Small Cloud and Cloud Shadow Removal for Image Enhancement and Analysis of Kilauea Volcano Eruption in Hawaiian

GISC – 4360K (Digital Image Processing)
Final Report

Prepared by:
Timothy Davis

Submitted to:
Dr. H. Cho

UNG | UNIVERSITY of NORTH GEORGIA™

Date: 4/25/2019

Abstract:

This project was developed to obtain a method by which we can remove or reduce clouds and unwanted noise from our imagery for analysis. This was done with the use of various tools in Arc Pro, such as but not limited to, the Resample Tool, Segmentation Tool, Reclassify Tool, Raster Calculator, and the Select Layer by Attribute tool. With the various types of classification, the idea is to find the one that works the best and to better understand what the other tools can do and how they might be useful for other types of data analysis and/or projects. In the final analysis I used the Raster Calculator to do most of my analysis. One of the first steps that were taken, was to change the background value from zero to NoData, so that it would not influence my analysis. I also used Raster Calculator to create a new raster with a gridcode of 1 or 0, with one being cloud and zero being everything else. I then used the Select Layer by Attribute to select the cloud polygons and delete them out of each band, resulting in a raster for each band with holes where the clouds were. I then composited the bands and mosaicked it with the data from the LandSat 8 imagery from the USGS website. With the entire array of Arc Pro's brand-new tool set there is a whole new horizon to geospatial data analysis that are now available.

Introduction:

In the summer of 2018, the large island of Hawaii experienced a large volcanic eruption on the Kilauea site. This eruption was very hard to analysis immediately after due to the amount of cloud cover

that was a result of the increased temperature on the island. In September of 2018 a satellite imagery company called Digital Globe captured a phenomenal image of the devastation left after the eruption. This data was donated to the Hawaiian Volcano Observatory in Hawaii for analysis and observation. Although the image is not cloudless, it is quite close to and with a stunning 0.8m pixel resolution it is an amazing piece of data to analyze (See *Figure 1).*
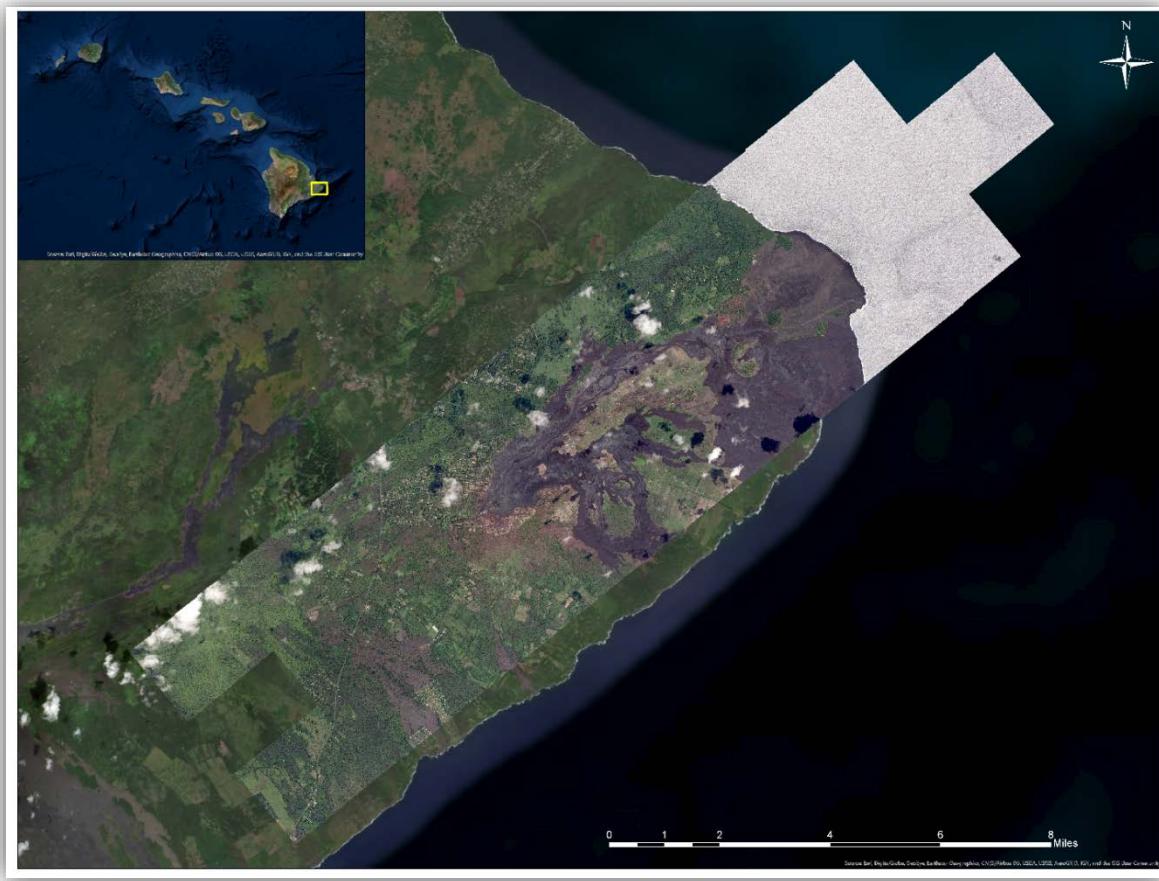
Figure 1: Digital Globe Imagery over Hawaii September 09/2018 overlaying ESRI Imagery Base Map

