments and suggestions. I would also like to thank my partner, my father, my sister, and the rest of my family for their encouragement, support, patience, and proofreading.

Abstract

Sulfur dioxide (SO_2) gas has been shown to be detrimental to human and environmental health and is emitted continuously from anthropogenic and volcanic sources. Sulfur dioxide is the main target gas used for the detection of hazardous volcanic plumes due to its ease of detection by satellite sensors. However, quantitative information on potential ground-level exposure to volcanic SO_2 (i.e., a volcanic gas 'hazard map') is currently unavailable for the vast majority of active volcanoes. Utilizing sulfur dioxide vertical column densities retrieved from the Ozone Monitoring Instrument on NASA's Aura satellite, Gridded Population of the World v.4, planetary boundary layer height data, and existing inventories of volcanic emitters, along with custom Python scripts and ArcGIS, this research presents a model to visualize and quantify exposure to persistent volcanic sulfur dioxide emissions. Results include a methodology for the creation of volcanic gas hazard maps and a series of global and regional maps displaying sulfur dioxide data and hazard information. This research is potentially applicable to many fields of study including policy and public health, volcanic hazard mitigation and forecasting, and environmental management.

Supplemental Materials:

All map products can be found in the supplemental materials attached with this thesis in ProQuest. Map products are available by contacting the author. There are 12 global maps for each year of data from 2005 to 2016 displaying uncorrected SO_2 VCDs in Dobson units with each region displayed in an inset map. There are 84 regional maps of corrected SO_2 vertical column density in Dobson units consisting of 12 maps for each of the 7 regions of interest corresponding to each year of data. There are 252 maps displaying regional SO_2 concentrations in parts-per-billion: 3 for each of the planetary boundary layer conditions for each of the 12 years of data for all 7 regions. Similarly, there are 252 SO_2 hazard maps displaying population count per area within the hazard zones bound by the WHO's AQG and regional NAAQS.

1 Introduction

1.1 Health Implications of Sulfur Dioxide Exposure

Exposure to sulfur dioxide (SO_2) gas, even at low concentrations, can cause severe health effects including respiratory illness, particularly among children with asthma. The U.S. Environmental Protection Agency (EPA) recommends an exposure limit of no more than 50-75 parts-per-billion (ppb) of short-term sulfur dioxide for humans and no more than 500 ppb over 3 hours per year for plants, animals, and infrastructure (U.S. EPA, 2017). Globally, volcanic sources passively degas nearly 63 kilotons/day (kt/day) of SO_2 (Carn et al., 2017), representing a significant source of natural air pollution in some volcanic regions and a chronic but often unrecognized hazard to human health (e.g., Hansell and Oppenheimer, 2004; Longo et al., 2005, 2008; van Manen, 2014; Brown et al., 2017). For example, Delmelle et al. (2002) measured up to 230 ppb of SO_2 at many sites around Masaya volcano in Nicaragua. Scientists have also described a "notable mortality crisis" in Iceland following the 1783-84 eruption of Laki volcano, with impacts reaching as far as Italy (Grattan et al., 2003). Hansell and Oppenheimer (2004) estimated that at least 455 million people globally were potentially exposed to volcanic gas hazards in 2004. Prior research has shown that volcanic eruptions and the subsequent SO_2 exposure threaten public health through increased morbidity and mortality on potentially significant scales (e.g., Brown et al., 2017). Current research is being conducted at Masaya in Nicaragua on the health impacts of the local population posed by Masaya's vumo, or persistent volcanic gas emissions (UNRESP, 2017). However, quantitative information on potential exposure to volcanic sulfur dioxide (i.e., a volcanic gas 'hazard map') is currently unavailable for the vast majority of active volcanoes.

In addition to detrimental health effects, noxious volcanic gases can cause severe disruption to local vegetation and cropland (van Manen, 2014). Delmelle et al. (2002) found that periods of high degassing resulted in a large decrease in coffee production in the Nicaraguan highlands near Masaya volcano. For the local communities of Pacaya and Panama who rely on agriculture for their livelihood, such devastation has major implications for their economic viability and community sustainability (Delmelle et al., 2002).

1.2 Review of Volcanic Hazard Maps

Calder et al. (2015) performed a preliminary review of 120 hazard maps from government or academic institutions and mentioned no hazard maps related to volcanic gas hazards. Previous studies involving volcanic hazard mapping focus on eruptive or

post-eruptive hazards maps with lahars and pyroclastic density currents making up most of the volcanic hazard maps followed by tephra, ballistics, and lava. Dunkley and Young (2000) cite three previous studies on volcanic hazard maps. Baxter et al. (1999) created a probabilistic hazard map on asphyxia risk due to carbon dioxide concentrations in the soil at Furnas volcano, Azores. Crandall et al. (1984) created a simplified administrative gas hazard map at the Dieng Plateau on Java, Indonesia, showing known gas vents and primary and secondary danger zones. Barquero and Fernandez (1990) created a qualitative gas hazard map showing zones of high corrosion to machinery, irritation to humans, and sulfur smell around Poas volcano in Costa Rica. These are all simplified maps based on ground-level observations (Crandall et al., 1984, and Barquero & Fernandez, 1990) or soil and air samples (Baxter et al., 1999).

A study of Masaya volcano in Nicaragua by van Manen (2014) involved sampling volcanic gases, including SO_2, at various locations and creating an interpolated layer of SO_2 concentrations. This was overlain with qualitative health effect data and effects on infrastructure, crops, and cattle. A more recent study is by Barsotti (2020) on the creation of probabilistic hazard maps during a single eruption in Bárðarbunga (Holuhraun), Iceland using differential optical absorption spectroscopy (DOAS) measurements of SO_2 and the CALPUFF atmospheric dispersion model. At the time of writing, similar volcanic gas exposure forecasts are also being made by the Icelandic Meteorological Offic for the ongoing eruption of Fagradalsfjall (Geldingadalir) on Iceland's Reykjanes peninsula (https://www.vedur.is/eldfjoll/eldgos-a-reykjanesi/gasmengun/). With the exception of Barsotti's (2020) probabilistic hazard map, there is a dearth of recent mapping studies involving volcanic gas hazards despite the potentially significant long-term health hazards of volcanic gas exposure at many active volcanoes. The advancement of satellite technology since the publication of many of these studies makes volcanic gas hazard mapping more quantifiable and more tractable.

1.3 Volcanic Sulfur Dioxide and Other Gases

Sulfur dioxide gas is commonly used for detection and monitoring of magmatic degassing at active volcanoes due to its detectability by remote sensing techniques through UV and IR absorption; its low or negligible background concentrations; and its potential atmospheric, environmental, and climatic impacts (e.g., Robock, 2000; Delmelle, 2002; van Manen, 2014). Volcanic SO_2 emissions may indicate the likely presence of other volcanic gases such as carbon dioxide (CO_2), water vapor (H_2O), hydrogen sulfide (H_2S), hydrogen chloride (HCl), and hydrogen fluoride (HF) as well as trace elements including radon, mercury, and lead. SO_2 can be used as a constraint parameter for flux estimations for these gases if their ratio to SO_2 is known (Carn et al., 2017).

Sulfur dioxide measurements can be used both independently and as an indicator for the potential presence of other gases that may cause detrimental health and environmental effects. Particulate matter can also accompany sulfur dioxide and other hazardous gases during both eruptive and noneruptive degassing. After emission, volcanic SO2 converts to sulfuric acid (sulfate; H_2SO_4) aerosol on variable timescales (hours to months) depending on the altitude of emission; this sulfate aerosol is an important source of fine particulate matter ($PM_{2.5}$) in the atmosphere, which is also a health hazard. However, Sears et al. (2013) found that satellite-retrieved sulfur dioxide was not a good proxy for the presence of satellite-detected volcanic ash for the purposes of aviation hazard avoidance.

1.4 Impact of Research

This research project has the potential to contribute to multiple fields. It is the first comprehensive visualization of population exposure to volcanic SO_2 emissions on a global scale and represents a novel use of NASA's Ozone Monitoring Instrument (OMI) satellite SO_2 measurements and other publicly available datasets. By identifying correlations and "hot spot" patterns between SO_2 concentrations, population, and volcano locations, this study locates areas of potentially high SO_2 exposure where populations may be at greater risk of adverse health effects or economic hardship. These regions are visualized in four series of maps created using data from 2005-2016 and highlight spatial and quantitative relationships between SO_2 concentrations and human population count.

This study also presents a new methodology and corresponding open-source Python computer codes for converting satellite-derived SO_2 vertical column amounts in Dobson Units (DU; 1 DU = 2.69×10^{16} molecules cm^{-2}) to concentrations in parts-per-billion (ppb) for subsequent comparison to global and regional ambient air quality standards. This allows for potential SO_2 exposure analysis and incorporation into GIS hazard maps for regions around major persistently degassing volcanoes.

2 Methodology

2.1 Data Preparation

Data preparation was performed on all datasets used in this research. This involved converting data types, importing data into pre-made geodatabases, and visualizing data in ArcMap. Some of these processes were performed with a custom Python script and some were done manually in ArcMap. Global data was stored in a global geodatabase to be distributed to regional geodatabases.

2.1.1 Sulfur dioxide Data

The SO_2 dataset used is daily SO_2 total vertical column amounts retrieved by the UV Ozone Monitoring Instrument (OMI) on NASA's Aura satellite from 2005 to 2016. The OMI has a spatial resolution at nadir of 13 by 24 kilometers. The current OMI SO_2 retrieval algorithm uses a principal component analysis (PCA) technique (Li et al., 2013) to retrieve SO_2 vertical column densities (VCDs) in Dobson Units representing the estimated total number of SO_2 molecules in the entire atmospheric column above a unit area. Because the OMI SO_2 sensitivity varies with altitude, the retrieved total column amount depends on the assumed vertical distribution of SO_2 and hence several fixed SO_2 vertical profiles are assumed in operational OMI retrievals. The OMI retrievals used here assume that all SO_2 is in the planetary boundary layer (PBL, or the lowest ~1 km of the atmosphere). For the present study, OMI pixels with a radiative cloud fraction >0.2 or solar zenith angle >70° (i.e., cloudy and/or high latitude pixels) were also excluded from data analysis. Data from OMI cross-track positions (or rows) affected by the OMI row anomaly (http://projects.knmi.nl/omi/research/product/rowanomaly-background.php) or near the edge of the swath (rows 1–10 and 51–60) and data from days potentially influenced by large transient volcanic plumes were also excluded. Details on the data filtering can be found in McLinden et al. (2016). These daily amounts were averaged to produce annual mean gridded SO_2 columns in DU as found in McLinden et al. (2016).

Annual mean SO_2 total column amounts data from 2005 to 2016 were provided in a data file format not native to ArcGIS software used for the analysis of data and construction of maps. The SO_2 data were arranged as an array of latitude, longitude, and column amount. A Python script was created for converting data file types from an array to a table and a second script was written to import the CSV tables into ArcMap as temporary layer files and then convert the temporary files into a multi-point feature class. A second method of first converting the raw data files to an ASCII file and then importing directly into Arc

via an “ASCII to Raster file” tool is possible, but Python scripts would still be needed as each year of SO_2 data is a multi-point feature class containing 5.76 million points.

To aid in visualization, increase data manageability, and improve data manipulation process speeds, a Python code broke up the data into eight sections by latitude and longitude in the following way:

Section 1: LON -89.95 to -179.95, LAT 0.05 to 79.95
Section 2: LON -0.05 to -89.85, LAT 0.05 to 79.95
Section 3: LON 0.05 to 89.85, LAT 0.05 to 79.95
Section 4: LON 89.95 to 179.95, LAT 0.05 to 79.95
Section 5: LON -89.95 to -179.95, LAT -0.05 to -79.95
Section 6: LON -0.05 to -89.85, LAT -0.05 to -79.95
Section 7: LON 0.05 to 89.85, LAT -0.05 to -79.95
Section 8: LON 89.95 to 179.95, LAT -0.05 to -79.95

The original spatial resolution of the gridded OMI SO_2 column amounts was 0.1 decimal degrees while population data from the Gridded Population of the World (GPW v.4) dataset (see section 2.1.2) had a resolution of 0.00833 decimal degrees. Because population count was used, all efforts to preserve area were taken, including resampling the SO_2 column amount data to the cell size of the population data and using area preserving coordinate systems. Many common analysis processes in ArcGIS, including those performed in the analysis of this research, will automatically resample the data to the coarsest cell size using a “nearest neighbor” resampling technique if cell sizes of two data sets differ (ESRI, n.d. a). Therefore, a resampling was done as a preliminary step to prevent multiple resampling processes in the subsequent analysis that could introduce errors and diverge values from the original datasets.

An interpolation method was necessary to decrease the cell size of the SO_2 layer. The sectionalized datasets were first converted from point data to raster data using the ArcMap Point to Raster tool. The datasets were then resampled using nearest neighbor interpolation to reduce the cell size from 0.1 to 0.00833 decimal degrees. For continuous data, bilinear to cubic resampling is usually used; however, resampling with nearest neighbor was chosen over other interpolation methods to preserve the true values of each cell. As with all interpolation methods, some changes from the original data were expected. The result of the resampling process was two data sets of the same cell size slightly offset as they had different points of origin. A further tool, Snap, could have been utilized to eliminate the offset between the two datasets. However, the Snap tool would have resulted in a second resampling with the possibility of further deviation from original data and was not necessary for this analysis.

2.1.2 Population Data

The Gridded Population of the World, Version 4 (GPWv4) is an open-source dataset from NASA's Socioeconomic Data and Applications Center (SEDAC) of population count estimates for the years 2005, 2010, 2015, and 2020. The dataset product is an aggregation of census data from national statistics offices and other organizations and, when necessary, official population estimates. This data was derived using the areal-weighting method, eliminating the need for additional modeling, and allowing for analysis with other data, "without concern for endogeneity" (CIESIN, n.d.a).

The GPWv4 dataset estimates population every 5 years. However, for this work population data corresponding to the annual mean SO_2 column amount data were needed. Therefore, the population data was interpolated. The difference between GPWv4 data years was taken and divided into four sections and successively added to each previous year. The SEDAC website specifically said this process was not appropriate for this data as the GPW data is not a time series. Instead, all years of data are created from the same input data and a growth rate is applied to extrapolate target years (CIESIN, n.d.b). However, the lack of open-source population datasets when the analysis began necessitated the use and interpolation of this data. A similar procedure was performed on GPWv4 population data in Li et al. (2017), but the details of the interpolation were not published.

2.1.3 Planetary Boundary Layer Data

The PBL is the portion of the atmosphere that directly interacts with the surface of the planet. It is the upper bound of the OMI VCD sulfur dioxide data (Fioletov et al., 2016) and was needed in this methodology for the conversion of SO_2 VCDs in DU to concentrations in parts-per-billion. Along with raster cell size, it provides the volume estimates needed for those calculations. The PBL data was sourced from von Engeln and Teixeira (2013). The study uses European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim reanalysis on climatological records from 1990 to 2009 derived from several satellite sources and outlines a methodology for finding PBL height with respect to local elevation. In von Engeln and Teixeira (2013), several gradient-based parameters are proposed in their methodology, but the minimum gradient of relative humidity performed the best based on sea- and land-based radiosonde evaluation, and the data set provided for use in this analysis is derived using the relative humidity parameter.

The dataset was presented as a set of yearly averaged diurnal PBL heights in kilometers with a spatial resolution of 1.0 degree. The diurnal set consists of four files corresponding to universal time: 00UT, 06UT, 12UT, and 18UT. Each file was converted to a raster using the ArcMap Make NetCDF Raster Layer tool and oriented correctly using the Flip,

Shift, and Define Projection tools in ArcMap. From these four files, three PBL raster products were derived and used in the methodology. The PBL product at 12 UT was unaltered and used as is as an example of real modeled conditions. The two other PBL layers used were minima and maxima to define a PBL height range for each region of interest. The Cell Statistics tool in ArcMap was applied to produce these minimum and maximum raster values per cell from all four diurnal PBL rasters. Analysis involving PBL was triplicated for each PBL height raster condition: 12 UT, minimum, and maximum.

2.1.4 Global Emitters Data

Fioletov et al. (2016) compiled anthropogenic and volcanic point sources of SO_2 detected from OMI SO2 measurements collected since 2005. A catalogue of SO_2 emitters, including 76 passively degassing volcanic sources, resulted from their study, and additional information such as estimates of annual mean SO_2 emissions from each source in kilotons were also included. In cases of SO_2 sources in close proximity, they estimated and labeled multiple as a single source. A follow-up study by Carn et al. (2017) added 15 more volcanoes to the catalogue. These two studies were the source of air mass factor (AMF) data (section 2.3.1), location, and supplemental emissions data for each volcanic SO_2 emitter and were converted to GIS point feature classes and used throughout this analysis.

2.1.5 Political Boundaries Data

A global country boundaries shapefile was created by Liangcun Jiang (2016) from University of California Santa Barbara using data provided by DIVA-GIS. This layer was chosen for its high resolution and its non-restrictive open-source terms of use. It is presented in the GCS_WGS_1984 coordinate system with a datum of D-WGS_1984. The global shapefile was used by first selecting and exporting specific regional areas. It was then imported into geodatabases and projected into area-preserving specific coordinate systems.

2.2 Global Data Visualization

Sulfur dioxide column amount data in DU (uncorrected) were displayed in ArcMap for all years (2005-2016). Passive volcanic SO_2 emitters (Fioletov et al., 2016; Carn et al.,

2017) were overlaid on each data set to identify point sources of SO_2. Seven regional areas of high volcanicSO_2 emissions were located.

This study was not intended to be a comprehensive analysis of all passive emitters. Instead, the focus was on finding regions of varying geographic location, scale, and population. The criteria used to determine these regions were the presence of persistent passive SO_2 degassing, the absence of anthropogenic point sources of SO_2 near the volcanic emitters, and landmass population. These regions are summarized in Table 1.

Table 1: Names, locations, and volcanoes of each region of focus for this study.

Name	**Location**	**Relevant Volcanoes**
Hawaii	Hawaii, USA	Kilauea
Central America	Central American Volcanic Arc (Guatemala, El Salvador, Nicaragua, Costa Rica)	Santa Maria, Fuego, Pacaya, Santa Ana, San Miguel, San Cristobal, Telica, Masaya, Turrialba, Poas
Sicily or Italy	Sicily and Aeolian Islands, Italy	Etna and Stromboli
D.R. Congo	Democratic Republic of the Congo	Nyiragongo and Nyamuragira
Java	Sunda Arc; Java, Indonesia	Krakatau, Slamet, Merapi, Bromo, Semaru, Raung, Ijen
Japan	Japanese and Izu Islands	Tokachi, Asama, Miyake-jima, Sakura-jima, Satsuma-Iojima, Suwanose-jima
Vanuatu	Vanuatu Island Arc	Ambrym, Aoba, Yasur, Gaua

Seven regions were chosen for this study, but a few additional regions with significant passive volcanic SO_2 emissions could be analyzed in a continuation of this study. These areas include Papua New Guinea (omitted for its proximity to two other regions: Java and Vanuatu), Central Mexico (omitted for its proximity to several anthropogenic point emitters), Northern Mariana Islands (omitted because SO_2 emissions are not consistent and no occupied land mass falls within SO_2 plume footprint), Montserrat (omitted because of low population and minimal land mass in SO_2 plume footprint), Kuril Islands (omitted for proximity to Japan and low population density), and the northern volcanic zone of the Andes (omitted for its low population density). Twelve maps corresponding to each year of global SO_2 and population data were created. An example map can be found in the Results Section of this work and a comprehensive catalog of all maps can be found in the supplemental materials.

2.3 Regional Visualization

Data organization and preparation for each region was performed using the following steps:

1. A regional geodatabase was created.
2. An equal area projection was either found or created.
3. Polygon feature classes of country or regional outlines were created and projected.
4. All external data was clipped to the polygon feature class and projected while importing into the area geodatabase using a Python script. Regional visualization is for SO_2 concentration (ppb), corrected SO_2 VCDs (DU), and hazard maps.

