Investigating the fractal dimension of clouds to minimize uncertainty in climate models

Melita Wiles
Department of Physics, The College of Wooster, Wooster, Ohio 44691, USA
(Dated: May 2, 2021)

We use the box counting method to determine the fractal dimension of the boundary of certain clouds. The fractal dimension is essentially a measure of roughness of the shape, which will give us a dimension in between one and two. Being able to determine the fractal dimension of a cloud could be beneficial to climate modelers, because a concern for modelers includes what is the cloud shaped like, since using any Euclidian shape is not sufficient for this type of modeling. A Mathematica code was created to find the fractal dimension of the boundary of cloud images. The code works by inputting an image of a cloud and the image was overlaid with a grid of squares and scanned through a loop to determine how many squares covered the boundary of the cloud. This was done many times, with many different sizes of boxes. With these results, we found the dimension of the boundary of the cloud. We compared the fractal dimension to two different parameters of clouds: the thickness of the clouds and the altitude placement of the cloud in the sky. Results show that the thickness of the cloud does not depend on the fractal dimension, but when examining clouds higher and higher up in the sky, the fractal dimension of the cloud increases. This means the roughness of the cloud is lower for higher clouds compared to lower clouds in the sky. Further analysis must be done to confirm these results of cloud classification.

INTRODUCTION

As our earth continues to warm, scientists are working more and more to create realistic and successful climate models. These models include oceanic, atmospheric, and land processes such as melting glaciers and ocean circulation. They are generated from mathematical equations that are created from data recorded that simulate transfers of energy and water. In Fig. 1, we can see the increase in surface temperature from 1990 to what it is predicted to look like from climate models in 2098. From this image we can see that the earth is indeed warming, and has harmful effects to our planet. For example, natural disasters such as hurricanes and drought are becoming more frequent and intense. These events can destroy habitats for animals and cities for people. Some scientists are trying to understand and predict what will happen by using physical system models [1].

Clouds prove to be the greatest uncertainty in climate models. This is due to many aspects of the cloud itself. First, the composition of the cloud can differ, the type of aerosols making up the cloud itself, along with the size and alignment of these particles. Not only does the composition create mystery, but also the type of cloud, its shape, height, and placement in the sky matters as well [2]. Clouds can trap heat and contribute to the heating of our planet, but also can reflect the sunlight and decrease the heating of the planet [3]. It is unknown if the net contribution of clouds to global warming is positive or negative. Researchers are trying to describe these complex systems currently [4].

When investigating this problem, something else to explore is the fractal-like shape of these clouds. Fractals are another type of geometry, different than typical Euclidean geometry. Euclidian geometry characterizes regular objects, while fractal geometry studies the irregular objects. Fractals are rough shapes that can be separated into parts, where each is a reduced-size copy of the whole. This characteristic is called self-similarity. Understanding the fractal dimension and geometry of clouds could be another possible way to model clouds. Fractals could possibly be another metric to classify clouds.

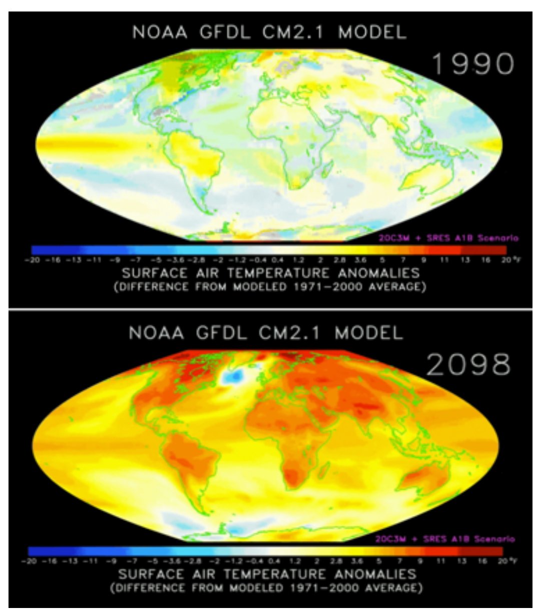

FIG. 1: Surface air temperature over the entire earth in (top) 1990, and predicted surface air temperature in (bottom) 2098 [1].