This project has been a profound training curve for me as we work through how to develop both cloud detection and removal, along with smoothing and edge detection, and noise removal on satellite imagery. Going forward, my hope is that this project will be useful for larger data sets and analysis of both satellite and aerial imagery. This research work is open source and is available to all for education, research, and analysis purposes. My hope is that this research will continue and that this data will be readily available to further and enhance future analysis in this area.

Literature Review:
Haze Detection and Removal in Remotely Sensed Multispectral Imagery
By: Aliaksei Makarau, Rudolf Richter, Rupert Müller, and Peter Reinartz

In this article Aliaksei Makarau, Rudolf Richter, Rupert Müller, and Peter Reinartz, (2014) try to determine the best way to deal with haze in "LandSat 8 OLI and WorldView-2" imagery. There are many things to consider when trying to detect and remove this kind of noise from imagery. One thing they look at is building a "Haze Thickness Map (HTM)" which for both hyper- and multi-spectral imagery the kinks have not all been worked out. They also looked at another method that is commonly used which is the "Haze Optimization Transform (HOT)" which is like a form of object-based analysis. And as a second half to the HOT there is the "Dark-Object Subtraction (DOS) that is used to remove shadows so that the indices of the vegetation can be calculated. They also touch on a method that they call "Background Suppressed Haze Thickness Index (BSHTI)" that does something like the HTM that we motioned earlier. There are also some details on using a "Digital Number (DN)" manipulation to de haze images as well. But in this article, they expound on the DOS method that was improved on by Pat S. Chavez, JR. (1988). But instead of trying to find one or two in a defined part of the image, they attempt to have it scan the entire image and apply this analysis to the hole locale image. With DOS method and the HTM together they brought a brand-new set of procedures to detect and remove haze. Their biggest win was that when they compared their dehazed image to a clear image of the same location from a different time period, the images were spectrally the same, meaning that the transformation had not distorted the DN values of the raster in the process. And in their conclusion, they talk about looking into manipulating "border effects" as a future audition to this method for a more comprehensive method of digital image processing and analysis Aliaksei Makarau, Rudolf Richter, Rupert Müller, and Peter Reinartz, (2014).

Objectives:

Clouded imagery is close to useless for geospatial analysis. Our area of study is the Kilauea volcano eruption site on the largest island of Hawaii that occurred in the summer of 2018. This project was developed for the express purpose of enhancing and eliminating small clouds and shadows from the satellite imagery by performing a pixel-based classification of the imagery and extracting the cloud and shadow pixels and overlaying them with cloudless pixels from another image of the same spatial location. This would allow for a replacement of missing data where the clouds are with visible data that would be collected on an earlier or later date.