2.3.1 Manual Division Workflow

Air mass factor (AMF) is defined as the ratio of slant column density (SCD) to vertical column density (VCD) for a specific gas measured using remote sensing techniques. The relationship between VCD and SCD is made more complex by the addition of atmospheric scattering and multiple reflecting surfaces, and hence the AMF is affected by factors such cloud cover, the angle between the retrieval instrument and the Sun, the surface albedo, the altitude of the source, and the specific gas being retrieved (McLinden et al., 2014). The PCA algorithm used to generate OMI SO_2 VCDs used in this analysis assumes a fixed air mass factor of 0.36, corresponding to summer conditions in the eastern USA. This will result in erroneous VCD values for regions that differ in elevation, latitude, or surface conditions due to variable AMFs. Fioletov et al. (2016) and Carn et al. (2017) calculated site-specific AMF values following the approach of McLinden et al. (2014). The new site-specific AMF values account for differences in the factors listed above and were used in correction of the SO_2 VCDs. The threshold level of the SO_2 VCDs used by Fioletov et al. (2016) to identify SO_2 sources was 0.1 DU (Fioletov et al., 2016). To further eliminate background noise values, a threshold value of 0.15 DU SO_2 was assumed in this analysis.

To correct AMFs for regional SO_2 column densities, a method of determining the influence of each volcano is necessary because AMF values are associated with individual volcanic emitters (Fioletov et al., 2016; Carn et al., 2017). In most cases, a single volcanic influence can be seen as a discreet “footprint” of the SO_2 VCD (i.e., the volcanic SO_2 plume). However, for some regions with multiple SO_2 sources there is difficulty determining the SO_2 influence from a single point. In some cases, two adjacent volcanoes have overlapping SO_2 “footprints” or non-volcanic (anthropogenic) SO_2 “footprints” are also present. In the first case, if SO_2 footprints cannot be delineated, the two are taken as one point and the AMF of the volcano with the highest elevation is used.

In the cases of industrial emitters, SO_2 from the two sources can't be decoupled without further data (e.g., vertical distribution of sulfur dioxide), if at all. The delineation process is outlined in Figure 1. The following steps detail this process.

1. A polygon layer was created for each year of SO_2 data by clipping to the general area and deleting political boundaries, if necessary, as in the case of Central America.

2. With SO_2 VCDs displayed in the background, the polygon layer was divided based on SO_2 "footprints" around emitters. This process was done by editing the layers using ArcMap's Editor: Cut Polygon tool manually. Supplemental annual mean SO_2 emissions in kilotons from Carn et. al. (2017) were used to aid in the delineation process. The polygons corresponding to each year were slightly different based upon the SO_2 "footprint" pattern and size. The output of this process was a divided regional polygon with sections of different volcanic emitter influence for each year.

3. SO_2 VCD data were projected and clipped to the regional outline using the Project and Clip tools within an ArcPy script. This data was displayed to show differences in SO_2 to make the process of delineation easier.

4. ArcMap's Contour tool was run on the regional SO_2 raster with an interval of 0.15 DU as the threshold amount and a setting of contour shell. This resulted in several polygons with upper ranges increasing incrementally by the interval amount while the lower bound remained at the minimum amount. The polygon corresponding to the minimum to 0.15 DU range was then selected and exported as an independent layer.

5. The output results from steps 2 and 4 were then updated with the correct AMF by creating a new field in their attribute tables and then input into the Union tool. For the polygon corresponding to the area below the 0.15 DU threshold, the standard AMF of 0.36 was used. For all other areas, the AMFs from Fioletov et al. (2016) and Carn et al. (2017) were used. The Union tool combines two or more vector layers into a single vector layer. It was used to combine both geographic footprint and attribute data of all input data. In this case, the output of the Union tool was a finalized area of interest (AOI) polygon divided both by volcanic influence and areas below the threshold (0.15 DU), with the correct AMF associated. These finalized areas of interest were then used to create two main products: maps of corrected SO_2 VCD in DU and maps of SO_2 concentration in ppb.

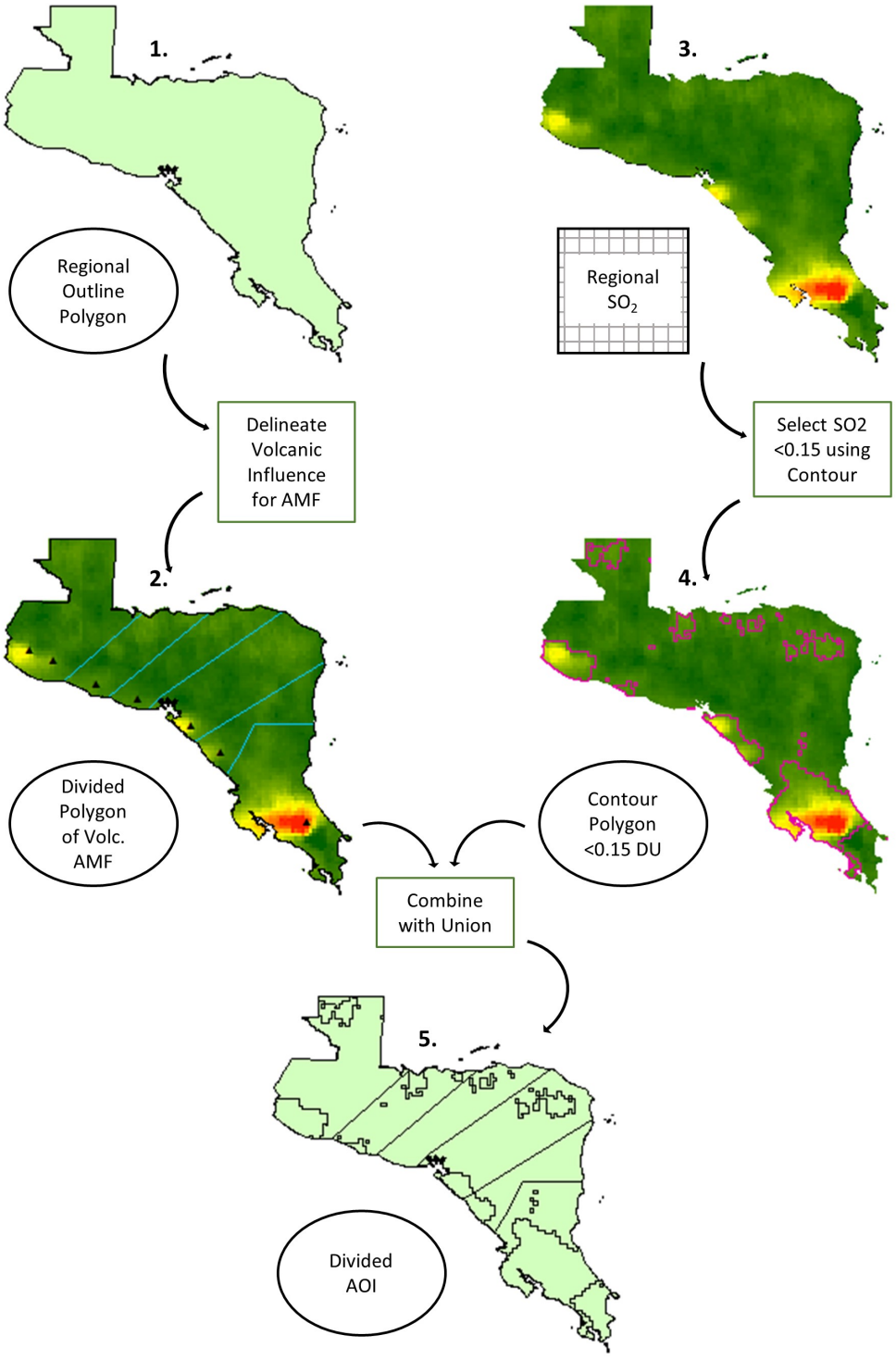

Figure 1: Example of the Manual Division Workflow on Central America. Using regional SO_2 and a regional polygon, appropriate AMFs are assigned to the correct areas. Areas below the threshold are assigned a default AMF. Steps are detailed in Section 2.3.1.

2.3.2 Parts-Per-Billion Conversion Workflow

The processes of correcting the AMF of the SO_2 in VCDs and creating ppb rasters were performed by the same Python script called Correcting Dobson units & Estimating Parts-per-Billion and are outlined in the same flowchart shown in Figure 2. This script can be found in Appendix A1. However, these processes could easily be performed separately. The steps involved are outlined below.

1. The divided AOI polygon (see Fig. 1) was the input into the Polygon to Raster conversion tool. The output of this tool was a raster for each year that displays AMF data created during the Manual Division Workflow. The Polygon to Raster conversion was necessary to enable the AMF values to take part in a Raster calculation as part of a later step.

2. To correct for the previously applied AMF and create corrected SO_2 VCD rasters, the following calculation in Equation 1 was performed, where the inputs were the AMF raster created in Step 1 and the SO_2 raster.

[1] $$SO_{2,Cr} = \frac{SO_{2,UC}}{AMF_{sd}} \cdot AMF$$

Where:

AMF is the raster created in Step 1.
$SO_{2,\,UC}$ is the uncorrected raster.
$SO_{2,\,Cr}$ is the newly created corrected SO_2 raster in DU.
AMF_{sd} is the previously applied AMF, 0.36.

3. The SO_2 molecule calculation required the corrected SO_2 VCD (DU) as an input. Total SO_2 molecules per column was calculated by multiplying the corrected SO_2 raster by the cell size area and the conversion factor. The cell size area was slightly different for each region due to the difference in equal area projections. The conversion factor was 2.687 x 10^{20} molecules m^{-2}. This is shown in Equation 2 below. The total SO_2 molecules were used as an input to the parts-per-billion calculation in Step 6.

[2] $$SO_{2,\,molttl} = SO_{2,Cr} * cell\ size * F_{DU}$$

Where:

$SO_{2,molttl}$ is a raster of the total SO_2 molecules.
Cell size is an area in meters.
$SO_{2,Cr}$ is the raster of corrected SO_2 DU.
F_{DU} is the conversion factor, 2.687x10^{20} molecules m^{-2}.

The parts-per-billion calculation required both the number SO_2 molecules and the total number of molecules in an atmospheric column within the area of the cell size of the raster and the PBL height. The number of SO_2 molecules was converted directly from the corrected SO_2 raster data, shown above. The total number of molecules was derived from the molecular form of the Ideal Gas Law, shown in Equation 3,

[3] $$n = \frac{PV}{k_B T}$$

Where:

n is the number of molecules.
P is the pressure in Pascals at median height of column (½ PBL altitude raster).
V is volume in m^3 or the PBL raster x cell size area.
k_B is the Boltzmann constant, 1.38×10^{-23} J K^{-1}.
T is temperature in Kelvin at median height of column (½ PBL altitude raster).

The altitude, temperature, and pressure rasters used in the Equation 3 were all created in the steps below.

4. There were three PBL height rasters used in this analysis: the minimum PBL, the maximum PBL, and the 12UT PBL. Each of these rasters was clipped and projected to the area of interest. The PBL rasters were looped in the Python script to produce temperature, pressure, and altitude rasters for each region and PBL condition. The altitude rasters were created by halving the PBL raster. To determine the temperature rasters, a modification of the Molecular-Scale Temperature equation under 86 km height was used, shown below in Equation 4.

[4] $$T_M = T\ \frac{M_0}{M}$$

Where:

T is the kinetic temperature in Kelvins.
M is the mean molecular weight of air.
M_0 is the sea-level value of M.

Within the range of altitude in the PBL, the molecular scale temperature in Equation 4 becomes a function of height, and the equation becomes linear with the form shown in Equation 5:

[5] $$T_M = T_{M,b} + L_{M,b} * (H - H_b)$$

Where:

$T_{M,b}$ is 288.15 K at altitudes below 11 km.
$L_{M,b}$ is the molecular-scale temperature gradient, also known as the lapse rate. Below 11 km, the lapse rate is -0.0065 K/m.

H is the altitude.
H_b is the geopotential height.

Below 11 km, H and H_b are the same; for the analysis, the altitude raster derived from the PBL rasters was used. Simplified, the Equation 5 becomes Equation 6, which was used to create the temperature rasters:

[6] $$T_M = 288.15\ K - (0.0065 \frac{K}{m} * H)$$

Where:

T_M is the molecular-scale temperature in Kelvin
H is the height, in this case the altitude raster, in meters

The pressure raster creation used the temperature raster as an input. Within the troposphere (below 11 km), a barometric Equation 8 is used to find P, pressure at ½ altitude. Equation 8 is modified from the following Equation 7.

[7] $$P = P_b * \left[\frac{T_{M,b}}{T_{M,b}+L_{M,b}*(H-H_b)}\right]^{\frac{g'_0 * M_0}{R^* * L_{M,b}}}$$

Where:

g'_0 is the dimensional constant to relate standard geopotential height to geometric height. It is the same as g'_0, but with different units. In this case, $g'_0 = 9.80665$ m/(s^2m).
M_0 is a set of values of molecular weights based upon carbon-12 isotope scale. In this case the molar mass of air, $M_0 = 0.0289644$ kg/mol.
$R^{\wedge *}$ is the gas constant. In this case it is 8.31432 * 103 N*m/ (kmol*K).
$L_{M,b}$ is the molecular-scale temperature gradient, or lapse rate. Below 11 km it is -6.5 K/km.
$T_{M,b}$ is the temperature at altitude. Below 11 km, this is 288.15 K.
P_b is the pressure at conditions of b. In this case, it is 101,325.0 N/m^2 or 101,325 Pa.
H_b is a set of geopotential height values, below 11 km it is 0 km.

Substituting the denominator for the molecular scale temperature, or the temperature raster, the equation was simplified further to:

[8] $$P = 101325\ Pa * \left[\frac{288.15\ K}{T_M}\right]^{-5.2525}$$

Equation 8 was then used to create pressure rasters for each PBL raster. These equations assume the atmosphere is an ideal gas and that temperature variation is linear according to the atmospheric lapse rate, $L_{M,b}$.

5. To calculate total molecules in an air column, corresponding pressure, temperature and PBL rasters were needed, along with the cell size of the corrected (and projected and clipped) SO_2 raster. The output total molecule raster was not saved due to limits of raster data type being unable to store very large numbers. Instead, the values were carried through the process to calculate SO_2 concentration in parts-per-billion. They were verified with a simple print statement easily converted into a comma-separated value (CSV) table. Equation 9 was modified from the Ideal Gas Law, Equation 3, and shows the process of creating the total molecule raster.

[9] $$Molecules_{Total} = \frac{P*cell\ size*PBL}{k_B*T}$$

Where:

$Molecules_{Total}$ is the total molecules in a column (unsaved raster).
P is a raster of pressure in Pa at ½ PBL.
PBL is the raster of the PBL height.
cell size is the area of each cell in the corrected sulfur dioxide raster.
k_B is the Boltzmann constant, 1.38 x 10^{-23} J K^{-1}.
T is a raster of temperature in K at ½ PBL.

6. To calculate parts-per-billion of SO_2, both the total SO_2 molecule rasters and total molecule in column rasters were used in Equation 10.

[10] $$SO_2\ PPB = \left[\frac{SO_{2,total}*F_{PPB}}{Molecules_{\ total} - SO_{2,total}}\right]$$

Where:

$SO_{2,\ total}$ is the total sulfur dioxide molecule raster from Step 3.
F_{PPB} is the Conversion factor for parts-per-billion, 1 x 10^9.
$Molecules_{total}$ is the total molecule in a column raster from Step 5.

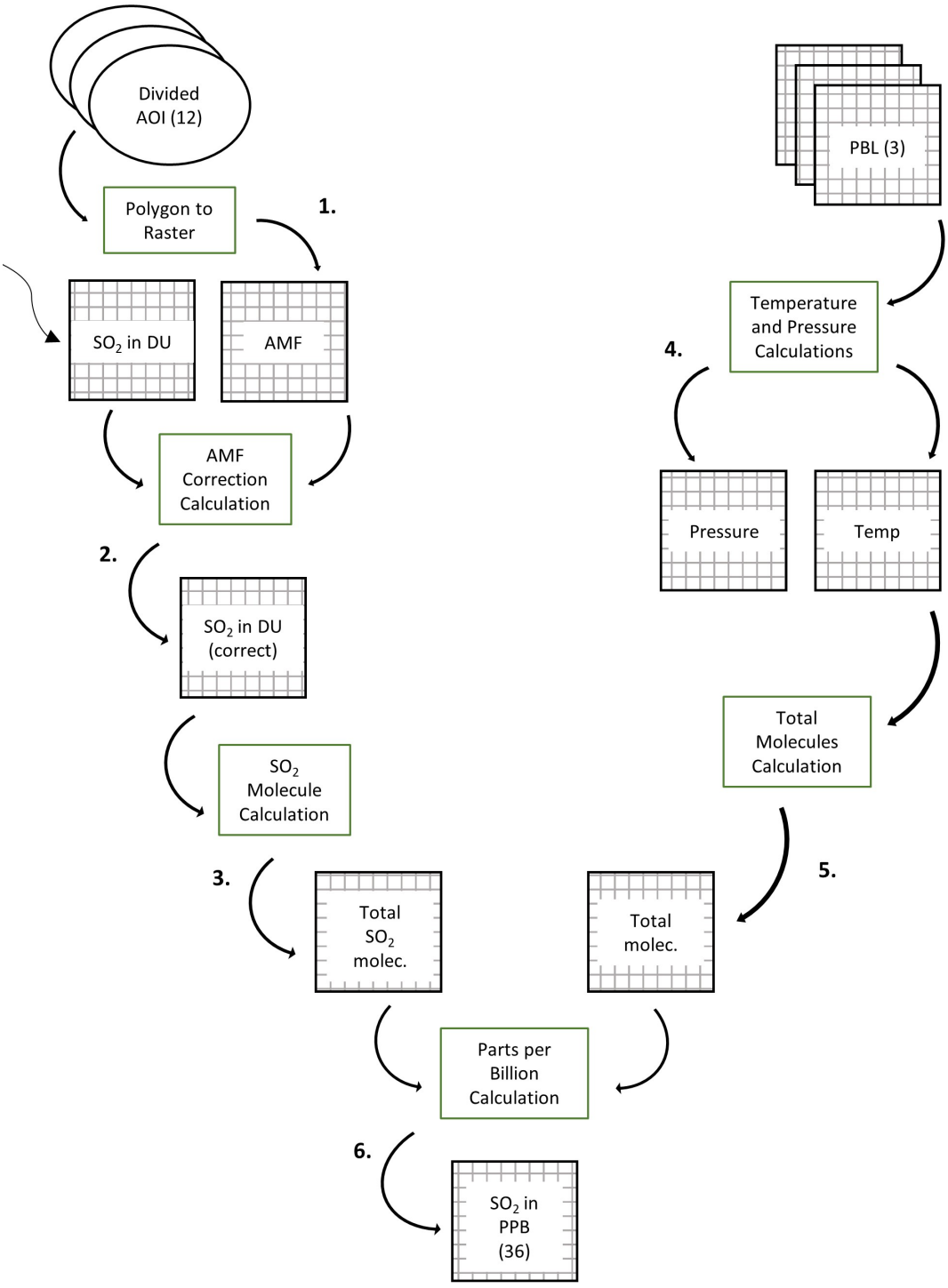

Figure 2: Workflow for ppb conversion and SO_2 VCD AMF correction. The outputs of this workflow were corrected AMF SO_2 VCD rasters (in DU) and SO_2 concentration (in ppb) rasters. Gridded boxes represent raster data, rounded ovals represent vector data, and green boxes indicate an action done to the data.

2.4 Hazard and Exposure Visualization Workflow

Adverse health effects due to volcanogenic SO_2 can be difficult to quantify. In this research, it is assumed that SO_2 coexists with other volcanic and possibly anthropogenic gases, making its health effects difficult to parse from other gases in the atmosphere. Initially, this study aimed to quantify SO_2 by health effect. However, SO_2 affects individuals differently depending on their sensitivity and duration of exposure, and morbidity studies show there are few discrete amounts at which health effects are ubiquitous for long-term exposure; these health effects can be chronic and experienced on a varying scale (Longo et al., 2008).