THEORY & APPROACH

Creating the Code

One way to understand these physical systems is to understand the dimension of clouds. Clouds themselves all have fractional dimension numbers, meaning each cloud's roughness can be described through its fractal dimension. Through Euclidean geometry, we learn about one two and three dimensional objects, but usually in nature lots of objects are in between one to two or two to three dimensions. We call this a fractal (fractional) dimension. For example, a newspaper is two dimensional, and a perfectly compact newspaper ball is three dimensional, but a crumbled up piece of newspaper is in between two and three dimensions. Clouds have the property of 'in between' dimension. Understanding what dimension these clouds are can be beneficial to understanding how these clouds contribute or do not contribute to global warming.

In 1918, Felix Hausdorff introduced the Hausdorff dimension or fractal dimension, which is a measure of roughness [5]. We know that a single point is of dimension zero, a line segment is of dimension one, a circle is of dimension two, and a sphere is dimension three, but as stated before, lots of shapes in nature have dimensions in between integer numbers.

There are multiple ways to find the dimension of these shapes, depending on how simple the shape is. When measuring coastlines, we can see when zooming into the coastline, as the zoom increases, we can see more and more detail. For example, when measuring the coastline of Great Britain, by using different length scales to measure the coastline, the perimeter will grow as we use a more precise measurement [6]. Meaning that since the coastline is not a straight line and there are boulders and rocks in the way, you have to measure each and every in and out where boulders, rocks and pebbles are located. Thus the perimeter increases when using a ruler compared to using a meter stick. The final answer is as you measure more and more closely, you approach infinity. So, we use a fractal dimension to classify and understand this coast line. This is done by measuring the perimeter using different length scales, like a ruler and then a meter stick. We then calculate the dimension by taking the logarithm of the perimeter over the logarithm of the inverse length scale, or magnification, to find the fractal dimension.

For clouds, we essentially use a similar method by creating a program in Mathematica, but break it down a little differently compared to finding the perimeter. In our experiment, because these images are very complex, we use the box counting method [7]. The box counting method is executed by covering the image of interest (here the clouds) with a grid of boxes of a specific area and counting how many boxes cover part of the boundary of the image image. Then this process is repeated by using a finer grid with increasingly smaller boxes. Examples of two-by-two and four-by-four grids can be seen in Fig. 2. Because the size of boxes on the grid is shrinking while repeating this box counting method, we can more accurately capture the structure of the image. Then we can find the dimension by dividing the logarithm of how many boxes included the boundary by the logarithm of the inverse length scale, which in this case is the inverse of the box size. Usually, to understand this number we plot the logarithm of the number of boxes interfering with boundary versus the logarithm of the inverse box size and take the slope as the dimension of the image. For our clouds, the fractal dimensions range from one to two dimensions. A straight line is of dimension one and a super wiggly line that completely fills up a two-dimensional area is two-dimensional. So, a steeper slope closer to two means a more wiggly line and a shallow slope closer to one corresponds to a straighter line [8].

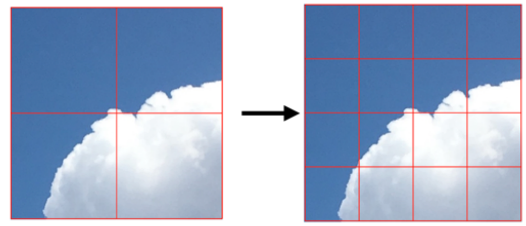

FIG. 2: Example of the box counting algorithm. First image: a two-by-two box grid overtop of a cloud image, all four boxes interfere with the boundary. Then we move to smaller boxes as seen in the second image, where there are 16 boxes and only seven interfere with the boundary.

We calculate the fractal dimension of the boundary of clouds, not the cloud itself. This is because our program is not equipped to analyze the entire three dimensional cloud, so we focus on the boundary of the cloud.

To actually calculate the fractal dimension of the boundary of clouds, we start by inputting an image, like any of the images seen in the first row of Fig. 3. Then we binarize these images so the images are strictly black and white. This makes it easier to apply the box counting method, so the box can recognize where either the back or the white boundary of the cloud. The binarization process for each cloud does differ, because not all cloud images are taken with the same brightness settings, so we have to manually determine what the threshold should be to best identify where the cloud is in the image. After this process, our images look something like what the second row of Fig. 3 looks like [9].