Materials and Methods:

So, despite all the amazing aspects of this imagery there are a few things that needed to be addressed for this data to be useful for analysis. First was that awesome resolution that became a problem in a hurry, as it took hours to draw and redraw every time you zoomed in or out in Arc Pro Desktop, process, classify, and analysis. To solve this problem, I just resampled the cell size from 0.8m to 5m resolution, and this drastically decreased the file size from 2.77 GB to 72.73MB allowing for a much smoother work flow. The first thing I tried was to do a Segmentation, this tool is in Arc Pro and can be found in the "Imagery" tab in "Image Classification" under "Classification Tools".
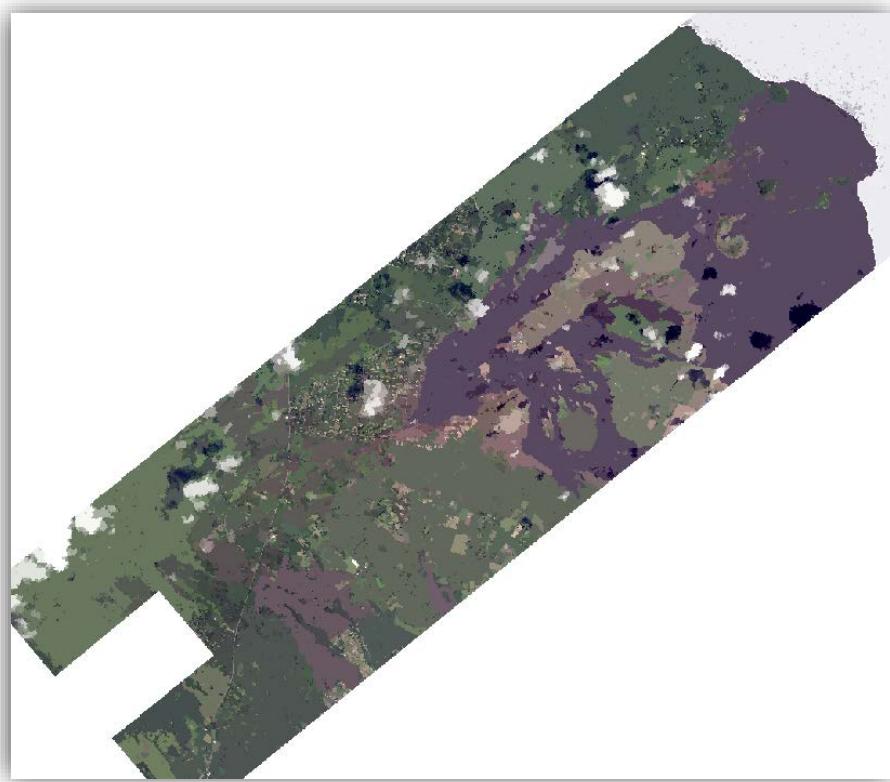
Figure 2: Segmented Digital Globe Imagery

Segmentation is essentially an object-based classification that groups pixels together based on color and/or grouping patterns (*See Figure 2*). There are three different settings that can be used to manipulate the hierarchical importance of the data in the object-based classification. First, for "Spectral Detail" or resolution, you can select any value between one and twenty, one being a very high range of cell values returning very few classifications, and twenty being a very low or small range of cell values returning many classifications. So, the higher the "Spectral Detail" more classes will be created, and the opposite holds true as well according to the Segmentation—ArcGIS Pro | ArcGIS Desktop (2019) website. For this project I used a "Spectral Detail" setting of 15.5 and it worked great in identifying the shadows and clouds. The second setting that can be manipulated is the "Spatial Detail" or resolution, which also has a similar range of values to pick from, of one to twenty. In this case the smaller the number picked will return a larger cluster of pixels, in other word, if you pick a low value similar pixels will be grouped together into one class and if you pick a higher value only similar pixels that are spatially close together will be grouped into one class as stated on the Segmentation—ArcGIS Pro | ArcGIS Desktop (2019) website. For our project we set it to 15, and it did a great job. And the last setting that can be manipulated is the "Minimum Segment Size". This has to do with the smallest linear mapping units in pixels, in other words if a class is smaller than a single linear mapping unit, then it will be assigned to the nearest neighbor class that is the most like it as noted on the Segmentation—ArcGIS Pro | ArcGIS Desktop (2019) website. For this setting in our analysis we used 20 pixels for our minimum segment size.