Therefore, for the purpose of this study, national ambient air quality standards (NAAQS) for each region were used in conjunction with World Health Organization air quality guidelines (AQGs) to quantify potential exposure risk. The WHO 24-hour AQG for SO_2 is 7.7 ppb (World Health Organization [WHO], 2018). The WHO AQGs are based on epidemiological studies that consider sensitive subgroups of the population, while NAAQS are often a compromise between what is feasible and health guidelines (Vahlsing & Smith, 2012). The standards used for each region are found in Table 2 below. The WHO AQG and NAAQS for each region are 24-hour averages of non-exceedance standards. They have been converted to parts-per-billion using a conversion of 1 ppb = 2.62 $\mu g/m^3$ if necessary.

Several countries did not have SO_2 NAAQS accessible. The Democratic Republic of the Congo, Vanuatu, Guatemala, and Honduras did not have accessibly published standards for sulfur dioxide (Green & Sánchez, 2013). Guatemala and Honduras were part of the Central American regional area of interest, so the NAAQS from the surrounding countries were used. These countries, Costa Rica, El Salvador, and Nicaragua, all agreed on a standard of 140 ppb (Vahlsing & Smith, 2012; Green & Sánchez, 2013). Italy followed the European Union's 24-hour AAQS of 48 ppb (European Commission, n.d.; Vahlsing & Smith, 2012). Japan has a 24-hour NAAQS of 48 ppb (Ministry of the Environment Government of Japan, n.d.). Indonesia's NAAQS for SO_2 is 140 ppb (Clean Air Initiative for Asian Cities [CAI-Asia], 2010; Vahlsing & Smith, 2012). The United States does not have a current 24-hour NAAQS for SO_2, with the previous 24-hour NAAQS being revoked in 2010. The previously revoked standard of 140 ppb was used in lieu of a current 24-hour standard (United States Environmental Protection Agency [U.S. EPA], 2017; Vahlsing & Smith, 2012) as it is the same as the standard for Hawaii (Pressler & Ige, 2016). A standard of 140 ppb was used for both the Democratic Republic of the Congo and Vanuatu for ease of comparison with other areas of interest.

The hazard maps created in this study display whether these countries or regions are potentially meeting their own standards and the WHO's guideline under each PBL condition used in this analysis. The maps indirectly show hazard effects through the

country and global standards. If alternative guidelines are created or if NAAQS are updated, the methodology for the creation of the hazard map is unchanged and can be implemented.

Table 2: National ambient air quality standards (NAAQS) and World Health Organizations Air Quality Guidelines (WHO AQGs) used in this study. These are 24-hour limits. Values in parentheses indicate yearly standards. Asterisks and superscripts indicate those standards are not to be exceeded that number of times per year. The mean of the standards is 99.5. All standards have been converted into ppb from $\mu g/m^3$.

Country	**Standard in ppb**	**Country**	**Standard in ppb**
Mean	99.5	Indonesia (Java)	140 (23)
Costa Rica	140*[1] (30)	Japan	40
DR Congo	none	Nicaragua	140 (30)
El Salvador	140	US (Hawaii)	140*[1] (30)
EU	48*[3] (8) (1HR-134)	Vanuatu	none
Guatemala/Honduras	none	WHO	7.7

The hazard maps display the WHO's AQG in orange hatch marks to denote the areas where SO_2 exceeds 7.7 ppb. The NAAQS values for all areas are higher than the WHO AQG and are therefore presumed to be more hazardous. These are shown in red hatch marks. Note that not all maps show the NAAQS; some regions' maxima never exceed 140 ppb, especially when the maximum PBL condition is used, which results in higher column volume and lower SO_2 concentrations within the column.

2.4.1 Methods for Estimating Population Count Exposure

Within ArcGIS there are several ways to determine the population potentially exposed to variable concentrations of SO_2. Essentially, the process is to find areas that fall within a range of SO_2 concentration (the standards listed above) and use these as a "cookie cutter" on the population raster to determine the amount of people that fall within these ranges. Two methods are discussed: the Extract Tools method and the Contour/Clip method. The Extract Tools method uses two ArcMap tools: Extract by Attributes and Extract by Mask. Extract by Attributes creates a new raster from a logical query, in this case cells where SO_2 concentration is between 7.7 ppb and the country/regional standard, and a second selection over the country/regional standard, where applicable. These areas are then used as a mask for the second tool, Extract by Mask, which extracts the cells of a target raster (population) that are covered by the mask (SO_2 ranges).

The second method uses different ArcMap tools: Contour (shell up) and Raster Clip. The Contour tool creates a ringed feature class based on a base level and an interval, similar to the contours of a topographic map. The contours of interest (WHO and country/regional standards) are then selected, and all other contours are deleted. In the cases where the maximum SO_2 concentration exceeds the country/regional standard and the country/regional standard contour is created, the WHO standard contour is erased from the country/regional standard using the Erase tool. This ensures that a population isn't counted in both the WHO standard area and the country/regional standard area. Next these contour areas are used to clip out areas from the corresponding population raster using the Raster Clip tool. This results in separate population count rasters with areal coverage corresponding to the WHO standard hazard zone and the country/regional hazard zone.

Both of these methods for isolating population based on SO_2 concentrations work visually. However, they differ statistically. The Extract method causes the rasters to resample during both the Extract by Attribute and Extract by Mask tools, which can alter the statistics of the rasters. The Contour/Clip method does not resample rasters if the option not to maintain the clipping extent is enabled during the Clip tool. Taking the WHO standard of 7.7 ppb as an example, when using the Contour/Clip method, the contour line is interpolated to the value of 7.7. A cell will be included if the center of the cell falls within the contour bounds. This differs from the Extract method, where a cell is included only if its value is above 7.7. Because the Contour/Clip method does not resample the data like the Extract method does, it was chosen to create hazard maps and to perform the temporal and statistical analysis in the next section. The Python scripts used to execute this process are called Population and Hazard Analysis 1: Contouring Standards and Population and Hazard Analysis 2: Clipping Population and can be found in Appendices A.2 and A.3, respectively.

2.5 Temporal Analysis

2.5.1 Statistics Collection

After determining the areal coverage of the hazard zones, the next step is to find the amount of people exposed to SO_2 at or above the AQG and NAAQS. This was done by summing the cell values of the population rasters that fall within those ranges. The mean and cell count were also found to verify the sum statistic. This is shown in the third column of Table 3. These sums were found for every hazard map created.

Each of these areas of interest were chosen because of their relationship with population, either high or low. It follows that since anthropogenic SO_2 cannot be parsed from the volcanogenic load, it is unclear how much of the SO_2 load is coming from anthropogenic

sources like power plants and metal smelters. To address this, a population-weighted average SO_2 column amount was used. The equation to determine population-weighted SO_2 column amount was modified from one used in Li et al. (2017). The calculation used is shown by Equation 11.

[11] $$\Omega_W = \frac{\sum P_i \Omega_i}{\sum P_i}$$

Where:

Ω_W is the population-weighted SO_2 column amount in either DU or concentrations in ppb.
P_i is population count.
Ω_i is SO_2 VCD (DU) or concentration (ppb).

The population-weighted average captures a combined effect of changes in population and SO_2 and was used as an index of potential severity of exposure and its temporal variation. A script called Population-weighted Averages was used to determine these values and can be found in Appendix A.4. To verify the results of the Python script, the numerator and denominator were also collected, as well as the total sum of the SO_2 in either DU or ppb, along with average SO_2 in either ppb or DU. For the SO_2 VCD data, the population-weighted average and average SO_2 column amount were collected for every year. For parts-per-billion data, the average and population-weighted SO_2 concentration were collected for every year and PBL used. A graphic example of these statistics can be found in the Results Section as well as in Appendix C.

Table 3: Statistics gathered on raster products. Three main product types are listed in the header and the statistics calculated are listed in the table.

Dobson unit Statistics	**Parts Per Billion Statistics**	**Population Hazard Analysis Statistics**
Average and population weighted average SO_2 column amount	Average and population weighted average SO_2 concentration	Population sum per hazard zone
Sum of product of SO_2 and DU column amounts	Sum of product of SO_2 and ppb concentration	Population mean per hazard zone
Sum of population (total regional)	Sum of population (total regional)	Population count per hazard zone
Sum of DU (total regional)	Sum of ppb (total regional)	---

One of the difficulties of working with rasters in ArcGIS is the lack of a straightforward way to retrieve statistics. Similarly, ArcPy does not have a function to directly call the sum as an object. Instead, the data must first be defined as an array, and then the Python numpy package function numpy.sum can be used. Alternatively, the product of the count and the mean can be used. However, because ArcPy reads NoData values as either 0 or -9999, in this case 0, the array must be flattened to find the true mean without the zeroes (note: there should be no naturally occurring zeroes as the Contour/Clip step eliminated any values under 7.7, and, in the case of finding the sum using Dobson units, any values below zero are to be considered background noise.). After flattening the array, the true mean and count can be found and verified against the array sum. An alternative method is to manually visualize the raster in ArcMap and find the sum via Properties; Symbology; Classified; Classify. The sum will be listed under the Classify tab. This method is not feasible for finding the sum on many rasters. After consideration, all statistics were found using a Python script that utilizes a flattened array to produce tables of statistics.

2.6 Map Production Processes

The final products of this research are a series of hazard maps for selected volcanic regions and a nascent methodology for producing the data presented in the maps. These series include a global SO_2 map for each year with areas of interests highlighted (12 maps), a series of maps for each region displaying SO_2 VCD in DU (84), a series of maps displaying SO_2 in ppb for each region and each PBL condition (252), a series of hazard maps for each region and each PBL condition displaying population overlaid with ppb hazard zones (252), resulting in a total of 600 maps. Within each map a certain number of elements are necessary to complete the cartographic product. Careful detail on the placement of elements, for example legends and scale bars, require a manual review process. However due to the large volume of maps, each map cannot be created individually, similar to the data processing and statistical analysis; the map production must be automated in Python. A script called Map Automation was created that utilized several map tools and can be found in Appendix A.5.

2.6.1 Classification and Display of Data

All regional maps containing SO_2 data use a geometric interval classification to visualize this data. This includes the VCD maps, the concentration maps, and the hazard maps. Both the VCD maps, and the concentration maps display SO_2 in seven classes, while the hazard maps display population count in five classes. Geometric interval classification was used to emphasize high SO_2 values. As stated in the ArcGIS Help Guide on Data Classification methods (ESRI, n.d. b),

> This algorithm was specifically designed to accommodate continuous data. It is a compromise between the equal interval, natural breaks (Jenks), and quantile methods. It creates a balance between highlighting changes in the middle values and the extreme values, thereby producing a result that is visually appealing and cartographically comprehensive.

The world map shows SO_2 VCD in DU displayed in five classes using the natural breaks (Jenks) classification method for ease of visibility. The natural breaks method was chosen over geometric classification because geometric classification was visually awkward for these maps because it classified most data into the two highest classes while natural breaks was optimal for comparison of different sources, both volcanic and anthropogenic.

2.6.2 Equal Area Projections

With the exception of the global SO_2 maps, all other data is projected into an equal area projection. Both the SO_2 data and the population count data are area dependent, meaning if the area of the original data is altered, the data may become erroneous. Many national grid or projections are not designed to preserve area, and many of the regions in this analysis do not have national projections available. Therefore, it was necessary to create custom projections for most of these regions that preserve area while minimizing distortion. Table 4 below displays the details of the regional map projections.

For Central America, the area of interest spans several countries from Guatemala to Costa Rica with both a significant North to South span and a moderate East to West span. An equal area projection is not readily available for this area, so a custom projection was created. The template projection was South America Albers Equal Area Conic projection, WKID: 102033. The new Central Meridian and Standard Parallels are shown in Table 4. The Democratic Republic of the Congo (DRC) has a square geographic orientation. However, the SO_2 footprint emanates from Nyiragongo and Nyamuragira along the Eastern border and covers the country in an East to West strip. A cylindrical equal area projection was used for this orientation as it reduces distortion along a straight narrow area. The central meridian and standard parallel were modified from WKID: 54034.

For mid-latitudes that do not exceed 30 degrees N to S with moderate extent (not continent, not city), such as Italy, Albers Equal area conic is typically used. Area is preserved while shape, angle and length become distorted away from standard parallels. For the area of Hawaii, an equal area projection that was suitable already existed. It is the Hawaii Albers Equal Area Conic, WKID: 102007. Due to Japan's square orientation running NE to SW, a Lambert azimuthal equal area projection was chosen to preserve the area, with shape distortion increasing radially away from intersection of meridian and origin. The nearest appropriate projection found was WKID: 102020, South Pole Lambert Azimuthal Equal Area, which was modified to suit Japan.

Java also lacked a ready-made equal area projection. An Albers Conic projection was chosen because of the East to West extent of the island and its tabular shape, the projection would display minimal distortion between the two standard parallels. This projection was modified from WKID: 102028. Vanuatu has a strong North to South linear orientation. For this shape, a cylindrical projection was chosen to minimize distortion. The World Cylindrical Equal Area projection, WKID: 54034, was modified to fit the area and range for Vanuatu. Two main sources were used in researching appropriate projections. The first being the Projection Wizard (Šavrič et al., 2016) and the second being the Esri Help Documentation on projections (n.d. c).

Table 4: This table details the coordinate systems and projection information on the regional areas of interest. Custom projections were created for all regions except Hawaii.

Region	**Coordinate system Name**	**Central Meridian and Std. Parallels**	**Geographic Coordinate System**	**Datum & Spheroid**	**Notes**
Central American Volcanic Arc	Central America Albers Equal Area Conic	-87.5: 11.4, 14.2	GCS South American 1969	South American 1969, GRS 1967 Truncated	Modified to fit region. Datum transformation needed.
DRC	Cylindrical Equal Area	29.2: -1.474	GCS WGS 1984	WGS 1984	Modified to fit region.
Hawaii	Hawaii Albers Equal Area Conic	-157: 8.0, 18.0	GCS North American 1983	North American 1983, GRS 1980	
Italy	Europe Albers Equal Area Conic	15.0: 36.5, 38.5	GCS European 1950	European 1950, International 1924	Modified to fit region.
Japan	South Pole Lambert Azimuthal Equal Area	136.75: NA	GCS WGS 1984	WGS 1984	Modified to fit region.
Java	Asia South Albers Equal Area Conic	115.0: 2.0, -7.0	GCS WGS 1984	WGS 1984	Modified to fit region
Vanuatu	World Cylindrical Equal Area	168.3: -17.25	GCS WGS 1984	WGS 1984	Modified to fit region
World	GCS WGS 1984	0.0	GCS WGS 1984	WGS 1984	Standard GCS

3 Results

The results of this study are two-fold. A methodology for the analysis of satellite-derived volcanogenic SO_2 data and the creation of four map products: global SO_2 VCD, regional SO_2 DU, regional SO_2 concentrations, and regional population hazard maps. Also included in this section is a series of statistical graphs and a table of population count exposure to SO_2 at or above the WHO AQG and regional ambient air quality standards (AAQS).

3.1 ArcPy Scripts

The primary product of the methodology for the analysis and map creation is a series of Python scripts formatted in ArcPy that were utilized throughout the analysis. The relevant Python scripts can be found in Appendix A, but the following sections provide a description and limitations of all scripts used in the analysis.

3.1.1 Data Import 1: Raw Data to CSV Script

This script converts a folder of raw .dat arrays into 3-column .csv tables. This script is to be used before and in conjunction with dBase Table to Feature Class script. Original data files must have specific lat lon SO_2 array structure. This script adds headers to the csv columns for lat, lon, and SO2_DU. Therefore, it is structure specific, and errors can occur if lat and lon are confused. File names are copied throughout the processes; therefore, file names must contain the words "mean" or "count".

3.1.2 Data Import 2: dBase Table to Feature Class

This script is to be used in conjunction with the Raw Data to CSV tool. CSVs must have a three column, headed format. File names are also copied throughout the processes; therefore, file names must contain the words "mean" or "count". Due to the large extent of data, WGS 1984 coordinate system was chosen, and all data will be adopted into WGS 1984 when converted to a feature class, unless the coordinate system is modified. This tool takes a folder of .csv files created from Raw Data to CSV and performs a series of conversions. First, it converts the CSVs into dBase tables. It then imports the dBase

tables into a working file geodatabase. Next, it creates an XY Event Layer from the dBases in the File Geodatabase. It then uses Copy Features to create permanent feature classes from the XY Event Layers.

3.1.3 Data Management: Sectionalization and Resampling

Point data of SO_2 VCD were broken up into eight sections for data manageability. Without this step, ArcMap cannot render all data points at once. This script is meant to be a fast alternative to manual selection and export. This code uses the Select by Attributes tool to select each portion of the total point data based on their lat and lon and the Point to Raster tool to export the selected points directly to raster data. The raster is then resampled using nearest neighbor interpolation to adjust cell size to match destination cell size. The code will loop through all years of data in a defined folder or geodatabase.

As written, this code only applies to global point data; i.e., the input data is a global dataset of points. A temporary layer file used in the Select by Attributes tool must be defined. Also, the location and name of output sectionalized point feature classes must be defined.

3.1.4 Data Organization: Clip and Project into GDB

This script iterates through all years of global SO_2 VCD in raster form, clips each SO_2 raster to a country or regional area feature class, then projects the clipped layer into the correct coordinate system, and finally imports it into a geodatabase. It repeats this process for planetary boundary layer raster data as well as population raster data and various other datasets. This code is necessary for data organization and automates the importation of raster into regional geodatabases.

Inputs or information needed for this script include the file locations of global SO_2 rasters from the previous step, planetary boundary layer rasters, population rasters, a country or regional outline, a coordinate system, and the file location of the regional geodatabase, which should be created beforehand. Otherwise, a simple addition to the code can create a geodatabase. The output is an organized geodatabase with clipped and projected SO_2 VCD, population, and planetary boundary layer raster data. The general template of the script allows for this process to be performed on any raster data in need of clipping, projecting, and importing into a geodatabase.

3.1.5 Correcting Dobson units & Estimating Parts-per-Billion

This script, found in Appendix A.1, closely follows the workflow outlined in Fig. 2. It iterates through PBL rasters to create intermediate products of temperature, altitude, and pressure rasters. These are then inputs into the calculation finding total molecules in an air column, which is used in the parts-per-billion calculation. This script also converts the area of interest polygons from the Manual Division Workflow to AMF rasters. These rasters are then input into a Dobson unit correction equation, the product of which is a raster of corrected SO_2 VCD (in DU). Lastly, the corrected SO_2 VCD raster is converted to parts-per-billion using a conversion factor equation.

The two primary outputs of this script are rasters of corrected SO_2 VCD for each region and year and rasters of estimated SO_2 concentrations for each region of interest, each year, and each PBL condition. The secondary products are rasters of AMF, temperature, pressure, and altitude for each year and region of interest.

It is not recommended that this script be altered in formatting to fit different data types. It is created solely for the purpose mentioned above. Furthermore, it uses the len and sort functions on the area of interest polygons and the SO_2 rasters to determine if there are equal numbers of each and ensures each year of polygon data corresponds to the matching year of SO_2 data.

3.1.6 Population and Hazard Analysis 1: Contouring Standards

This script, found in Appendix A.2, creates polygons using the Contour tool configured at the WHO AQG of 7.7 ppb and the national ambient air quality standards (NAAQS). If SO_2 in parts-per-billion exceeds the NAAQS of the region, then the WHO AQG polygon is erased where the regional NAAQS polygons overlaps to eliminate counting the population twice. The inputs include the parts-per-billion SO_2 raster data. The outputs are polygon feature classes covering areas of the region exceeding both WHO AQGs and NAAQS with no overlap.

This script utilizes the Contour tool, which creates a series of features based on a given interval. Therefore, the features that do not correspond to the NAAQS or AQG must be deleted. Note that the Erase tool that clips out the NAAQS polygon from the AQG polygon is only available with an advanced license for ArcMap.