After the binarization process, we create a matrix of zeros and ones, where all of the zeros are the black areas and the ones correspond to the white areas on the image. Each value corresponds to a pixel in the image. Then we create boxes and loop through the rows of the image

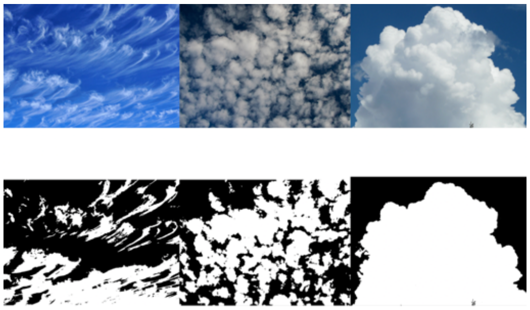

FIG. 3: Three different clouds used in the study. From left to right: cirrus, altocumulus, and cumulus clouds. The top row are the images and the bottom row are the binarized versions of the top row.

to see where the images changes from zeros to ones, or from black to white to outline the boundary of the cloud. For example, when our box size is 2 by 2 pixels, we go through and average the ones and zeros in the 2 by 2 square. If the average is zero or one, that means we have not hit any boundary, but if the average is somewhere between zero and one (1/4, 2/4, or 3/4) we count that box as interacting with the boundary of the image. Then after looping through each row of the image we add up the number of boxes interacting with the boundary of the cloud and take the logarithm of that, which is then graphed versus the logarithm of the inverse box size. The slope of the graph will give us the fractal dimension.

Testing the Code

To test our program and make sure it was running correctly we inputed an image of a fractal with a predetermined fractal dimension. We used the Koch curve seen in Fig. 4, whose boundary has a fractal dimension of 1.26. When running our fractal through our program, we got a fractal dimension of 1.18, which is about a 6% error. This number was decent, but could be improved. In general, as we increase the size of the box, the accuracy of the dimension decreases, so we fit a line of the actual dimension to our plot of points to see what parts of the graph fit the line best, or in other words, what part of the graph was most accurate. This plot can be seen in Fig. 5, where the red line has a slope of 1.26. Because of

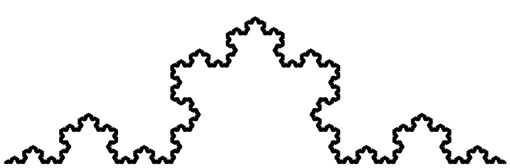

FIG. 4: A computer simulation of the Koch curve at the seventh step of the curve.

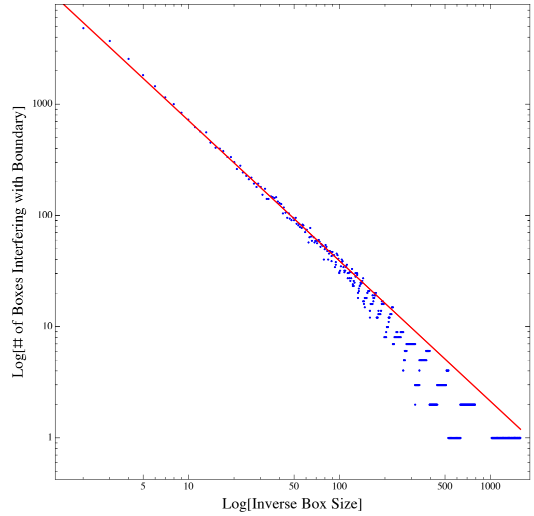

FIG. 5: Logarithm of the number of boxes interfering with boundary plotted against the logarithm of the box size for the Koch curve seen as represented in blue dots and lines. The red line represents the line of best fit that gives a slope of 1.287, which is the computed fractal dimension of the curve. The actual fractal dimension value for this curve is 1.26. Our value is off by approximately 2%.

our findings here, we decided that for further analysis we would only look at points from an inverse box size of 10 to 70. Because we looked more at the middle section of data, our accuracy went from 6% error to 2% error, and we found a fractal dimension of 1.287 for the Koch curve boundary. We made this choice because naturally at the lowest and highest points on the graph, the data tends to trail off because of accuracy issues.

Because of the results from the Koch curve test, data collection could begin. We used cloud images from multiple sources, including the Cloud Appreciation Society [10]. Cloud images were collected in terms of the thickness of the cloud, so we collected images of clouds that were almost transparent, but also thick clouds that were completely opaque. We tested ten thin and ten thick cloud images. With these images we wanted to explore the relation, if any, that these images have with their fractal dimension.