The resulting image with segmentation was not useful for our study although it greatly helped in understanding what type of analysis was needed to be done for this image analysis and for cloud removal. It was understanding how segmentation takes cell values and applies a classification to each

cell based on its values and its neighboring values that helped us to see we needed to analyze this image on a band-by-band biases for better results.

First, we needed to change the background value of zero to "NoData" so that it would not affect our analysis. With RGB set to zero the background is black where as if it is set to "NoData" it will be masked or removed as seen in *Figure 3.* This was done for the individual bands that we had for this image of Red, Green, and Blue with the "Raster Calculator" tool.
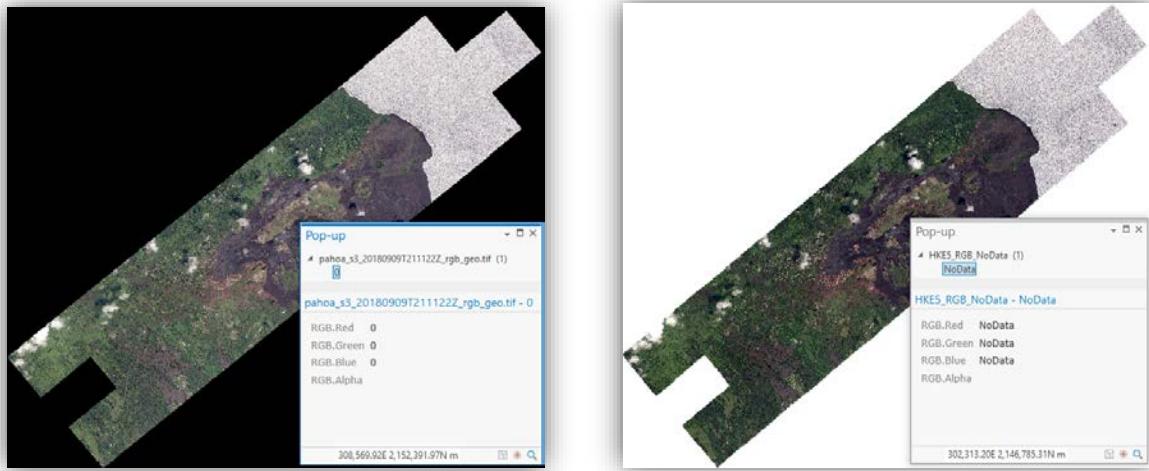
Figure 3: Digital Globe Imagery with background set at zero and with background set to NoData

We then used the "Raster Calculator" to create a new raster with a "gridcode" of either 1 or 0 where 1 is clouds and 0 being everything else. To determine what was cloud I took ten samples of cloud pixels and got their maximum and minimum for each band and put them in the Raster Calculator in the following way. If Red is less than or equal to (<=) Red Max and if Red is greater than or equal to (>=) Red Min, then assign gridcode a value of 1, otherwise assign the value 0. This is the code line that I used in the Raster Calculator: ("R" <= R_Max and "R" >= R_Min, and "G" <= G_Max and "G" >= G_Min, and "B" <= B_Max and "B" >= B_Min, 1, 0) With this code we are only looking for pixels that intersection of these values, so for example if a pixel has an "R" value in the range and a "G" or "B" value outside the range it will not be selected as cloud. I have named this analysis procedure Multi Band Cell Value Intersection Range (MBCVIR) for object base selection. The best part of this is that though clouds do not have well defined edges as Lixin Sun, Rasim Latifovic and Darren Pouliot, (2017) noted in their attempt to detect and remove haze, the automated nature of it allows for a fast but imperfect selection. This imperfection in cloud detection can and is a tricky thing to discern as Keith D. Hutchison, (1999) noted that spectral reflectance form clouds can vary depending on the thickness and moisture density of the cloud. Be that as it may we tried to select a range that most closely resembled cloud for this part of the analysis.

With this new raster I used the "Select Layer by Attribute" tool to manually remove noise that is not cloud such as buildings and other lighter colored objects by selecting them and reassigning their gridcode value. First, I add a clause that selects rows if their gridcode equals one and if their shape area is less than 300 square meters. Then I manually looked over the raster and selected any polygons that might have been set as cloud but are not, and then I changed my selection to select only the rows that have a gridcode value of 1, and then I deleted them. The last step in this part of the analysis was to take the raster with the cloud polygons deleted out and to run the "Dissolve " tool on it to make it one large polygon mask with cloud holes in it.