3.1.7 Population and Hazard Analysis 2: Clipping Population

The contour polygons created in Part 1 are used to clip the population rasters resulting in population rasters of areas corresponding to WHO AQGs and NAAQS. A for loop is created that iterates through each of the three AQG or NAAQS contour polygon that differ by planetary boundary layer condition for each single yearly population raster and clips the population raster to each of these standards contour polygon. This is done for both the WHO AQG and the regional NAAQS. The inputs are the workspace, the AQG and NAAPS, the contour polygons from Step 1, and the population rasters. This script can be found in Appendix A.3.

3.1.8 Obtaining Statistics

As described in the Methodology, there is no straightforward way to obtain basic statistics of rasters in ArcMap or ArcPy, especially rasters with irregular or non-continuous coverage, which can lead to false zeroes or NoData figures. This script uses a tool and two Python functions to determine the statistics given in Table 3. The Raster to NumPy Array tool converts the raster data into a number array where some statistics can be called using the sum function. However, the array will still contain default NoData values which can skew most statistics with false zeroes. Therefore, the flatten function is needed. Flatten will convert the array into a list of numbers. From this list, all non-zero numbers can be isolated, eliminated, and statistics can be found using the sum, mean, and shape functions. The shape function returns the cell count in this case. The statistics are then arranged using a simple print statement resulting in a block of text that is comma separated.

This code is easily customizable for inputting alternative raster data, finding additional statistics, or returning statistics in a different format. One limitation on this script is that raster data inputs must contain no natural zeroes.

3.1.9 Population-weighted Averages

This script, found in Appendix A.4, uses similar code found in the Obtaining Statistics script to determine sums of parts-per-billion and Dobson unit rasters. It uses these sums Equation 11 (Section 2.5.1).

For the population-weighted concentrations, the script loops through the SO_2 concentration rasters in a 3:1 loop with the population rasters by year as there are three

ppb rasters per year of data. The VCD rasters have a 1:1 ratio with the population rasters, so a different loop is used to determine the equation inputs. All rasters are multiplied by a Boolean layer to eliminate cells below the SO_2 VCD threshold of 0.15 DU. For verification, the population-weighted average, the unaltered average, the numerator, the denominator, and then the sum of the SO_2 raster, either in ppb or DU, are printed in a comma separated print statement.

This code should only be used and altered to find the above equation and its parts. To determine statistics of alternative rasters, the Obtaining Statistics code should be altered and used. This code has two parts that perform similar equations but differ by the for loops used. The first uses a 3:1 ratio and therefore the inputs must have a 3:1 ratio. The second has a 1:1 ratio and may be more suitable to reconfigure for other data if necessary.

3.1.10 Map Automation

Three scripts were written for each main map product. The script shown in Appendix A.5 is the most complex, involving the use of population rasters within each NAAQS and WHO AQG region. Not all regions exceeded their NAAQS. In these cases, NAAQS rasters were not created. If NAAQS rasters exist, then both the NAAQS layer and WHO AQG layers are displayed. If the NAAQS raster doesn't exist for the region, year, and planetary boundary layer, then only the WHO AQG is displayed.

This script iterates through these defined layers of an ArcMap map document (.mxd) and creates a map file from a set of predetermined layers. For this script to run, a map document must already have been created, with all iterative layers turned off in the table of contents. An ArcMap layout is the reference for the map files and should contain all relevant map elements. This specific script requires a title that is labeled "title" in its description properties. A layer file (.lyr) must be created from the layer with the largest range with all symbology finalized. These files were saved as PNGs; however, other file types can easily be defined instead.

3.2 Map Products

3.2.1 Global Maps

The methodology was successful in producing a series of maps displaying SO_2. Twelve global maps corresponding to each year of data were created. These maps show the location of the major volcanic emitters and areas of interest for this study. An example map is shown in Figure 3 below. The regions of interest are boxed in blue.

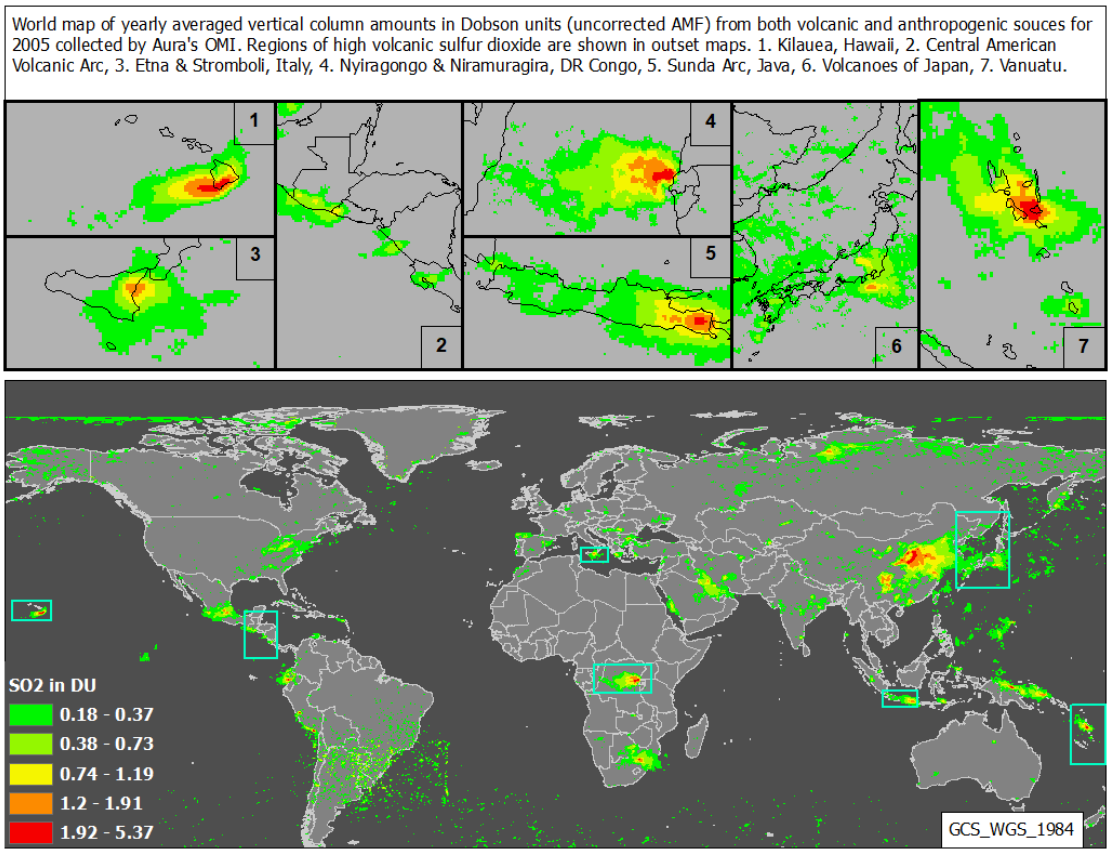

Figure 3: Global map displaying raw (uncorrected) sulfur dioxide column amounts in Dobson units for 2005.

3.2.2 SO_2 Column Maps

Twelve maps displaying corrected SO_2 VCDs were produced for each year of data and region of interest, for a total of 84 maps. An example map is shown below in Figure 4. These maps include a range of SO_2 as well as the locations and names of known volcanic emitters (Fioletov et al., 2016; Carn et al., 2017).

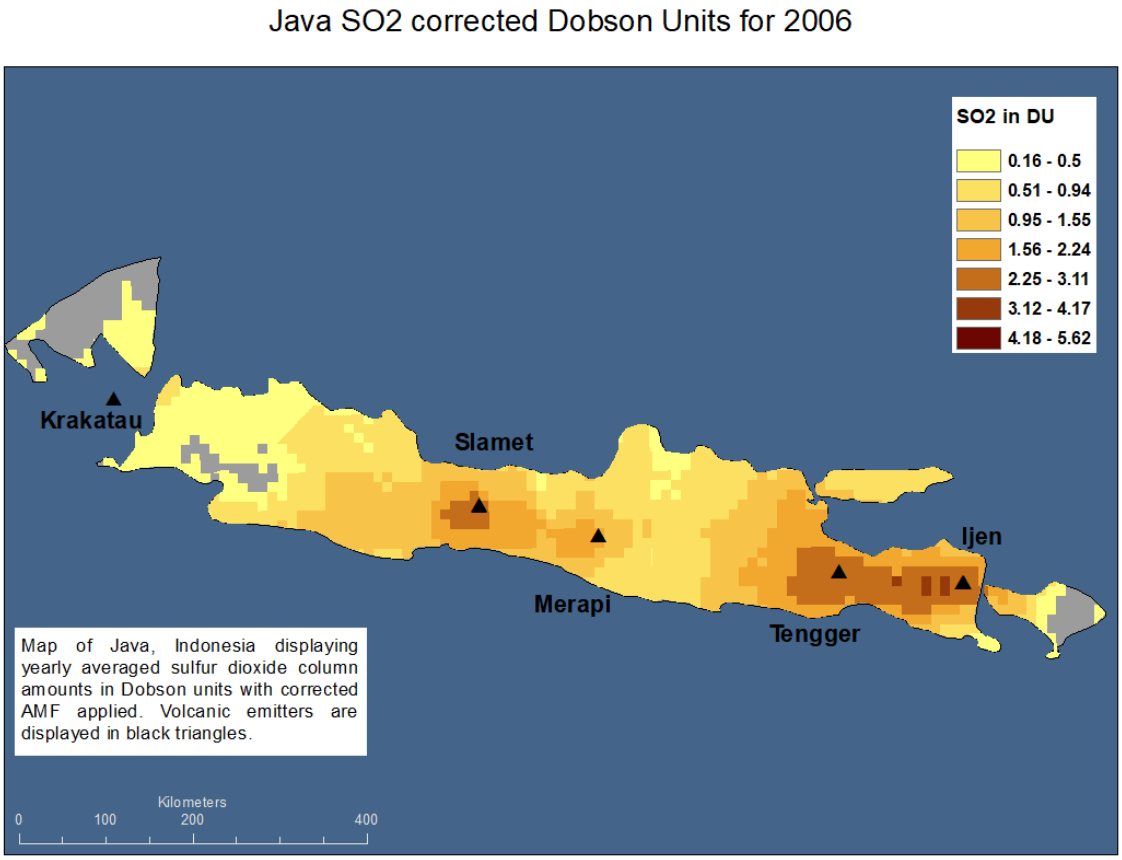

Figure 4: Map of Java displaying annual mean SO_2 vertical column amounts in Dobson units for 2006.

3.2.3 SO_2 Concentration Maps

A map of SO_2 concentrations for each region, year, and PBL condition was created resulting in 252 maps. An example map is shown below in Figure 5.

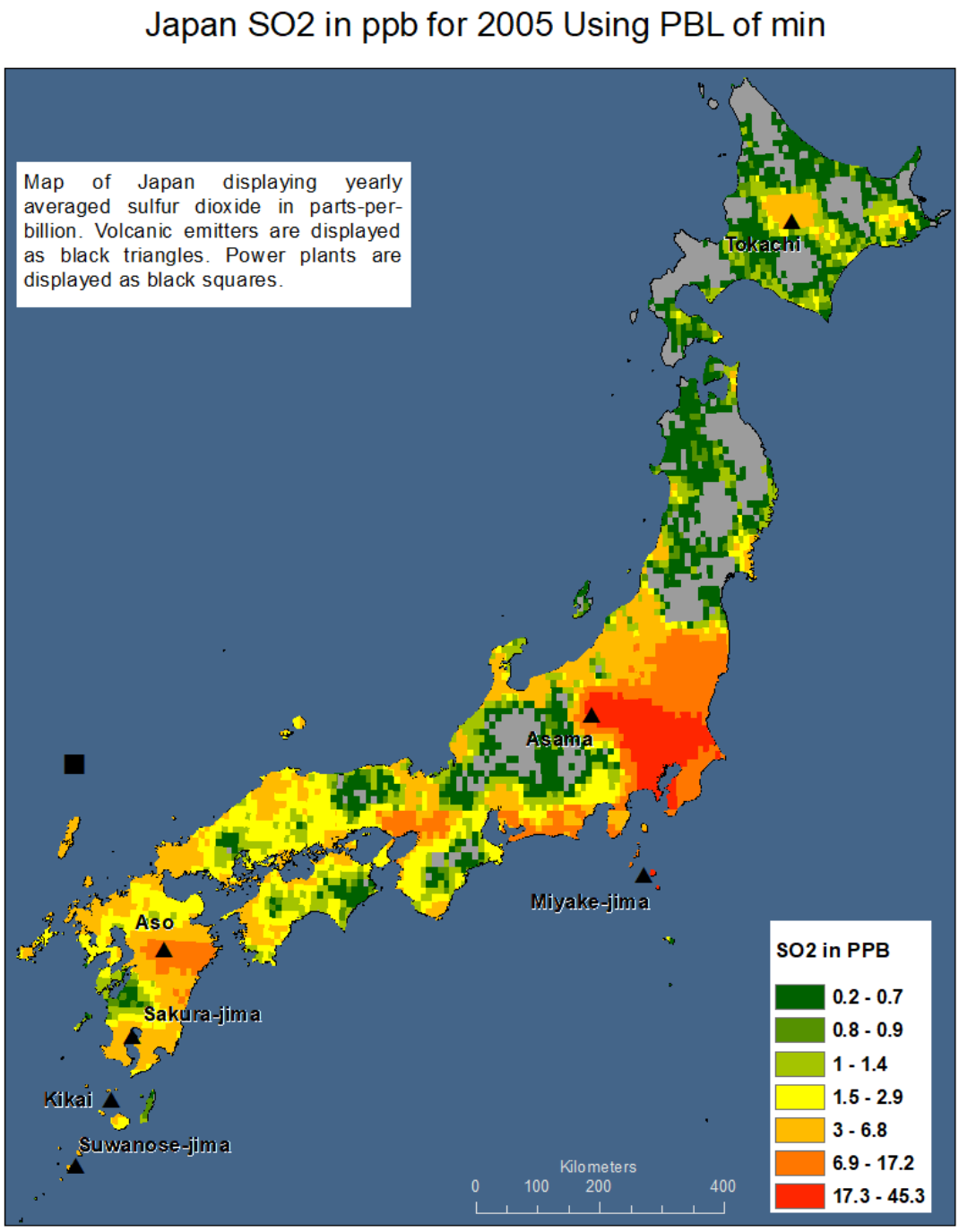

Figure 5: Map of Japan displaying sulfur dioxide in parts-per-billion using a planetary boundary layer condition of min for the year 2005.

3.2.4 Hazard Maps

A hazard map was created for every year of data, every area of interest, and every PBL condition, resulting in 252 maps. Hazard zones are denoted by hashed areas. These areas correspond to where estimate SO_2 concentrations exceed the ambient air quality standard used for the region or the WHO's air quality guideline of 7.7 ppb. Underlain is the region's population count per square kilometer (each region cell size will differ due to projection differences). An example hazard map is shown below in Figure 6.

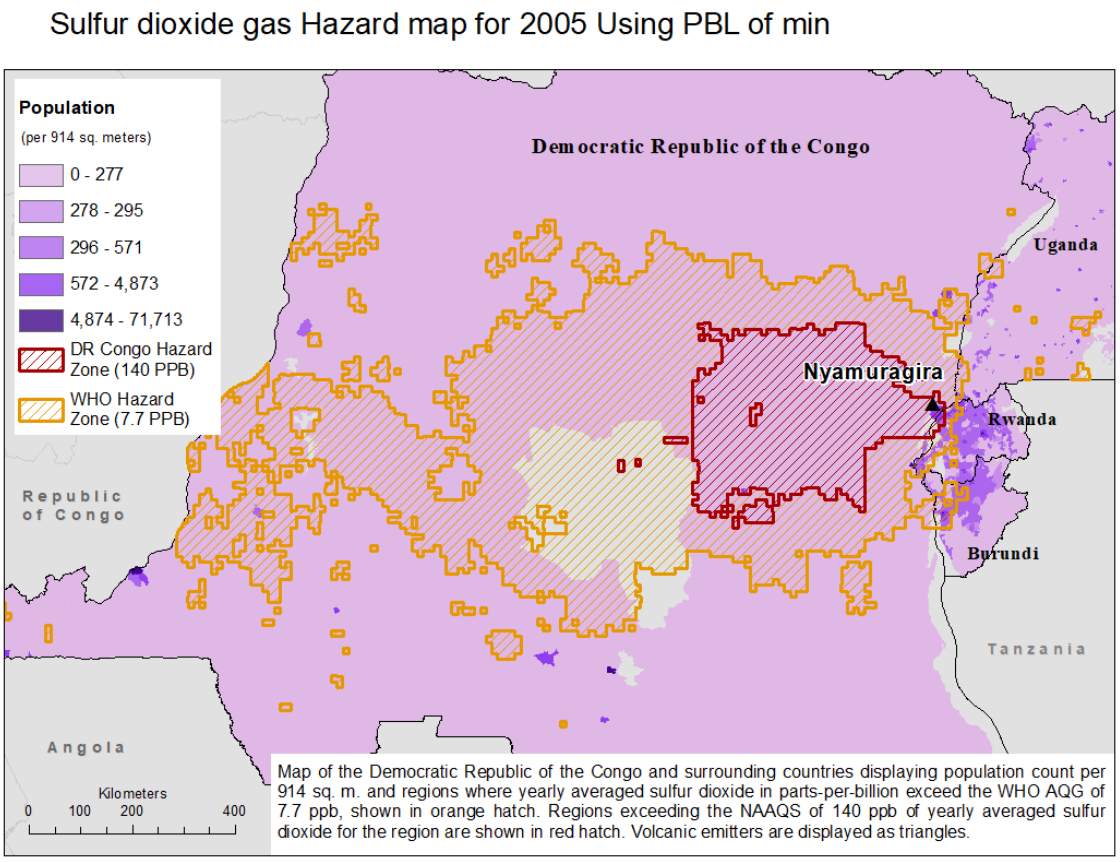

Figure 6: Map of the D.R. Congo and surrounding regions displaying population count and standards exceedance. Data displayed is for 2005 using a PBL condition of min and a regional standard of 140 ppb.

3.3 Statistical Analysis

3.3.1 Population Analysis Results

Each area of interest for each year and planetary boundary layer exceeded the WHO's AQG of 7.7 ppb. However, not every area of interest for every year exceeded their NAAQS. The table that follows outlines each region's potential population count exposed to each hazard zone. The largest potential number of people exposed to the WHO's AQG of 7.7 ppb was 70,654,456 people for 2015 using the minimum PBL height within the region of Java, Indonesia. The largest potential number of people exposed to a region's NAAQS was in the D.R. Congo during 2015 using the minimum PBL height. The D.R. Congo did not have its own NAAQS, so a NAAQS of 140 ppb was used. Potentially 4,723,771 people were exposed under these conditions. It is important to note that comparison of regions must consider different NAAQS. For example, Italy and Japan's potential population exposed to their NAAQS of 48 ppb and 40 ppb, respectively, may appear inflated compared to the other regions with regional AAQS of 140 ppb.

Table 5: Population exposed to the WHO's AQG for SO_2 of 7.7 ppb for each region and year/PBL condition. Population exposed to regional NAAQS are shown in parentheses. If parentheses are not present, the region and year/PBL condition did not exceed the AAQS for that region.