We also collected images of nine categories of clouds, including cirrus, cirrocumulus, altocumulus, which are lower level clouds, altocumulus, nimbostratus, and stratus clouds, which are mid level clouds, and lastly cumulus, cumulonimbus, and stratocumulus clouds, which are higher altitude clouds. We tested two images from each category because of limited resources for images. Again, we wanted to explore the relationship between height of the cloud in the sky to the fractal dimension of each image.

Since there is a relationship between high level clouds being more thinner clouds and lower level clouds being thicker clouds, we wanted to not only find the fractal

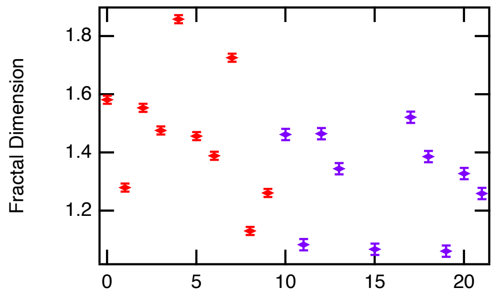

FIG. 6: Fractal dimension versus image number. Purple diamonds (numbered 0-9) are considered thin clouds and red diamonds (numbered 10-20) are considered thick clouds. There is no obvious trend for the fractal dimension here when switching from thin to thicker clouds. Error bars show that the thin clouds (in purple) have a 1.4% error and the thick clouds have a 2% error.

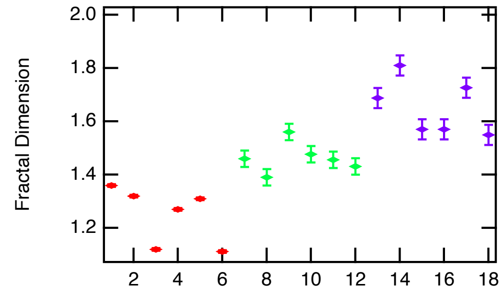

FIG. 7: Fractal dimension versus image number. Red diamonds (numbered 1-6) are low level clouds, green diamonds (numbered 7-12) are mid level clouds, and purple diamonds (numbered 13-18) are high level clouds. As the color changes from red to green to purple, the fractal dimension increases. Error bars show a 4% error for low clouds, a 3% error for mid clouds, and a 0.5% error for high clouds.

dimension of all the clouds listed above, but also examine any trends related to this idea and the fractal dimension. So, for example we wanted to find out if higher level clouds, which are generally thicker clouds would have a lower fractal dimension compared to lower level, thinner clouds.

RESULTS & ANALYSIS

When approaching this problem, we focused on two parameters of clouds besides the fractal dimension to test if there was any relationship. First, we examined the thickness of the cloud and compared it to the fractal dimension of the clouds. These clouds were categorized as thin and thick clouds. Usually, with thinner clouds, they are easier to see through and with thick clouds one is unable to see through the cloud.

The calculated fractal dimension for the different thin and thick clouds are shown in Fig. 6. The first ten clouds plotted are the thin clouds and the thicker clouds are plotted on the x-axis as clouds 10-20. As seen in the plot, there is no immediate relationship or change when looking at thin or thick clouds. The plot has a lot of noise and it is hard to draw any conclusions from this data other than that using the fractal dimension of clouds to understand the difference between thin and thick clouds may not be the best direction to go in, because there are no clear relationships seen.

The second parameter we looked at was the height at which the cloud was at in the atmosphere. We primarily looked at three different groups of clouds: low level, mid level, and high level clouds. These results can be seen in Fig. 7 and Fig. 8. In Fig. 7, if you look specifically at the way the low level cloud (clouds 1-6 on the x-axis) fractal dimension compares to the mid level clouds (7-12 on the x-axis), there is an overall increase. Similarly looking at the increase from the mid clouds to the high level clouds (clouds 13-18 on the x-axis), this makes it evident that there may be a relationship between the height of the cloud in the sky and the fractal dimension of the cloud boundary. After averaging the low, mid, and high level cloud fractal dimensions and plotting them seen in Fig. 8, the relationship is even more clear; as the height of the cloud in the sky increases, the fractal dimension seems to increase as well. Or, the higher up the cloud is, the rougher the cloud becomes. This is also obvious if you think about a cumulus, low-level cloud, the puffy, cotton ball like clouds, like the third image in the top row in Fig. 3. These clouds seem pretty smooth compared to more wispy, cirrus clouds that are more jagged, like the first image in Fig. 3, which are higher up in the sky.