Figure 4: The Mosaic Operator set to First, the Mosaic Colormap Mode set to Match, and the Color Matching Method set to Match Histogram.

Then with the dissolved polygon that had the cloud holes in it I ran the "Extract by Mask" tool on each band to cut out the clouds and the result, three bands that had NoData spots where the clouds were. I then ran the "Composite Bands" tool to join the three bands together and form the new RGB image without any clouds.

Now to patch the holes in our image we turned to LandSat 8 30m spatial resolution imagery from the USGS Earth Explorer website, Survey, U. (2019). This data is free to members and comes in a tar.gz file format to minimize file size as much as possible to facilitate download and upload. Once the file was twice unzipped, we needed only three of seven bands available which in LandSat 8 RGB are bands 4,3,2 as stated by Vermote, E., Justice, C., Claverie, M., & Franch, B. (2016). These bands also had varying grayscale that needed to be addressed for color rand to match that of the Digital Globe Image. To fix this I used a tool developed by Dr. H. Cho called "Rescale Gray Levels" that takes a raster input and rescales it to the user defined gray level, in my case I set it to 256 with a range of [0-255]. I then took the LandSat 8 Imagery tile and clipped it to the Digital Globe tile that had as the background value NoData so that the clip would follow the image and not the tile outline.

Then I took the new clipped LandSat 8 Image and the new composited Digital Globe image without and clouds and ran the "Mosaic" tool to join them together. This tool has many settings and there are two different versions of the "Mosaic" tool in Arc GIS Pro that can be used.

Results & Conclusion:

In my first attempt to use the segmentation tool for this project I found that it worked very well in identifying the clouds with a form of object-based analysis, but I couldn't figure out how to use the segmented raster for cloud removal.
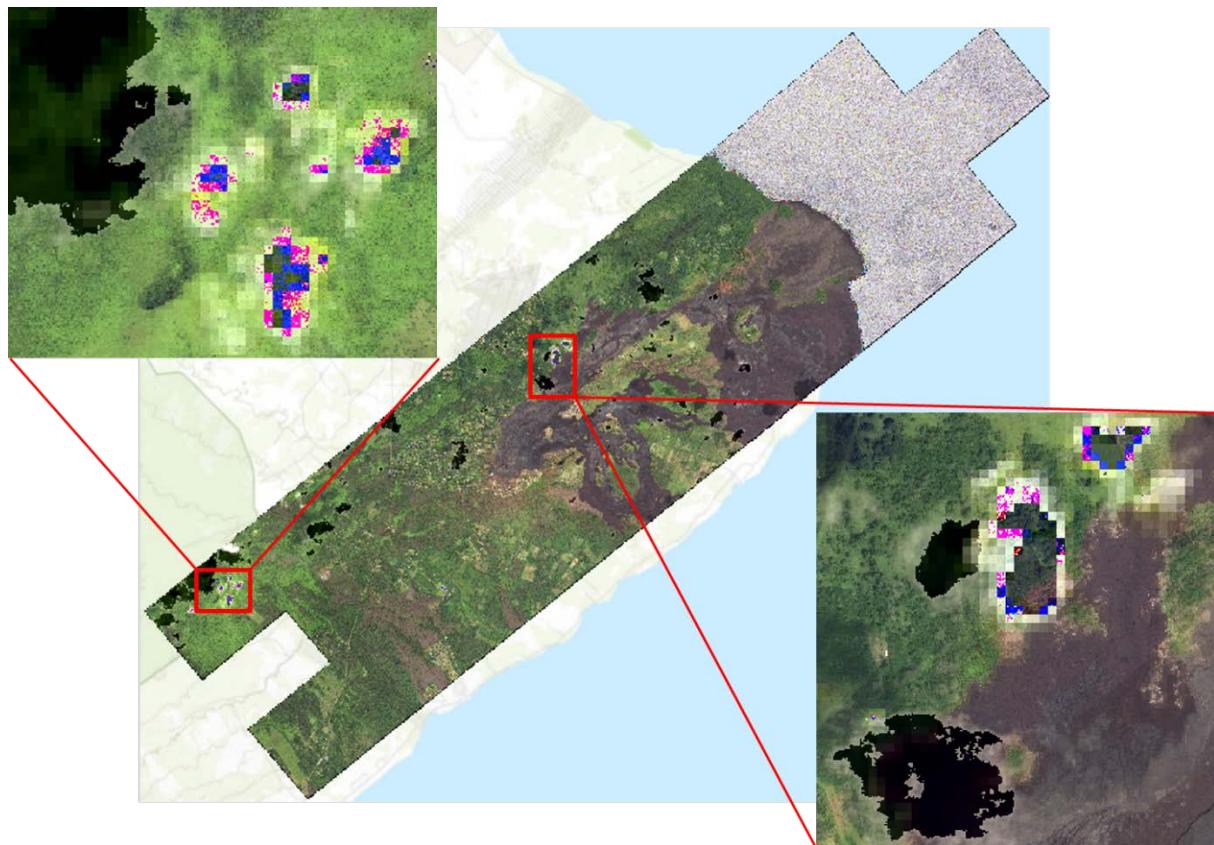