	Cent. Am	**DR Congo**	**Hawaii**	**Italy**	**Japan**	**Java**	**Vanuatu**
NAAQS	140 ppb	140 ppb	140 ppb	48 ppb	40 ppb	140 ppb	140 ppb
2005 12UT	8,569,078	7,157,711	117,643	1,769,830	30,508,772	26,946,972	63,387
2005 Max	2,296,047	7,157,711	99,844	1,657,499	9,093,111	25,656,376	63,387
2005 Min	8,689,576	14,834,395 (2,940,477)	124,451	1,476,467 (691,332)	9,014,560 (1,092,710)	40,270,864	65,345
2006 12UT	5,211,209	5,498,451	81,106	2,232,928	17,509,982	53,802,448	75,331
2006 Max	1,297,566	5,498,451	36,417	2,215,215	2,196	48,603,868	75,331
2006 Min	5,421,870	18,556,632 (2,775,564)	122,836	1,532,341 (838,484)	31,899,698	70,183,712	75,391
2007 12UT	4,027,704	5,796,497	89,702	1,964,453	286,830	25,520,234	71,006
2007 Max	557,328	5,796,497	52,209	1,914,667	4,771	23,571,346	71,006
2007 Min	4,125,920	14,387,235 (2,367,374)	104,698	1,657,006 (800,792)	19,884,526	43,496,064	72,239
2008 12UT	6,237,655	4,586,350	96,757	2,316,039	2,998,163	22,556,834	46,480
2008 Max	3,054,225	4,586,350	73,455	2,269,062	2,753	20,912,744	45,891
2008 Min	6,377,400	15,841,896 (2,055,734)	133,770	1,971,713 (831,866)	33,148,386	40,081,368	57,106
2009 12UT	7,037,292	4,251,352	139,769	2,125,441	17,686,068	16,527,802	68,905
2009 Max	3,911,513	4,251,352	116,612	2,083,790	1,085,366	14,794,632	68,546
2009 Min	7,115,880	8,526,974 (2,055,837)	143,411	1,674,199 (732,907)	36,555,640	37,421,000	69,243
2010 12UT	7,815,005 (1,009,717)	3,900,318	137,744	1,535,698	333,305	45,176,060	95,769
2010 Max	3,930,989	3,900,318	70,183	1,464,030	75,885	40,659,992	95,769
2010 Min	7,900,458 (1,009,717)	18,646,034 (1,770,517)	143,413 (3,328)	1,866,506 (237,706)	20,667,674	62,813,828	100,896

2011 12UT	6,383,891 (812,667)	6,560,795	112,197	2,556,667	1,128,269	22,149,146	58,605
2011 Max	5,020,679	6,560,795	92,741	2,544,305	865,968	19,769,300	58,488
2011 Min	6,422,458 (812,667)	23,779,958 (2,030,766)	138,323	1,918,284 (862,476)	28,409,562	38,404,196	64,710
2012 12UT	3,115,984	13,392,885	120,222	1,809,377	1,149,050	1,799,193	94,733
2012 Max	1,399,719	13,392,885	85,324	1,712,143	751,093	1,305,487	91,540
2012 Min	3,133,489	46,790,240 (4,723,771)	150,491	1,283,385 (826,117)	14,958,373	22,865,248	97,924
2013 12UT	7,170,287	13,017,342	125,833	1,802,835	1,567,190	12,931,365	121,268
2013 Max	4,209,567	13,017,342	65,169	1,765,421	831,468	12,026,145	121,268
2013 Min	7,302,762	28,352,860 (4,689,391)	158,690	2,365,975	23,807,288	35,730,108	121,372
2014 12UT	10,853,015	8,490,196	151,140	2,019,281	1,509,206	31,259,500	106,297
2014 Max	4,816,913	8,490,196	123,841	2,005,622	454,975	27,388,858	101,030
2014 Min	10,946,777	31,189,502 (3,087,234)	161,994	1,694,730 (862,229)	31,405,362	54,047,660	108,098
2015 12UT	6,862,530	13,387,539	152,405	1,576,888	20,766,888	49,894,616	90,783
2015 Max	4,113,161	13,387,539	139,041	1,558,547	432,153	44,126,696	90,789
2015 Min	6,963,723	34,720,584 (4,169,246)	165,534	1,047,840 (877,440)	35,074,300	70,654,456	100,040
2016 12UT	9,603,710	5,720,107	109,035	2,168,111	879,464	33,554,378	118,226
2016 Max	5,269,416	5,720,107	65,879	2,161,644	139,108	30,142,026	117,493
2016 Min	9,780,837	11,990,374 (2,309,287)	145,023	1,695,957 (869,343)	26,814,408	50,087,020	123,023

3.3.2 Temporal and Population-Weighted Averages Analysis

The results of the Population-weighted Averages script are a graph of sums of population and SO_2, the population-weighted average of SO_2, and average SO_2, as found in the script description of the Results Section above. This was done for both corrected SO_2 VCDs and SO_2 concentrations. Some of these results were graphed and an example graph is shown below in Figure 7. These statistics were derived from the SO_2 DU footprint, where SO_2 in DU was above the threshold of 0.15 DU, which dynamically shifts each year. These graphs show the population potentially exposed to SO_2 at a level above the threshold and the total sum of SO_2 in DU (corrected), for each region and year. It also allows comparison between average SO_2 and population-weighted average, per year, a representation of the severity of potential exposure. Graphs for all regions displaying the statistical results can be found in Appendix C.

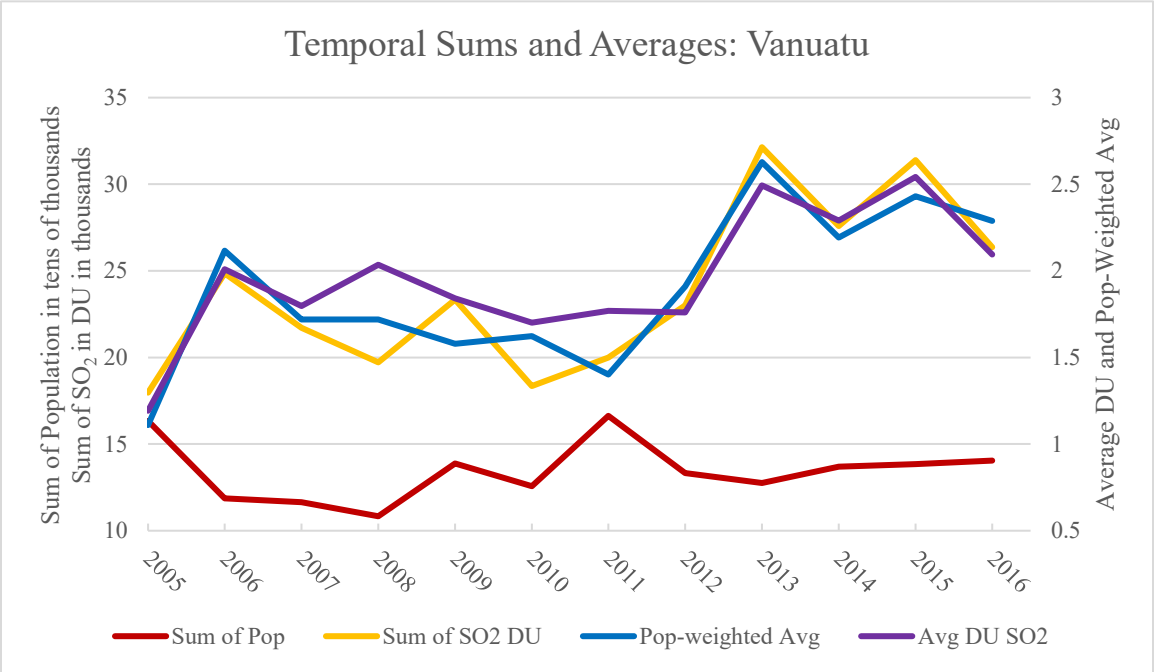

Figure 7: Temporal trends in the estimated sum of population, sum of corrected SO_2 VCD (in DU), average corrected SO_2 VCD (in DU), and population-weighted average of corrected SO_2 VCD (in DU) for Vanuatu. Population is shown in tens of thousands. SO_2 VCDs in DU are shown in thousands.

4 Findings and Discussion

4.1 Anthropogenic Influence

While attempts were made to eliminate sources of anthropogenic SO_2 from the study, a few regions included known point sources like power plants and refineries. These sources are present in Guatemala, Honduras, and Java. The emissions from these sources are relatively low compared to the volcanic sources in the area (Fioletov et al., 2016). However, they are usually at lower altitude, so may have more influence on the PBL. Despite not having any anthropogenic point sources of SO_2 itself, Japan has trans-boundary effluence pollution from sources in China and South Korea that may have a significant influence on SO_2 shown in this study. While not shown in the maps of this study due to country boundary clipping of data, the trans-boundary SO_2 can be seen when not clipped. Further investigation may be needed to determine the role that trans-boundary pollution may play in Japan's air quality.

Anthropogenic climate forcing resulting in variable climate change may impact the model used to derive SO_2 concentrations from SO_2 VCDs due to the dependence on atmospheric temperature and pressure. Rising temperatures and additional greenhouse gas flux in the atmosphere may have consequences on any continuation of this study and may be acknowledged or considered in future work.

4.2 Planetary Boundary Layer Influence

The PBL data is the main contributor to the calculation of SO_2 concentrations. The cell sizes are uniform throughout a region and all regions' cell sizes are around 1,000 meters squared, depending on latitude. The PBL controls the upper bounds of the column and determines the volume of the column, which directly influences the concentration of the SO_2 in parts-per-billion. The range of PBL conditions for each cell was represented by the min and max PBL rasters used throughout this study, while the 12UT PBL raster was used as an example of real PBL conditions from the original data source. The resulting maps created using the min and max PBL rasters are to be considered on a cell-by-cell basis. Because the cells of these rasters contain either minimum or maximum PBL values derived from the four diurnal PBL products, they do not necessarily display the real-world diurnal conditions modeled by von Engeln and Teixeira (2013).

However, a comparison between the min, max, and 12UT rasters shows that for some regions, the min or max vary little from the 12UT condition, suggesting the min or max conditions may be present during the diurnal cycle. This observation is supported visually

when comparing all PBL rasters in ArcMap (not shown), and statistically. For example, a comparison of potential population exposure by PBL and year found in Table 5 shows that for the Democratic Republic of the Congo and Vanuatu the 12UT values are the same as the max values whereas the Central America potential population exposure values for 12UT and min are similar. A future analysis could involve the use of the PBL time rasters nearest to the overpass time of OMI of 1:30 pm local time, for example, using the 00UT raster for Vanuatu and Hawaii or the 06UT PBL for Central America.

4.3 Population-Weighted Averages

The population-weighted averages (PWA) were determined for both SO_2 VCDs and concentrations for each year, region, and PBL condition (ppb only). The PWA was used to estimate which regions are most at risk for potential exposure. The ranges for ppb population-weighted averages for all regions is shown below in Figure 8. The highest ppb population-weighted average is for the D.R. Congo at 103.4 ppb PWA for 2009 with a PBL of min. The lowest is Japan at 2.4 ppb for 2010 with a PBL of max.

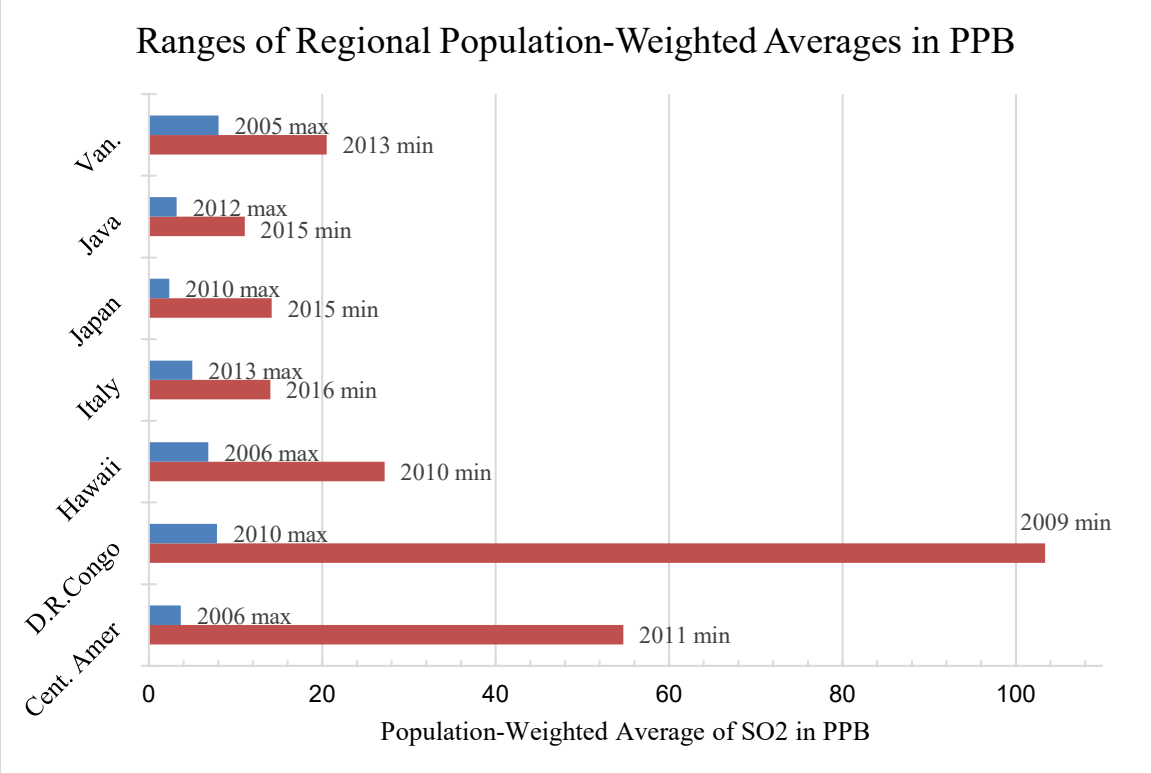

Figure 8: Maxima and minima population-weighted averages for SO_2 in parts-per-billion for each region. The PBL condition and year is also shown.

Population-weighted averages of SO_2 in DU for each year and region are shown in Figure 9 below. Java has consistently lower population-weighted averages for both SO_2 in DU and ppb, yet Table 5 shows that Java has the highest population (70,654,456 people) for potential exposure to SO_2 in ppb over the WHO's AQG of 7.7 ppb while never exceeding the NAAQS of 140 ppb for any year or PBL condition. Both Figure 8 and Figure 9 show Central America as one of the regions most at risk for potentially hazardous exposure. Vanuatu has the highest population-weighted average SO_2 in DU but is one of the lower regions for population-weighted averages in ppb, possibly due to the behavior of the PBL. The Democratic Republic of the Congo is in the top three highest population-weighted average for SO_2 in DU but has the highest population-weighted average in ppb. These two cases illustrate the impact of the planetary boundary layer in determining potential exposure and calls for further investigation into mixing models and validation studies.

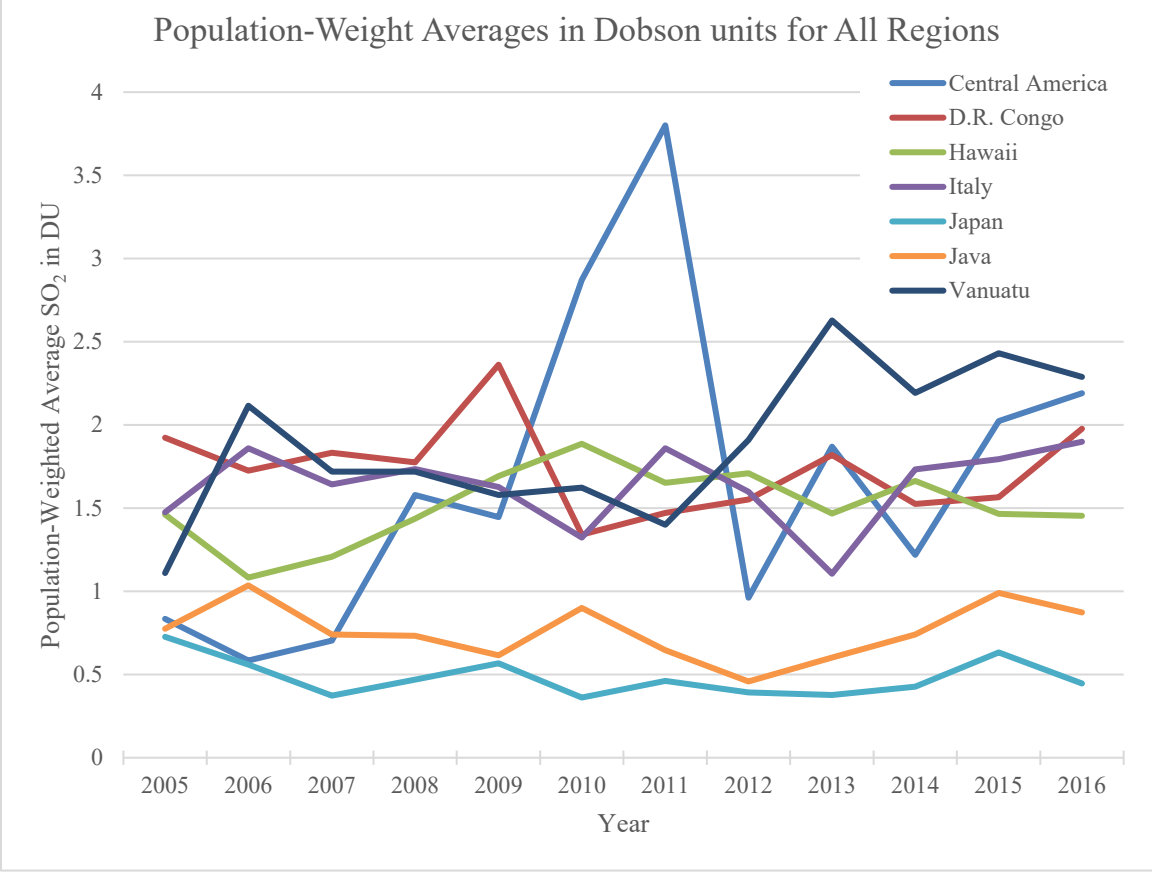

Figure 9: Population-weighted averages SO_2 in Dobson units for each year and region.

4.4 Factors of Vulnerability

Continuation of this work could include further analysis of public health policies in these regions. While the assumptions inherent in this study make specific conclusions impractical, the results of the study show that these regions may be experiencing elevated levels above what is recommended by the WHO AQG, and in some cases above the region's own NAAQS. The WHO uses epidemiological studies to update their air quality guidelines based on the health impacts on the most vulnerable populations and offer suggestions to governments on their national standards (WHO, 2018). Vahlsing & Smith (2012) suggest that while most nations in their survey use the WHO suggestions when creating their standards, non-science-based business and economists may have an influence on the standards set. Comparisons of Table 5 population estimates from the WHO's AQG and the NAAQS may indicate that millions of people are not accounted for in exposure estimates if only NAAQS are used, some of which may be the particularly vulnerable populations.

Future work could examine the following questions. Which countries monitor air pollution, specifically around volcanic emission sites? Do these countries monitor respiratory disease cases and produce longitudinal studies that impact public policy? Do these countries have action plans for evacuation when air quality is hazardous? Which countries have robust medical infrastructure to care for affected populations?

In addition, there are several other factors of vulnerability to consider. For example, a history of political and civil unrest in the Goma region of the Democratic Republic of the Congo puts potentially exposed populations at greater risk due to limited resources. Japan has experienced recent natural and technological disasters which could potentially strain resources needed to respond to other hazards. Each region will have unique factors of vulnerability and capacity that should be considered in a full hazard analysis.

4.5 Limitations and Improvements

Comparisons between regions were limited by difference in population, land area coverage, number of emitters per region, and a difference in NAAQS. This work could further be refined by limiting scale to individual emitters as data resolution allows, enabling comparisons between individual emitters. Furthermore, this study confined the yearly average column amounts to land mass, clipping at land borders, to assess exposure of human populations. This omitted near-shore emissions that populations may be exposed to, as well as not accounting for potential environmental or infrastructure damage that off-shore emissions may cause. This limitation also eliminated the ability to directly compare results with previous studies like Carn et al. (2017), where the yearly

summed mass of emitters was calculated in kilotons. For future work and verification purposes, SO_2 data can be analyzed before clipping or without clipping.

Future work could also involve additional population data sets. GPW v.4 (2016) is an open-source, academic, robust data series from CIESIN that extrapolated population count data for every 5 years. This study interpolated this data to produce yearly data sets, which is not recommended. A potential improvement would involve the use of yearly data sets not requiring interpolation, such as World Pop (n.d.) population count, a resource not available at the initiation of this study.

The PBL data provided by von Engeln and Teixeira (2013) has a spatial resolution of 1 degree. Compared to the other datasets used, this is coarse. If this data was refined or if other suitable PBL data was found, it could improve the conversion of SO_2 column to concentration. Raster resampling associated with some GIS analysis tools along with the differing cell sizes of input data can result in some spatial deviation from original data. An error analysis would be needed on any further refinement of this research, especially if this methodology is applied to larger scaled areas of interest.