Again, to make sure we were getting results that made sense, we checked our results for the fractal dimension of different clouds with other papers from places that might have better binarization techniques or better imaging systems. For example, our calculations show that cirrus clouds have a fractal dimension of 1.56, and a paper by Batista -Tomas reports in the Quarterly Journal of the Royal Meteorological Society [11] that the fractal dimension of cirrus clouds is approximately 1.39, giving a 10% error, as well as cumulonimbus clouds having a dimension of 1.23, when our program gives a value of 1.31, which is a 7% error.

There seems to be a relationship between the altitude of the cloud and the fractal dimension of the cloud, but no clear relationship between the thickness of the clouds and its fractal dimension. This is an interesting result, because as stated before, lower level clouds are generally thinner clouds and higher level clouds are generally categorized as thicker clouds. Although our results do not follow that general idea or result. This could be because

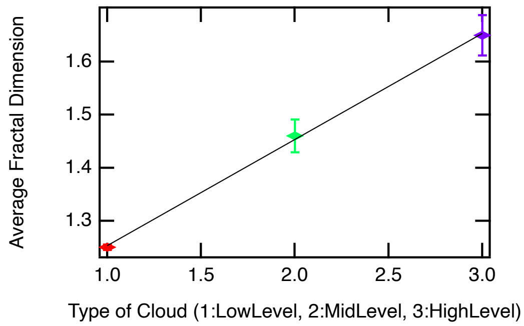

FIG. 8: Average fractal dimension of low, mid and high level clouds versus type of cloud: red diamond (numbered 1) represents low level clouds, the green diamond (numbered 2) represents mid level clouds, and the red diamond (numbered 3) represents high level clouds. As seen here, there is an direct trend, with a trend line in black with a fit of r-squared value to be 0.99. Error bars seen are 4% for low clouds, 3% for mid clouds, and 0.5% for high clouds.

challenges finding suitable images for analysis, because it was difficult to find high resolution images, and also there were no clear cloud image databases that had categorized clouds as thick or thin clouds. Thus, we were making educated guesses on how to categorize each cloud image into the category of thick or thin.

Error Analysis

In Fig. 4, all of the data points on the plot are exact, but when taking a section of the plot to use for results instead of the entire plot to find the slope of the data like we did, error is present. The error bars on Fig. 6, 7, and 8 were found by calculating the percent error between our calculated dimension values (values 10-70 in Fig. 4) and the values calculated between 0 and 80 on our log-log plot. While this is not a rigorous way to calculate error, the errors calculated gave us a better understanding of how much our data would change due to a slight change in data collection.

CONCLUSION

The purpose of this lab was to explore a new way to classify clouds to improve climate models and categorization of clouds in general. Through the box counting method, we were able to create a program that could calculate the fractal dimension of the boundary of cloud images. The code took in an image and binarized it according to the image's brightness. Pixel by pixel the image was covered with a grid, as seen in Fig. 2 to count how many boxes covered that actual boundary of the cloud. After varying the size of the box used, we plotted the results. We plotted the logarithm of the number of boxes that covered the boundary versus the logarithm of the inverse box size. To make sure that our program was running smoothly, we checked its accuracy by feeding the program an image with an already known fractal dimension. We used the Koch curve. After fitting a line to the graph and choosing which set of points were going to give us the best results, we found the dimension of the fractal to be 1.287, compared to the known value of 1.26, which gives a 2% error.

We run our cloud images through, comparing fractal dimension to thickness of a cloud and also fractal dimension to altitude placement of clouds. Results show that there is a lack of a relationship between the thickness of a cloud and its fractal dimension, as seen in Fig. 6. Although, when comparing the altitude placement of the cloud in the sky and its fractal dimension, results show that clouds placed higher up in the sky tend to have a higher fractal dimension, seen in Fig. 7. The average of each category: low level, mid level, and high level clouds, were averaged and fit to a linear trend line in Fig. 8 to strengthen the argument that higher clouds are rougher, or have a higher fractal dimension.