Figure 5: This was produced using the Mosaic To Raster with the Cell Size set to 5, the Mosaic Operator set to Sum, and the Mosaic Colormap Mode set to Match

This analysis is not feasible with different data sets and data collected with different spatial, spectral, and temporal resolution as seen in *Figure 4 and 5*. It may be useful if you have two datasets that were collected around the same time and with the same sensor, that this analysis would be a viable method for removing small clouds. It has also proven to be difficult and time consuming and as a result I hope that in the future there will be a better and more automated version of this tool that will do most of the tedious work through automation. I also had a lot of trouble with smaller clouds and thinner and haze like clouds that were around the larger and denser cloud mass. These smaller clouds and haze are according to Bo-Cai Gao, and Rong-Rong Li, (2017) removed with what they describe as the "empirical technique" which they describe in detail in their article. These methods are very useful, but they require very in-depth knowledge on the subject and extensive research on each topic for one to even understand what they are writing about.

As an extension to this project there would need to be some form of cloud shadow detection and removal that would accompany the cloud detection and removal for a complete analysis and image enhancement. I did not attempt it due to time restraints and deadlines that need to be met for this project, and as I have explored the different hurtles to overcome, I found that M. Joseph Hughes, and Daniel J. Hayes, (2014) also showed the complexity of distinguishing cloud shadow and shadows resulting from hill shade and land based objects that cast shadows. Also, darker objects such as water bodies and or valleys may have a similar reflectance as that of cloud shadows. So perhaps in the future and with more time this project will be more fully expounded and completed to better serve the needs for research and analysis.

Work Cited:

Sun, L., Latifovic, R., & Polios, D. (2017). Haze removal based on a fully automated and improved haze optimized transformation for Landsat imagery over land. *Remote Sensing*, *9*(10), 972.

Hughes, M., & Hayes, D. (2014). Automated detection of cloud and cloud shadow in single-date Landsat imagery using neural networks and spatial post-processing. *Remote Sensing*, *6*(6), 4907-4926.

Hutchison, K. D. (1999). Application of AVHRR/3 imagery for the improved detection of thin cirrus clouds and specification of cloud-top phase. *Journal of Atmospheric and Oceanic Technology*, *16*(12), 1885-1899.

Chavez Jr, P. S. (1988). An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. Remote sensing of environment, 24(3), 459-479.

Makarau, A., Richter, R., Müller, R., & Reinartz, P. (2014). Haze detection and removal in remotely sensed multispectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(9), 5895-5905.

Gao, B. C., & Li, R. R. (2017). Removal of thin cirrus scattering effects in Landsat 8 OLI images using the cirrus detecting channel. Remote Sensing, 9(8), 834.

Segmentation—ArcGIS Pro | ArcGIS Desktop. (2019). *Pro.arcgis.com*. Retrieved 26 April 2019, from https://pro.arcgis.com/en/pro-app/help/analysis/image-analyst/segmentation.htm

Survey, U. (2019). *EarthExplorer - Home. Earthexplorer.usgs.gov*. Retrieved 26 April 2019, from https://earthexplorer.usgs.gov/

Vermote, E., Justice, C., Claverie, M., & Franch, B. (2016). Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sensing of Environment*, *185*, 46-56.