Improvements to the methodology should include investigation into the automation of the manual division workflow to determine regions of volcanic influence for assigning AMF. Visual delineating between emitters involved some amount of data-supported guesswork. Modifying an existing toolset, like those found in the hydrology tools of ArcGIS or a least-cost path analysis, may be appropriate and would eliminate the guesswork.

The data and methodology used in this research limit the potential health and environmental effects to those caused only by volcanic SO_2. This study does not investigate other potentially significant sources of volcanic air pollution such as sulfuric acid (sulfate) aerosols or other particulate matter (e.g., Halliday et al., 2019; Longo et al., 2005; Tofte et al., 2017). Longo et al. (2005) measured both SO_2 and fine aerosol concentrations on the island of Hawaii, finding that some measurements of both exceeded health standards. After emissions, SO_2 oxidation occurs on various timescales (hours to days) depending on factors such as plume altitude and atmospheric humidity, resulting in acidic aerosols in the form of clouds or fog. In conditions further away from volcanic vents, volcanic fog, or vog, can consist mainly of these particulate aerosols (Longo et al., 2005; Tofte et al., 2017). This process can occur readily in tropical or sub-tropical climates. Many regions analyzed in this study fall within these climates, yet these hazardous aerosols are unaccounted for in this research. Future work could include models that incorporate climate and weather data to estimate SO_2 oxidation and amounts of sulfuric aerosols to further refine the potential hazards of volcanic gas, similar to Tofte et al. (2017).

4.6 Further Work

Determining the extent to which volcanic SO_2 is impacting populations requires the incorporation of dispersion modeling and validation techniques like the use of ground sensors. This study assumed that SO_2 is well mixed throughout the PBL column. This assumption does not allow for small scale changes in SO_2 concentration due to wind dispersal, atmospheric conditions like temperature and humidity, terrain, and other parameters. Future work could refine the estimate of population exposure to SO_2 by incorporating atmospheric models such as AERMOD. AERMOD is a modeling system developed by the American Meteorological Society and the U.S. EPA. It uses a minimum of two data processor inputs, AERMET for meteorological data and AERMAP for terrain data, to model "air dispersion based on planetary boundary layer turbulence structure and scaling concepts, including treatment of both surface and elevated sources, and both simple and complex terrain" (Support Center for Regulatory Atmospheric Modeling (SCRAM), n.d.). There are numerous cases of AERMOD's use for dispersion modeling of SO_2 from oil refineries and other anthropogenic sources (e.g., Chen et al., 2012, Amoatey et al., 2018, and others). However, its application for modeling volcanic gas is confined to discrete eruption events (e.g., Jenkins et al., 2015). Continuation of this work could contribute to early applications of AERMOD for modelling volcanic gas emissions over longer and variable periods of time.

Population exposure values can potentially be validated by near- or on-ground air quality monitoring station data. Comparisons of SO_2 at ground level to levels predicted by AERMOD and detected by OMI could lead to analysis on gas layer thickness, dispersal rate, and residence time and further refine the extent to which volcanic gas plays a role in volcanic hazards. An example of this would be the use of on-ground or near-ground monitoring stations with compact spectrometers that measure volcanic SO_2 plume levels near population centers identified in this study, similar to the methodologies of Hidalgo et al. (2015) and others. The Network for Observation of Volcanic and Atmospheric Change (NOVAC, n.d.) may be another source for previously collected data for use in ground validation.

In addition to expanding this study to other years and regions, the continuation of this research could support work in various fields, including:

- Defining areas for mitigation and implementation of hazard monitoring, e.g., a low-technology mitigation strategies is the use of SO_2-resistant trees and shrubs as windbreak vegetation to protect local communities and economically important crops (Delmelle et al., 2002; van Manen, 2014).
- Using refined temporal data, such as daily SO_2 data, with a similar methodology presented in this work to compare SO_2 loads to daily AQG and NAAQS to gain further insight into exposure to volcanic gases.
- Using refined temporal data and measured SO_2/X gas concentration ratios of volcanic plumes to estimate other volcanic gas emissions (e.g., Aiuppa et al.,

2008; Delmelle et al., 2002; Lages et al., 2020) to include other potentially hazardous gases in a hazard analysis.

- Investigating NAAQS policy making factors in conjunction with toxicology and morbidity studies to refine risk and hazard of these populations' long-term exposure to SO_2.
- Creating probabilistic hazard maps for hazard analysis and forecasting (e.g., Barsotti, 2020).
- Refining spatial and/or temporal resolution of study by using data from TROPOMI, the Tropospheric Monitoring Instrument on Copernicus Sentinel-5 Precursor platform. Like OMI, TROPOMI uses an algorithm to retrieve SO_2 data from UV (310-405 nm) measurements. At nadir, TROPOMI has a spatial resolution of 3.5 km by 5.5 km and can "discern very fine details in the SO_2 distribution and will be able to detect point sources with annual SO_2 emissions of about 10 kT yr^{-1} or lower." (Theys et al., 2017).

5 Conclusions

This study presented a novel methodology for the creation of volcanic gas hazard maps. Four different map products were created including global SO_2, SO_2 VCDs, estimated SO_2 concentrations for variable PBL heights, and hazard maps displaying population and regions of potentially heightened exposure based on WHO air quality guidelines (AQG) and regional or national ambient air quality standards (NAAQS). This study also made the first attempt to estimate the number of people potentially exposed to volcanic SO_2 exceeding these standards on regional scales for several volcanic regions. Temporal analysis involved using population-weighted averages of SO_2 per year from 2005 to 2016 to evaluate the relative risk of each region. Maps of seven different regions reveal the potential for thousands of people to be exposed to hazardous levels of SO_2.

Due to health, infrastructure, and subsequent economic impacts of long-term exposure to ambient volcanic gas, and the fact that SO_2 may indicate the presence of other hazardous volcanic gasses, future work on quantifying or mapping volcanic gas hazards and their effects on public health are needed. This study contributes to the foundation for such work and suggests on- or near-ground sensors for validation purposes as well as vertical column mixing models for areas of continuation.

6 Reference List

Aiuppa, A., Giudice, G., Gurrieri, S., Liuzzo, M., Burton, M., Caltabiano, T., McGonigle, A.J.S., Salerno, G., Shinohara, H., & Valenza, M. (2008). Total volatile flux from Mount Etna. *Geophysical Research Letters, 35*(24), L24302. https://doi.org/10.1029/2008GL035871

Amoatey, P., Omidvarborna, H., Affum, H.A., & Baawain, M. (2018). Performance of AERMOD and CALPUFF models on SO2 and NO2 emissions for future health risk assessment in Tema Metropolis. *Human and Ecological Risk Assessment, 25*(3), 772-786. https://doi.org/10.1080/10807039.2018.1451745

Barquero, J. & Fernandez, E. (1990). Erupciones de gases y sus consecuencias en el Volcan Poás, Costa Rica. *Boletin de Vulcanologia Universidad Nacional, Costa Rica, 21*, 13-18. [In Spanish.]

Barsotti, S. (2020). Probabilistic hazard maps for operation use: the case of SO2 air pollution during the Holuhraun eruption (Bárðarbunga, Iceland) in 2014-2015. *Bulletin of Volcanology, 82*, 56. https://doi.org/10.1007/s00445-020-01395-3

Baxter, P.J., Baubron, J-C, & Coutinho, R. (1999). Health hazards and disaster potential of ground gas emissions at Furnas volcano, São Miguel, Azores. *Journal of Volcanology and Geothermal Research, 92*(1-2), 95-106. https://doi.org/10.1016/S0377-0273(99)00070-0

Brown, S.K., Jenkins, S.F., Sparks, R.S. J., Odbert, H., & Auker, M.R. (2017). Volcanic fatalities database: analysis of volcanic threat with distance and victim classification. *Journal of Applied Volcanology, 6*, 15. https://doi.org/10.1186/s13617-017-0067-4

Calder, E., Wagner, K., & Ogburn, S. (2015). Volcanic hazard maps. In S. Loughlin, S. Sparks, S. Brown, S. Jenkins, & C. Vye-Brown (Eds.), *Global Volcanic Hazards and Risk* (pp. 335-342). Cambridge University Press. https://doi.org/10.1017/CBO9781316276273.022

Carn, S. A., Fioletov, V.E., McLinden, C.A., & Krotkov, N.A. (2017). A decade of global volcanic SO2 emissions measured from space. *Nature Scientific Reports 7*, 44095. https://doi.org/10.1038/srep44095

Center for International Earth Science Information Network - CIESIN - Columbia University. (n.d.a). *What are the advantages of GPW over other products?* SEDAC User Feedback and Frequently Asked Questions (FAQs). https://sedac.uservoice.com/knowledgebase/articles/917760-what-are-the-advantages-of-gpw-over-other-products

Center for International Earth Science Information Network - CIESIN - Columbia University. (n.d.b). *GPWv4 releases data for five years: 2000, 2005, 2010, 2015, 2020. Is it advisable to compare data sets of different years in order to assess changes in population density over time? Are there constraints in using the GPWv4 target year data sets as a time series?* SEDAC User Feedback and Frequently Asked Questions (FAQs). https://sedac.uservoice.com/knowledgebase/articles/799152-gpwv4-releases-data-for-five-years-2000-2005-20

Center for International Earth Science Information Network - CIESIN - Columbia University. (2016). *Gridded Population of the World, Version 4 (GPWv4): Population Count*. [Dataset]. NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4X63JVC

Chen, J.A., Zapata, A.R., Sutherland, A.J., Molmen, D.R., Chow, B.S., Wu, L.E., & Hesse, R.C. (2012). Sulfur dioxide and volatiles organic compound exposure to a community in Texas City, Texas evaluated using AERMOD and empirical monitoring data. *American Journal of Environmental Sciences, 8*(6), 622-632. https://doi.org/10.3844/ajessp.2012.622.632

Clean Air Initiative for Asian Cities (CAI-Asia) Center (2010). *Air quality in Asia: Status and trends - 2010 Edition*. Clean Air Initiative for Asian Cities (CAI-Asia) Center. Retrieved from http://www.indiaenvironmentportal.org.in/files/AirQualityinAsiaStatusandTrends2010.pdf

Crandall, D.R., Booth, B., Kusumadinata, K., Walker, G.P.L, & Westercamp, D. (1984). *Source-book for volcanic-hazards.* UNESCO.

Delmelle, P., Stix, J., Baxter, P.J., Garcia-Alvarez, & J., Barquero, J. (2002). Atmospheric dispersion, environmental effects and potential health hazard associated with the low-altitude gas plume of Masaya volcano, Nicaragua. Bulletin of Volcanology, 64(6), 423-434. https://doi.org/10.1007/s00445-002-0221-6

Dunkley, P N & Young, S R. (2000). Volcanic hazard mapping for development planning. *BGS Technical Report WC/00/20*. http://nora.nerc.ac.uk/id/eprint/509938/1/WC00020.pdf

ESRI (n.d. a). *Cell size and resampling in analysis. ArcMap Help Documentation.* https://desktop.arcgis.com/en/arcmap/10.7/extensions/spatial-analyst/performing-analysis/cell-size-and-resampling-in-analysis.htm

ESRI (n.d. b). *Data classification*. ArcMap Help Documentation. *https://desktop.arcgis.com/en/arcmap/10.7/extensions/geostatistical-analyst/data-classification.htm*

ESRI (n.d. c). *List of supported map projections*. ArcMap Help Documentation. https://desktop.arcgis.com/en/arcmap/10.7/map/projections/list-of-supported-map-projections.htm

European Commission. (n.d.). *Air Quality Standards*. Environment. https://ec.europa.eu/environment/air/quality/standards.htm

Green, J., & Sánchez, S. (2013). Air Quality in Latin America: An Overview. *Clean Air Institute*. Retrieved from https://www.yumpu.com/en/document/read/41258091/air-quality-in-latin-america-an-overview-clean-air-institute

Halliday, T.J., Lynham, J., & de Paula, A. (2019). Vog: Using volcanic eruptions to estimate the health costs of particulates. *The Economic Journal, 129*(620), 1782-1816. https://doi.org/10.1111/ecoj.12609

Hansell, A. & Oppenheimer, C. (2004). Health hazards from volcanic gases: A systematic literature review. *Arch. Environ. Health, 59*(12), 628–639. https://doi.org/10.1080/00039890409602947

Jiang, L. (2016). Global country boundaries [shapefile]. Retrieved from https://www.arcgis.com/home/item.html?id=2ca75003ef9d477fb22db19832c9554f

Jenkins, S. F., Barsotti, S., Hincks, T.K., Neri, A., & Phillips, J.C. (2015). Rapid emergency assessment of ash and gas hazard for future eruptions at Santorini Volcano, Greece. *Journal of Applied Volcanology, 4*, 16. https://doi.org/10.1186/s13617-015-0033-y

Lages, J., Moussallam, Y., Bani, P., Peters, N., Aiuppa, A., Bitetto, M., & Giudice, G. (2020). First in-situ measurements of plume chemistry at Mount Garet Volcano, Island of Gaua (Vanuatu). *Applied Science, 10*(20), 7293. https://doi.org/10.3390/app10207293

Li, C., Joiner, J., Krotkov, N. A. & Bhartia, P. K. (2013). A fast and sensitive new satellite SO_2 retrieval algorithm based on principal component analysis: application to the Ozone Monitoring Instrument. *Geophysical Research Letters, 40*(23), 6314–6318. https://doi.org/10.1002/2013GL058134

Li, C., McLinden, C., Fioletov, V., Krotkov, N., Carn, S., Joiner, J., Streets, D., He, H., Ren, X., Li, Z., & Dickerson, R. R. (2017). India is overtaking China as the world's largest emitter of anthropogenic sulfur dioxide. *Scientific Reports, 7*, 14304. https://doi.org/10.1038/s41598-017-14639-8

Longo, B.M., Rossignol, A., & Green, J.B. (2008). Cardiorespiratory health effects associated with sulphurous volcanic air pollution. *Public Health, 122*(8), 809-820. https://doi.org/10.1016/j.puhe.2007.09.017

Longo, B.M., Grunder, A., Chuan, R., & Rossignol, A. (2005). SO_2 and fine aerosol dispersion from the Kilauea plume, Kau district, Hawaii, USA. *Geology, 33*(3), 217–220. https://doi.org/10.1130/G21167.1

McLinden, C.A., Fioletov, V., Beorsma, K.F., Kharol, S.K., Krotkov, N., Lamsal, L., Makar, P.A., Martin, R.V., Veefkind, J.P., & Yang, K. (2014). Improved satellite retrievals of NO_2 and SO_2 over the Canadian oil sands and comparisons with surface measurements. *Atmospheric Chemistry and Physics, 14*, 3637-3656. https://doi.org/10.5194/acp-14-3637-2014

McLinden, C. A., Fioletov, V., Shephard, M.W., Krotkov, N., Li, C., Martin, R.V., Moran, M.D., & Joiner, J. (2016), Space-based detection of missing sulfur dioxide sources of global air pollution. *Nature Geoscience, 9*(7), 496–500. https://doi.org/10.1038/ngeo2724

Ministry of the Environment, Government of Japan (n.d.). *Environmental quality standards in Japan - Air quality*. Ministry of the Environment, Government of Japan. https://www.env.go.jp/en/air/aq/aq.html

Network for Observation of Volcanic and Atmospheric Change (NOVAC) (n.d.). Home. *NOVAC*. https://novac-community.org/

Pressler, V., & Ige, D.Y. (2016). *State of Hawaii annual summary 2016 air quality data*. State of Hawaii Department of Health. https://health.hawaii.gov/cab/files/2019/07/aqbook_2016.pdf

Robock, A. (2000). Volcanic eruptions and climate. *Reviews of Geophysics, 38*(2), 191-219. https://doi.org/10.1029/1998RG000054

Šavrič, B., Jenny, B. & Jenny, H. (2016). Projection Wizard – An online map projection selection tool. *The Cartographic Journal, 53*(2), 177–185. https://doi.org/10.1080/00087041.2015.1131938

Sears, T. M., Thomas, G. E., Carboni, E., Smith, A. J. A. & Smith, Grainger, R. G. (2013). SO2 as a possible proxy for volcanic ash in aviation hazard avoidance. *JGR: Atmospheres, 118*(11), 5698-5709. https://doi.org/10.1002/jgrd.50505

Support Center for Regulatory Atmospheric Modeling (SCRAM) (n.d.). *Air quality dispersion modeling – Preferred and recommended models.* United States Environmental Protection Agency (U.S. EPA). https://www.epa.gov/scram/air-quality-dispersion-modeling-preferred-and-recommended-models

Theys, N., De Smedt, I., Yu, H., Danckaert, T., van Gent, J., Hörmann, C., Wagner, T., Hedelt, P., Bauer, H., Romahn, F., Pedergnana, M., Loyola, D., & Van Roozendael, M. (2017). Sulfur dioxide retrievals from TROPOMI onboard Sentinel-5 Precursor: algorithm theoretical basis. *Atmospheric Measurement Techniques, 10*(1), 119–153. https://doi.org/10.5194/amt-10-119-2017

Tofte, K., Chu, P-.S., & Barnes, G.M. (2017) Large-scale weather patterns favorable for volcanic smog occurrences on O'ahu, Hawai'i. *Air Quality, Atmosphere & Health, 10*(10), 1163-1180. https://doi.org/10.1007/s11869-017-0502-z

United States Environmental Protection Agency (U.S. EPA) (2017). Timeline of sulfur dioxide National Ambient Air Quality Standards (NAAQS). United States Environmental Protection Agency. https://www.epa.gov/so2-pollution/table-historical-sulfur-dioxide-national-ambient-air-quality-standards-naaqs

UNRESP (2017). The project. *UNRESP.* https://unresp.wordpress.com/about-unresp/

Vahlsing, C., & Smith, K.R. (2012). Global review of national ambient air quality standards for PM_{10} and SO_2 (24 h). *Air Quality, Atmosphere & Health*, *5*, 393-399. https://doi.org/10.1007/s11869-010-0131-2

van Manen, S.M. (2014). Perception of a chronic volcanic hazard: persistent degassing at Masaya volcano, Nicaragua. *Journal of Applied Volcanology, 3*, 9. https://doi.org/10.1186/s13617-014-0009-3

von Engeln, A., & Teixeira, J. (2013). A planetary boundary layer height climatology derived from ECMWF reanalysis data. *Journal of Climate*, *26*(17), 6575–6590. https://doi.org/10.1175/jcli-d-12-00385.1

World Health Organization (WHO). (2018). *Ambient (outdoor) air pollution.* World Health Organization. https://www.who.int/en/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health

World Pop. (n.d.). *Population Counts.* World Pop. *https://www.worldpop.org/project/categories?id=3*

Appendix A Python Scripts

These scripts are designed to be used for the continuation of this or similar work. Only scripts relevant to the methodology were included; however, all scripts may be available by contacting the author. Scripts can be copy/pasted into a code editor or IDE and will only work if ArcGIS is installed. Note that pathnames, variables, and possibly more will need to be updated based on the user's analysis and database structure. These codes were written in Python 2.7 for ArcGIS 10.7. Future work may require converting codes to different Python versions. Descriptions of codes in the Results Section are intended to be copied with code to provide further guidance on limitations and use.