Our first result, that there is no clear relationship between the fractal dimension and the thickness of the cloud comes as a surprise when paired with the second result, that fractal dimension decreases as the altitude placement if the cloud in the sky increases. This is because generally clouds that are thicker are at higher points in the sky and thinner clouds are lower level clouds. Although these results are found by using small datasets, so trends and results could be strengthened or become more clear when using larger datasets. Also, we personally classified clouds as either thin or thick, there was no strong scientific reasoning behind why some clouds were considered thinner or thicker, other than the way the clouds looked (if they were transparent or opaque). The other clouds used at different altitudes were taken from a website that classified each cloud as either low, mid, or high level clouds, so we are more confident about those results.

Overall, further analysis must be done in order to determine if the fractal dimension is a strong metric for classifying clouds. Although, we do know that the fractal dimension of clouds is not something to overlook when creating climate models. The roughness of each cloud does differ, as this study does show, and that is important when creating a model with clouds. You cannot just use a circle or a shape from Euclidian geometry to describe the shape of a cloud, that is incorrect science. Thus this study could be helpful when determining how rough to make a cloud in a model.

FUTURE WORK

Further analysis of our work here could be useful and relevant. Using larger data sets or more accurate data sets can help strengthen our results and prove or disprove our result that the fractal dimension decreases for clouds higher up in the sky, and could help create new results about the thickness of the cloud compared to the fractal dimension. It could also be the case that there is no relation, which is still a result in itself.

Another way to better this program is to build off of it, and create somewhat of a backwards program. By that, we mean create a program similar to the one used by where the user inputs a fractal dimension, and maybe some other parameters like size of a cloud, and the program outputs an approximate cloud with that specific dimension. This could be helpful to modelers when trying to replicate the exact shape of the cloud, since that can be difficult.

Lastly, this program created can also be used for other objects, it is not only able to calculate the fractal dimension of clouds, but any object's boundary. For example, the program can be used to calculate the fractal dimension of leaves. Studies have shown the the fractal dimension of a leaf is related to the cell density of surface area of the leaf, which could be used as a new approach to taxonomical study of plants.

ACKNOWLEDGMENTS

I would like to thank my lab instructor Dr. Lehman for her help with helping me create a realistic direction to take my project in and constant guidance and teaching assistants Andrew Kunkel and Katie Shideler for their support and help during lab time. Thank you to Dr. Lindner for helping me come up with this idea and guiding me through it. Finally, thank you to Manon Grugel-Watson and the Physics Department for the tools and lab space.

[1] Lauren Harper. What are climate models and how accurate are they? https://news.climate.columbia.edu/2018/05/18/climate-models-accuracy/, 2018. Accessed on April 17 2021.

[2] Kazuo Gotoh and Yasuhiko Fujii. A fractal dimensional analysis on the cloud shape parameters of cumulus over land. *Journal of Applied Meteorology*, 37(10):1283 – 1292, 01 Oct. 1998.

[3] G. Pretor-Pinney. *The cloudspotter's guide.* Perigee, New York, 2014.

[4] P. Chylek. Light scattering, aerosols, clouds, climate, hendrik van de hulst, and i. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 206:333 – 337, 2018.

[5] Wikipedia contributors. Hausdorff dimension — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Hausdorff_dimension&oldid=999655108, 2021. Accessed on April 17 2021.

[6] Wikipedia contributors. Coastline paradox — Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Coastline_paradox&oldid=1015095299, 2021. Accessed on April 17 2021.

[7] Wolfram Research Inc. *Mathematica 11.2*, 2020.

[8] Shaun Lovejoy. Area-perimeter relation for rain and cloud areas. *Science (New York, N.Y.)*, 216:185–7, 05 1982.

[9] Fractal cloud shapes. https://geo.libretexts.org/Bookshelves/Meteorology/Book\%3A_Practical_Meteorology_(Stull)/06\%3A_Clouds/6.07\%3A_Fractal_Cloud_Shapes, 2020. Accessed on April 17 2021.

[10] Gavin Pretor-Pinney. Cloud appreciation society. https://cloudappreciationsociety.org/, 2016. Accessed on April 17 2021.

[11] A.R. Batista-Tomas, O. Diaz, A.J. Batista-Leyva, and E. Altshuler. Classification and dynamics of tropical clouds by their fractal dimension. *Quarterly Journal of the Royal Meteorological Society*, 142(695):983 – 988, 2016.