A.1 Correcting Dobson units & Estimating Parts-per-Billion

```
# Correcting Dobson units & Estimating Parts-per-Billion
# Sanna Mairet

# Description: This script creates parts-per-billion SO2 rasters from several
# pre-synthesized data including dissected regional feature classes (see Fig. 1),
# SO2 in Dobson units (uncorrected), while creating and using pressure,
# temperature, and altitude rasters. It also creates and uses AMF rasters
# to apply to SO2 in Dobson unit data to for the creation of rasters of
# corrected SO2 in Dobson units. The workflow is outlined in Figure 2.

import arcpy, os, sys, traceback, numpy
from arcpy import env, da
from arcpy.conversion import PolygonToRaster
from arcpy.sa import *

arcpy.CheckOutExtension("Spatial")

#Allow script to overwrite existing feature classes
arcpy.env.overwriteOutput = True

#Set workspace path
arcpy.env.workspace = "E:\\HawaiiAnalysis.gdb\\"
#arcpy.env.workspace = arcpy.GetParameterAsText(0)
```

```
#Set Variables
# Create/find token for exact cell size
cellsizeresult =
arcpy.GetRasterProperties_management("E:\\HawaiiAnalysis.gdb\\SO2_2005_0
083clipproj", "CELLSIZEX")
cellsize = cellsizeresult.getOutput(0)
print 'cellsize is ', cellsize
f_cellsize = float(cellsize)
cellsizearea = f_cellsize**2
print 'cellsize area is ', cellsizearea

# Variable Constants
Boltzmann_constant = 1.38E-23
Dobson_units_to_molecules = 2.687E+20
amf_standard = 0.36
PPB_conv_factor = 1E9
# PPB_conv_factor is to get data into ppb format

# This snippet lists the polygon and raster layers used in the equations to
# convert DU to PPB.
attpoly = arcpy.ListFeatureClasses("*final", "Polygon", "AOIs2000sLayers")
so2rasters = arcpy.ListRasters("*clipproj", "All")
pblrasters = arcpy.ListRasters("projclippbl*","All")

# This for loop creates altitude, temp, and pressure rasters from the PBL rasters
for pblrast in pblrasters:
    Rpbl=Raster(pblrast)
    print 'creating altitude, temp, and pressure rasters from ', pblrast

    altrast_out = arcpy.env.workspace + "\\" + pblrast[8:] + "_median_alt_m"
    print altrast_out
    altrast = (Rpbl / 2) * 1000
    altrast.save(altrast_out)
    print 'altrast created from ', pblrast

    tempk_out = arcpy.env.workspace + "\\" + pblrast[8:] + "_tempk"
    tempk = 288.15 - (0.0065 * altrast)
    tempk.save(tempk_out)
    print 'tempk raster created from ', pblrast
```

```
    prespa_out = arcpy.env.workspace + "\\" + pblrast[8:] + "_prespa"
    prespa = 101325 * ((288.15 / tempk) ** -5.2525)
    prespa.save(prespa_out)
    print 'prespa raster created from ', pblrast

    if len(attpoly) == len(so2rasters):
    # len() determines the length of a list. This is to make sure the same
    # number of items are in the raster and polygon list. Sort below will
    # sort them into years so 2005 poly corresponds with 2005 raster.
        attpoly.sort()
        print attpoly
        so2rasters.sort()
        print so2rasters
        irast= -1
        for fc in attpoly:

        # This snippet converts the AMF polygon to a raster
            attrastamf = arcpy.env.workspace + "\\" + "AMF"+ fc[3:7]
            arcpy.PolygonToRaster_conversion(fc, "AMF", attrastamf,
"MAXIMUM_AREA", "",f_cellsize)
            attrastamfR= Raster(attrastamf)
            print 'AMF raster complete for ' + fc[3:7] + 'called ' + attrastamf
        # Each SO2 data set came pre-corrected for an AMF of 0.36. This needs to
        # to be undone and a new AMF applied. This produces a set of rasters for
        # each year that are correct Dobson units.
            irast += 1
            inrasters = so2rasters[irast]

            Rinrasters= Raster(inrasters)
            correctedfordu =  (Rinrasters/amf_standard)*attrastamfR
            correctedfordu_out = arcpy.env.workspace + "\\" + inrasters +
"DUcorrected"
            correctedfordu.save(correctedfordu_out)
            print 'Corrected Dobson units completed. New maximum is ',
correctedfordu.maximum
            print 'Old maximum was ', Rinrasters.maximum
            print correctedfordu_out
```

```
        # This snippet calculates total molecules in the atmospheric column
        #Rpbl is multiplied by 1000 because it needs to be converted to meters from
kilometers
            ttlmol_elev = (prespa * Rpbl * 1000 * cellsizearea) / (tempk *
Boltzmann_constant)
            print 'total molecules maximum using ' + pblrast + ' is ',
ttlmol_elev.maximum
            print 'total molecules minimum using ' +pblrast + ' is ',
ttlmol_elev.minimum

        #Finding the total number of so2 molecules in the atmospheric column
            so2molttl = (correctedfordu* cellsizearea * Dobson_units_to_molecules)
            print 'total SO2 molecules maximum using ' + pblrast + ' is ',
so2molttl.maximum
            print 'total SO2 molecules minimum using ' + pblrast + ' is ',
so2molttl.minimum

        # Finding parts per billion
            ppbso2 = (PPB_conv_factor * so2molttl) / (ttlmol_elev - so2molttl)
            ppbso2.save(arcpy.env.workspace + "\\" + "PPB"+ inrasters[4:8] +
"_pbl"+ pblrast[11:])
            print ppbso2
            print 'ppb max using elevation ' + pblrast + 'is: ', ppbso2.maximum
    else:
        print "Error! Source data mismatch"
```

A.2 Population and Hazard Analysis 1: Contouring Standards

```
# Population and Hazard Analysis 1: Contouring Standards
# Sanna Mairet

# Description: This code creates contours (polygon) at the WHO standard and the
# regional standard. If the parts-per-billion exceeds the regional standard, then
# the WHO standard contour is erased where the regional standard polygon
# overlaps.
```

```
import arcpy, os, sys, traceback, numpy
from arcpy import env, da
from arcpy.sa import *

arcpy.CheckOutExtension("Spatial")

# Allow script to overwrite existing feature classes
arcpy.env.overwriteOutput = True

# Set workspace path
arcpy.env.workspace = "E:\\DRCongoAnalysis.gdb\\"

# Set County/Regional Standard
regstd = 140
whostd= 7.7

# List parts-per-billion SO2 data
so2ppbrast = arcpy.ListRasters("rfPPB*" ,"All")

# Begin loop on ppb data creating contours based on WHO AQG
for ppb in so2ppbrast:
   rppb = Raster(ppb)

   out7_7cont = arcpy.env.workspace + "\\con7_7" + ppb
   continv77 = 7.7
   basecont = 0
   zfac = 1
   conttype = "CONTOUR_SHELL_UP"
   Contour(rppb, out7_7cont, continv77, basecont, zfac, conttype)
   print out7_7cont, "created"

   # Delete all attribute rows that aren't 7.7 to max
   field = "ContourMin"
   with arcpy.da.UpdateCursor(out7_7cont, field) as cursor:
      for row in cursor:
         if  row [0] < 7.69 or row[0] > 7.71:
            cursor.deleteRow()
            print row, "row deleted"
```

```
        else:
            print  out7_7cont, "row 7.7 preserved"
    # If ppb SO2 data exceeds regional AAQG, then create a new contour
    # for this data
    if rppb.maximum > regstd:
        outregstdcont = arcpy.env.workspace + "\\con" + str(regstd) + ppb
        Contour(rppb, outregstdcont, regstd, basecont, zfac, conttype)
        print ppb, rppb.maximum

        #Delete all attributes that aren't regional standard
        with arcpy.da.UpdateCursor(outregstdcont, field) as cursorregstd:
            for rowregstd in cursorregstd:
                if rowregstd[0] != regstd:
                    cursorregstd.deleteRow()
                    print outregstdcont, "rows deleted"
                else:
                    print  regstd, "row preserved"

        # Erase the regional contour polygon from the WHO AQG
        # contour polygon to eliminate overlapping data
        new7_7 =  arcpy.env.workspace + "\\con7_7" + ppb +"clp"
        arcpy.Erase_analysis(out7_7cont, outregstdcont, new7_7)
        print new7_7, "created via erase"
        arcpy.Delete_management(out7_7cont)
        print out7_7cont, "deleted"

    else:
        print rppb.maximum, "is less than ", regstd
```

A.3 Population and Hazard Analysis 2: Clipping Population

```
# Population and Hazard Analysis 2: Clipping Population
# Sanna Mairet

# Description: The WHO AQG and NAAQ standards contour polygons
# created in part 1 are used to clip the population rasters resulting
```

```
# in population rasters of areas corresponding to WHO AQGs and NAAQS.

import arcpy, os, sys, traceback, numpy
from arcpy import env, da
from arcpy.sa import *

arcpy.CheckOutExtension("Spatial")

# Allow script to overwrite existing feature classes
arcpy.env.overwriteOutput = True

# Set workspace path
arcpy.env.workspace = "E:\\italyanalysis.gdb\\"

### These lists are the contours created in popanalysis_part1_contour. These
### contours show the areas derived from the SO2 ppb rasters that fall within
### zones outlined by the WHO SO2 standard and regional standard (regstd)
regstd = 48
whostd = 7.7
contpoly = arcpy.ListFeatureClasses("con7_7rf*", "Polygon")

contpoly_rs = arcpy.ListFeatureClasses("con"+ str(regstd)+"rf*")

poprast = arcpy.ListRasters("projclippop*", "All")


contpoly.sort()
poprast.sort()
contpoly_rs.sort()
print contpoly
print poprast
print contpoly_rs

# Begin clip methodology on WHO standard polygons.
ipop = 0

for pop in poprast:
   rpop =Raster(pop)
```

```
    # This matches the single population raster for each year to the
    # three polygon rasters for each year (differ by PBL)
    for ipoly in range(ipop*3, ipop*3+3):
        ppoly = contpoly[ipoly]

        # popzone is the output of the clipping. It is the population that
        # falls within the specified range of SO2 ppb
        popzone = arcpy.env.workspace + "\\Z" + str(ppoly)

        # The raster clip tool requires an "envelope" of corner
        # coordinates. The describe tool can retrieve the extent
        # coordinates and creates a variable
        desc = arcpy.Describe(ppoly)
        xmin = desc.extent.XMin
        xmax = desc.extent.XMax
        ymin = desc.extent.YMin
        ymax = desc.extent.YMax
        ppolyextentst = " ".join([str(e) for e in [xmin,ymin,xmax,ymax]])

        #This is the tool that clips a yearly population raster three
        # times for each PBL polygon
        arcpy.Clip_management(rpop, ppolyextentst, popzone, ppoly, "#",
"ClippingGeometry", "NO_MAINTAIN_EXTENT")
        rpopz = Raster(popzone)

        # To get the sum of the population raster within each SO2 zone, we need
        # to first convert it to an array. The array will contain 0's as NoData numbers,
        # so next it will be flattened and then the 0's eliminated. From here, stats can
be printed.
        arrpopzone = arcpy.RasterToNumPyArray(popzone)
        flarrpz = arrpopzone.flatten()
        nze = [elem for elem in flarrpz if elem > 0]
        print numpy.sum(nze)


    ipop +=1

# This begins the same process above but for the polygons corresponding
# to regional standards instead of WHO standard of 7.7.
```

```
print "beginning on country standard process"
for jpoly in contpoly_rs:

    # This grabs the year in the name of the polygon
    year = str(jpoly[11:15])
    print year

    for jpop in poprast:

        # This matches the year to the year in the population raster
        if  year in jpop:
            rpop = Raster(jpop)
            print jpop

            popzonej = arcpy.env.workspace + "\\Z" + str(jpoly)

            # Extent coordinates are needed for the clip tool
            desc = arcpy.Describe(jpoly)
            xmin = desc.extent.XMin
            xmax = desc.extent.XMax
            ymin = desc.extent.YMin
            ymax = desc.extent.YMax
            jpolyextentst = " ".join([str(e) for e in [xmin,ymin,xmax,ymax]])

            arcpy.Clip_management(rpop, jpolyextentst, popzonej, jpoly, "#",
"ClippingGeometry", "NO_MAINTAIN_EXTENT")
            rpopzj = Raster(popzonej)

            arrpopzonej = arcpy.RasterToNumPyArray(popzonej)
            flarrpzj = arrpopzonej.flatten()
            nzej = [elem for elem in flarrpzj if elem > 0]
            print numpy.sum(nzej)


    else:
        print "no further polygons corresponding to regional standards"
```

A.4 Population-weighted Averages

```
# Population-Weighted Averages
# Sanna Mairet

# Description: This code calculates population-weighted averages and other
#statistics on population data and ppb SO2 data within a threshold footprint.

import arcpy, os, sys, traceback, numpy
from arcpy import env, da
from arcpy.sa import *

arcpy.CheckOutExtension("Spatial")

# Allow script to overwrite existing feature classes
arcpy.env.overwriteOutput = True

# Set workspace path to your geodatabase
arcpy.env.workspace = "E://VanuatuAnalysis.gdb"

# List variables like the population raster, parts-per-billion raster and
# sulfur dioxide rasters
poprast = arcpy.ListRasters("projclippop*", "All")
ppbrast = arcpy.ListRasters("PPB*" ,"All")
so2rast = arcpy.ListRasters("*corrected","All")

# Sort these rasters so years will automatically line up
poprast.sort()
ppbrast.sort()
so2rast.sort()

# This snippet creates a new layer to filter out areas within rasters that are below
# the threshold of 0.15 DU.
for so2 in so2rast:
    Rso2 = Raster(so2)
```

```
    aboveTHSO2 = Rso2 > 0.15
    aboveTHSO2save = arcpy.env.workspace + "\\aboveTH_" + so2[4:8] +
"_DUC"
    aboveTHSO2.save(aboveTHSO2save)

# Create new variable from these new threshold rasters
aboveTHrast = arcpy.ListRasters("aboveTH*", "All")
aboveTHrast.sort()

# Begin for loop for data
if len(aboveTHrast) == len(poprast):
    ipop = 0
    for pop in poprast:
        rpop = Raster(pop)
        iabTHrast = aboveTHrast[ipop]
        rabth = Raster(iabTHrast)

        # population rasters are multiplied by Boolean raster to eliminate
        #areas below threshold.
        popabth = rpop * rabth
        popaboveTH = arcpy.env.workspace + "\\popaboveTH" + iabTHrast[8:12]
        popabth.save(popaboveTH)

        Rpopabth = Raster(popaboveTH)
        arrRpopabth = arcpy.RasterToNumPyArray(Rpopabth)
        flatrpopabth = arrRpopabth.flatten()
        nzrpopabth = [elem for elem in flatrpopabth if elem >0]


        # This matches the single population raster for each year to the
        # three ppb rasters for each year (differ by PBL)
        for ippb in range(ipop*3, ipop*3+3):

            rppb = ppbrast[ippb]
            Rrppb = Raster(rppb)

            abthPPB = rabth * Rrppb
            aboveTHppb = arcpy.env.workspace +"\\AbTh" + rppb
            abthPPB.save(aboveTHppb)
```

```
        temp_numerator = Rpopabth * Rrppb
        sumrast = arcpy.env.workspace + "\\TA_abTH_popX" + rppb
        temp_numerator.save(sumrast)
        rsum = Raster(sumrast)

        # To obtain statistics, the raster must be converted to an array
        # and the flattened to eliminate NoData zeroes.
        arrrsum = arcpy.RasterToNumPyArray(rsum)
        flatrsum = arrrsum.flatten()
        nzsum = [elem for elem in flatrsum if elem > 0]

        # Numerator and denominator refer to equation from Li et al. (2017)
        numerator = numpy.sum(nzsum)
        denominator = numpy.sum(nzrpopabth)
        popweightedavg = numerator / denominator


        arrrppb = arcpy.RasterToNumPyArray(aboveTHppb)
        flatrppb = arrrppb.flatten()
        nzrppb = [elem for elem in flatrppb if elem > 0]

        # Print statement easily converted into a CSV
        print rppb[3:] , popweightedavg, "%12.3f" % numerator, "%12.3f"
%denominator, "%12.3f" % numpy.sum(nzrppb), "%12.3f" %
numpy.shape(nzrppb), "%12.3f" % numpy.shape(nzrpopabth)

    ipop +=1
```

A.5 Map Automation

```
# Map Automation
# Sanna Mairet


# Description: This code is to automate map production. For this code to work,
```

```
# a map document (mxd) must be created with all layers check OFF in the TOC
# and a layout created. A layer file must be created from the layer with
# the largest range of values with all symbology finalized.

import arcpy, os, sys, traceback, numpy
from arcpy import env, da
from arcpy.conversion import PolygonToRaster
from arcpy.sa import *

arcpy.CheckOutExtension("Spatial")

#Allow script to overwrite existing feature classes
arcpy.env.overwriteOutput = True

#Set workspace path
#arcpy.env.workspace =("E:\\DRCongoAnalysis.gdb\\")
#mxd is map document
mxd = arcpy.mapping.MapDocument ("E:\\drcongomappinghaz.mxd")
#df is data frame in map document
df = arcpy.mapping.ListDataFrames (mxd, "Layers") [0]
#layer file holds the layout and symbology info- no actual data
lyrpop = arcpy.mapping.Layer ("E:\\layerfiles\\hazard\\drcongo\\pop.lyr")
lyrwho =
arcpy.mapping.Layer("E:\\layerfiles\\hazard\\drcongo\\conwhostd.lyr")
lyrreg =
arcpy.mapping.Layer("E:\\layerfiles\\hazard\\drcongo\\conregstd.lyr")

poplayers = arcpy.mapping.ListLayers(mxd, "rfproj*")
wholayers = arcpy.mapping.ListLayers(mxd, "con7*")
reglayers = arcpy.mapping.ListLayers(mxd, "con14*")
print reglayers
print wholayers
print poplayers

i = 0
for who in wholayers:
   descwho = arcpy.Describe(who)
   namewho = descwho.name
   yearwhomatch = descwho.name[11:15]
```

```
    pblwhomatch = descwho.name[11:21]
    print yearwhomatch, namewho, pblwhomatch
    print "in who loop"

    for pop in poplayers:
        descpop = arcpy.Describe(pop)
        namepop = descpop.name
        year = descpop.name[13:17]
        #str(pop[11:15])
        print year, namepop
        print "in pop loop"

        if year == yearwhomatch:
            print namewho, namepop, "year matched in pop loop"

            for reg in reglayers:
                descreg = arcpy.Describe(reg)
                namereg = descreg.name
                pblregmatch = descreg.name[11:21]
                print "in reg loop"

                if pblregmatch == pblwhomatch:
                    print pblregmatch, pblwhomatch
                    print "reglayers in play"
                    pop.visible = True
                    who.visible = True
                    reg.visible = True

                    arcpy.mapping.UpdateLayer(df,pop,lyrpop,True)
                    arcpy.mapping.UpdateLayer(df, who, lyrwho, True)
                    arcpy.mapping.UpdateLayer(df, reg, lyrreg, True)

                    #change title of map
                    MapTitle = arcpy.mapping.ListLayoutElements(mxd,
"TEXT_ELEMENT","title")[0]
                    MapTitle.text = "Sulfur dioxide gas Hazard map for " + year +"
Using PBL of "+ namewho[19:]
                    print "new title is ", MapTitle.text
```

```
                #export to png- check out further for ways to customize
                out_png = "E:\\SO2hazardmaps\\drcongo\\" + namewho +
".png"
                arcpy.mapping.ExportToPNG(mxd, out_png, "PAGE_LAYOUT")
                print 'exported '
                pop.visible = False
                who.visible = False
                reg.visible = False
                i += 1
                break

            else:
                print pblregmatch,"doesn't match ", pblwhomatch,": reglayers not
in play"
                print pblregmatch
                pop.visible = True
                who.visible = True

                arcpy.mapping.UpdateLayer(df,pop,lyrpop,True)
                arcpy.mapping.UpdateLayer(df, who, lyrwho, True)

                #change title of map
                MapTitle = arcpy.mapping.ListLayoutElements(mxd,
"TEXT_ELEMENT","title")[0]
                MapTitle.text = "Sulfur dioxide gas Hazard map for " + year +"
Using PBL of "+ namewho[19:]
                print "new title is ", MapTitle.text

                out_png = "E:\\SO2hazardmaps\\drcongo\\" + namewho +
".png"
                arcpy.mapping.ExportToPNG(mxd, out_png, "PAGE_LAYOUT")
                print 'exported '
                pop.visible = False
                who.visible = False
        else:
            print "pop and who don't match"
            print namewho," / ", namepop

    i += 1
```

Appendix B Map Products

All map products can be found in the supplemental materials attached with this thesis in ProQuest. Map products are also available by contacting the author. There are 12 global maps for each year of data from 2005 to 2016 displaying uncorrected SO_2 VCDs in Dobson units with each region displayed in an inset map. There are 84 regional maps of corrected SO VCDs in Dobson units consisting of 12 maps for each of the 7 regions of interest corresponding to each year of data. There are 252 maps displaying regional SO_2 concentrations in parts-per-billion: 3 for each of the planetary boundary layer conditions for each of the 12 years of data for all 7 regions. Similarly, there are 252 SO_2 hazard maps displaying population count per area within the hazard zones bound by the WHO's AQG and regional NAAQS.

Appendix C Temporal Analysis Figures

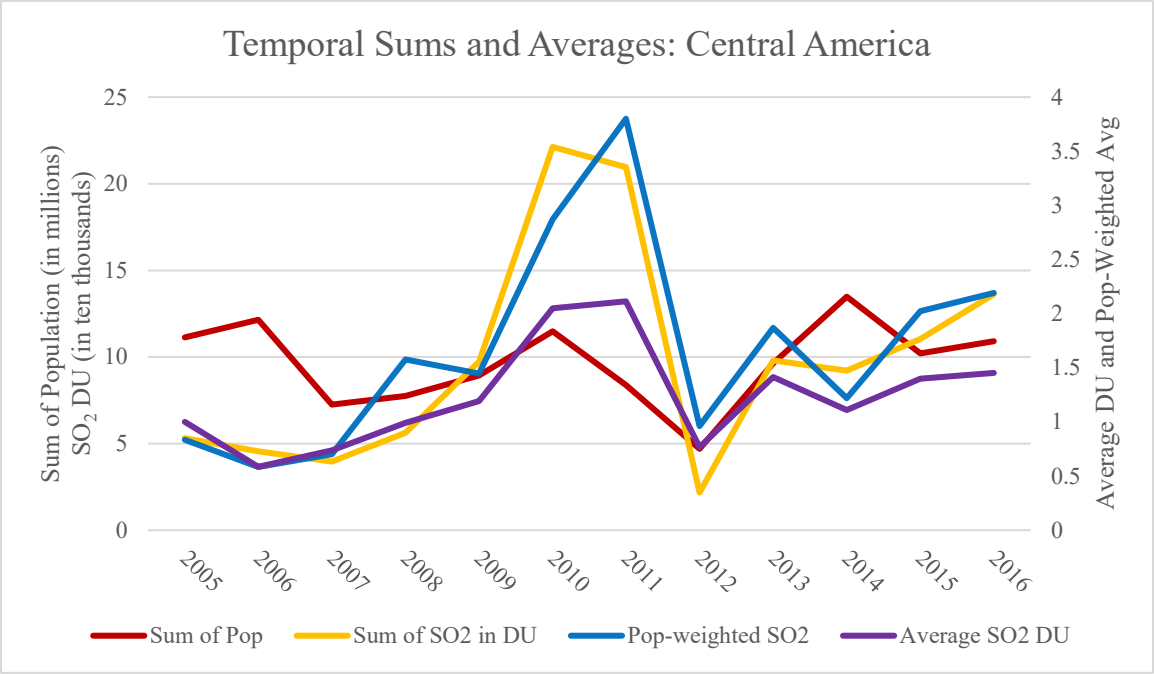

Figure A10: Temporal trends of estimated sum of population, sum of corrected SO_2 VCD (in DU), average corrected SO_2 VCD (in DU), and population-weighted average of corrected SO_2 VCD (in DU) for Central America. Population is shown in millions. SO_2 in DU column amounts are shown in ten thousands.

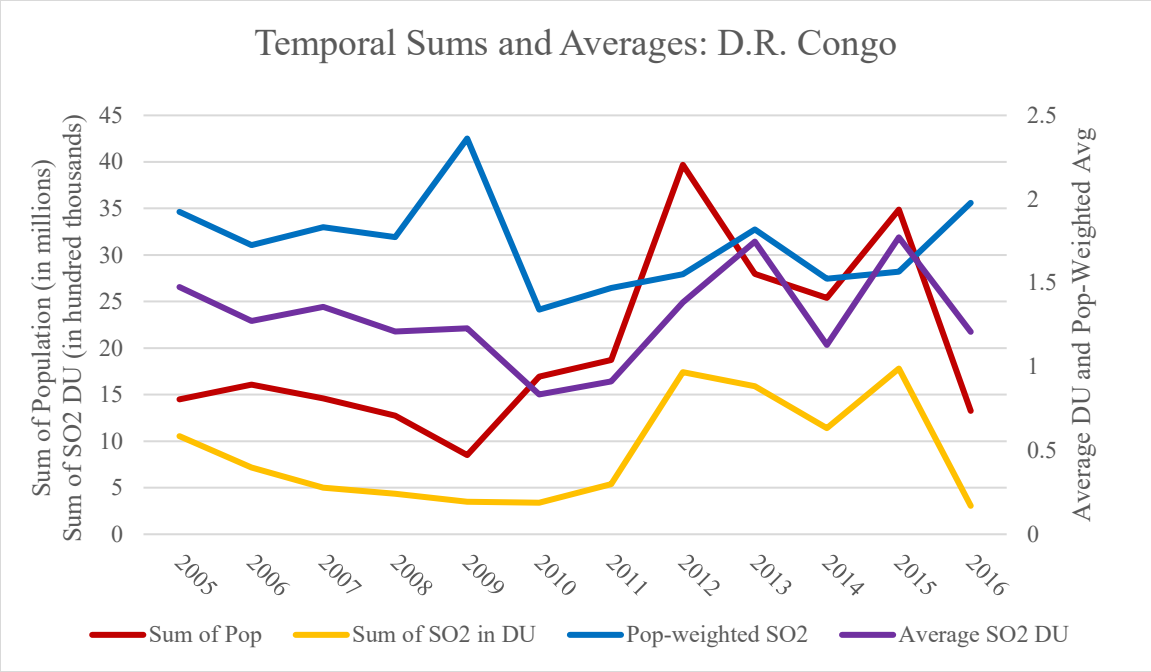

Figure A11: Temporal trends of estimated sum of population, sum of corrected SO_2 VCD (in DU), average corrected SO_2 VCD (in DU), and population-weighted average of corrected SO_2 VCD (in DU) for the Democratic Republic of the Congo. Population is shown in millions. SO_2 in DU column amounts are shown in hundred thousands.

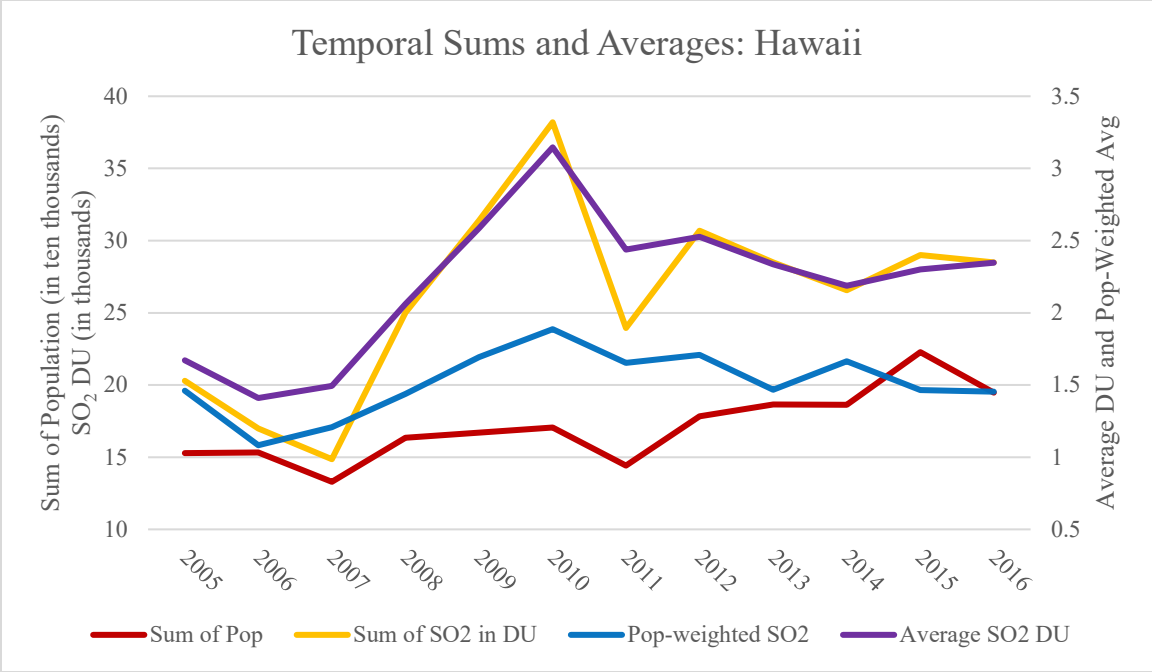

Figure A12: Temporal trends of estimated sum of population, sum of corrected SO_2 VCD (in DU), average corrected SO_2 VCD (in DU), and population-weighted average of corrected SO_2 VCD (in DU) for Hawaii. Population is shown in ten thousands. SO_2 in DU column amounts are shown in thousands.

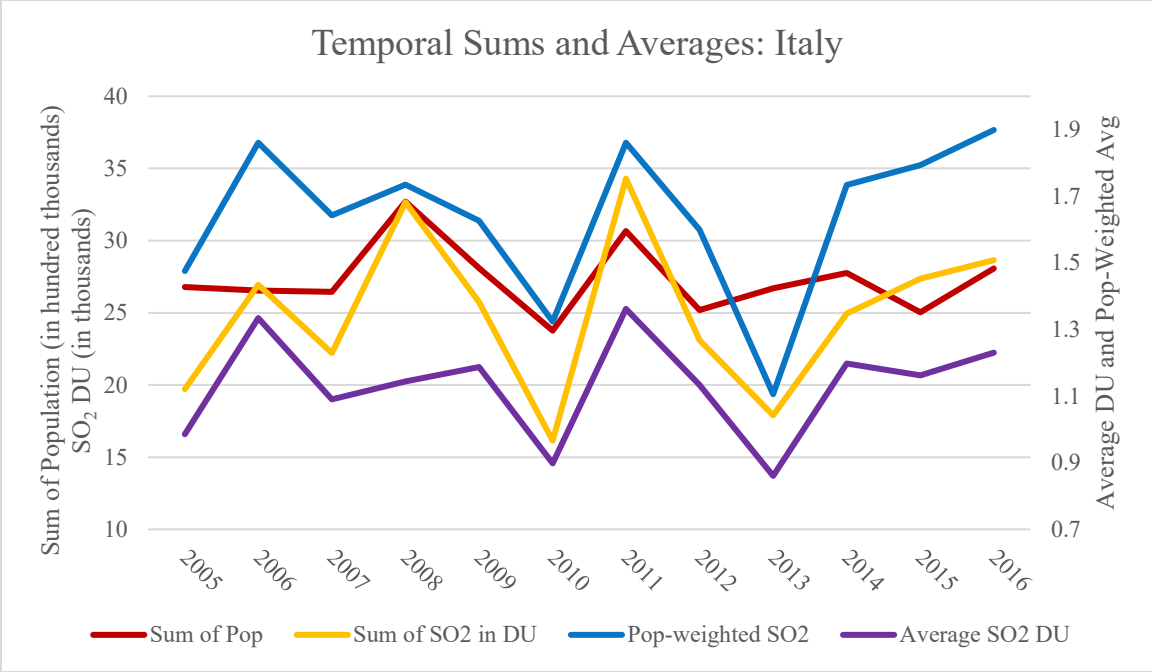

Figure A13: Temporal trends of estimated sum of population, sum of corrected SO_2 VCD (in DU), average corrected SO_2 VCD (in DU), and population-weighted average of corrected SO_2 VCD (in DU) for Italy. Population is shown in hundred thousands. SO_2 in DU column amounts are shown in thousands.

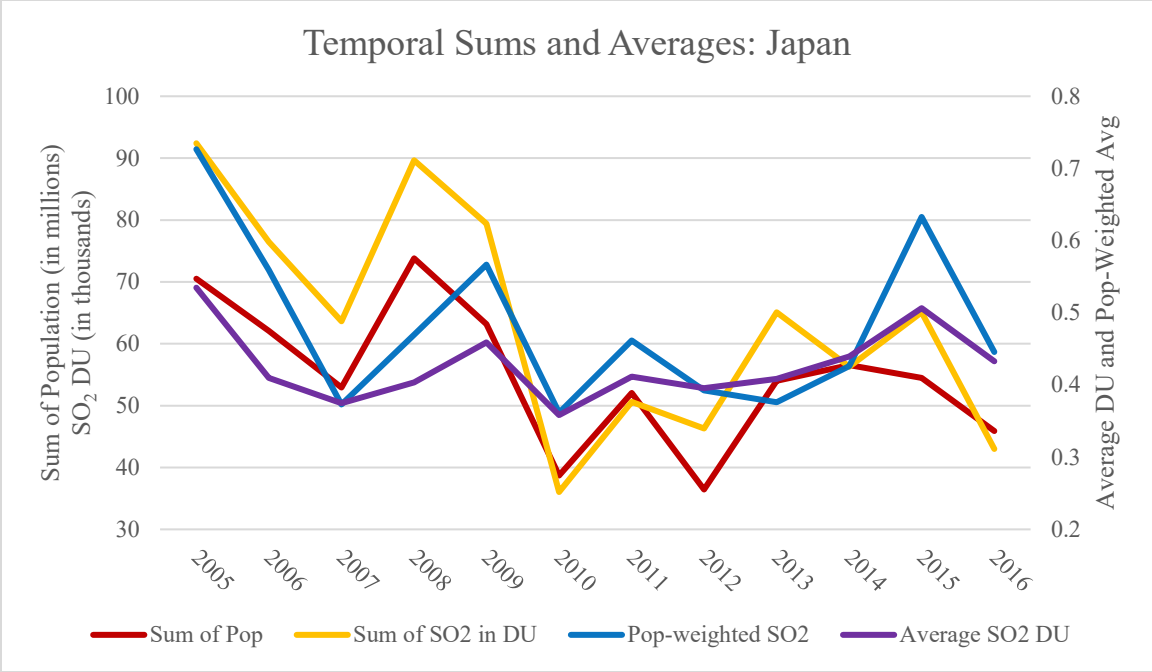

Figure A14: Temporal trends of estimated sum of population, sum of corrected SO_2 VCD (in DU), average corrected SO_2 VCD (in DU), and population-weighted average of corrected SO_2 VCD (in DU) for Japan. Population is shown in millions. SO_2 in DU column amounts are shown in thousands.

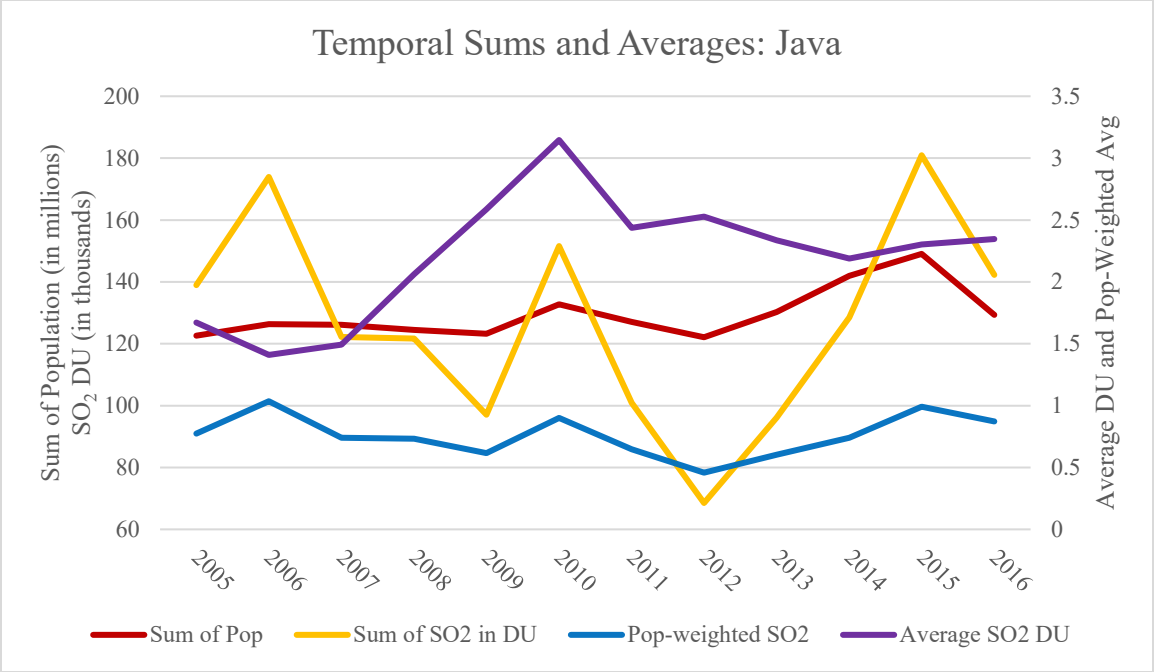

Figure A15: Temporal trends of estimated sum of population, sum of corrected SO_2 VCD (in DU), average corrected SO_2 VCD (in DU), and population-weighted average of corrected SO_2 VCD (in DU) for Java, Indonesia. Population is shown in millions. SO_2 in DU column amounts are shown in thousands.

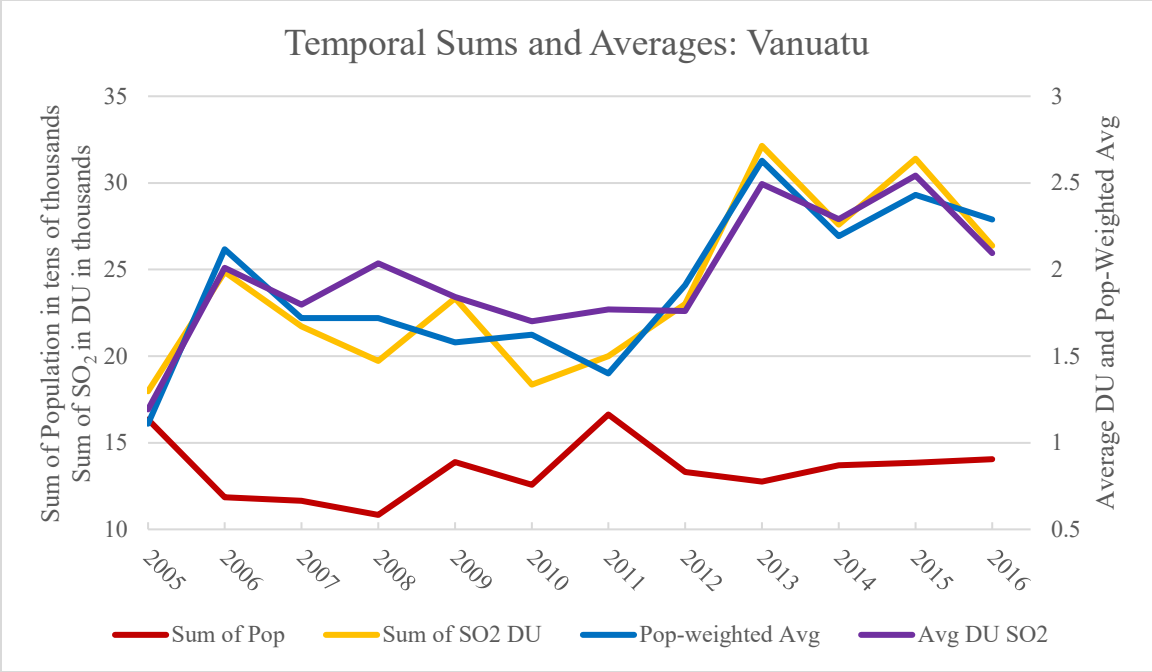

Figure A16: Temporal trends of estimated sum of population, sum of corrected SO_2 VCD (in DU), average corrected SO_2 VCD (in DU), and population-weighted average of corrected SO_2 VCD (in DU) for Vanuatu. Population is shown in tens of thousands. SO_2 in DU column amounts are shown in thousands. (Note: this figure is identical to Figure 7 in the Results Section.)