

Remote Sensing of Clouds and Biomes via Application of Temporal Analysis to Basic Methods

Alex Sturges and Adam Corness

Monday 6th November 2023

Abstract

Within satellite imagery, differentiating between cloud, land, ocean, and their respective subcategories is important for meteorological and ecological analysis. This data is used in many aspects of our daily life, such as maps, weather forecasts, and natural disaster detection. We developed techniques to perform cloud detection and removal on satellite imagery using temporal analysis, producing results that are qualitatively successful in equatorial and tropical regions, but fail toward the poles. A set of 23 images with clouds removed are produced by this in order to investigate biome changes over the course of a year. An unsupervised machine learning algorithm, SLIC, is then applied on these resulting images to segment and classify the African continent into different biomes, based on spatial and colour similarity. From qualitative analysis, this performed well on the majority of images, but 3-4 images have had mis-classifications of the biomes, resulting in an over-detection of desert regions. For both cloud removal and biome detection, an animation was made showing the variation of the Earths' surface throughout the year.

Contents

1	Introduction	2
2	Detection of clouds	3
2.1	Colour thresholding of images	3
2.2	Temporal Analysis of Colour Space	5
3	Cloud Removal	9
4	Biome detection	10
4.1	Method	10
4.2	Results	13
5	Discussion	15
5.1	Cloud detection and removal	15
5.2	Biome detection	15
6	Conclusion	16

1 Introduction

Remote sensing is the practice of acquiring information about a source from some distance, without any direct physical contact. Many types of remote sensors exist with applications across many fields; mostly within earth science disciplines. In this project, satellite imagery is considered - an example of a passive remote sensing technique, which observes reflected radiation from the Earth's surface and atmosphere. EUMETSAT comprises of a family of Geo-stationary satellites, with a focus on observing rapidly changing weather conditions over Europe, Africa and the Indian Ocean. Images used in this investigation were acquired by the EUMETSAT MSG-SEV1 satellite in, with its' main area of focus being the African continent. This satellite returns information relating to reflected radiation in 12 different spectral bands; our area of interest focuses on the first 3 channels, which consist of two visible light channels (VIS0.6 and VIS0.8) and one near Infra Red channel (IR1.6)[1]. Visible light channels are very useful in detecting thicker clouds that reflect a large amount of light, along with sensing different types of terrain/vegetation. The NIR channel also senses some reflected light, albeit with lower sensitivity. Value of this channel lies in its' ability to detect variations in atmospheric temperature [2], which allows the detection of cloud in areas that visible channels struggle to identify (i.e. more transparent clouds). Combining these three different channels provides enough information to differentiate between features of acquired images (i.e. clouds and terrain).

The data set used comprises of a set of 2190 unique images, split into three sets of 730 images of size 3712×3712 for each channel. We are looking over the range of one calendar year, from January 1st 2019, to January 1st 2020, meaning two images per day have been acquired for each channel. These images were taken at roughly the same two times for each day across the year, with the first taken at 12:27pm and the second at 12:43pm, differing only by a few seconds day to day. A monochrome image for each separate channel is acquired by the satellite, with colour images being produced by combination of the three (shown in figure 1).

When starting this project, our main goals were to produce a short animation showing the variation in African terrain throughout the course of the year, with separate biomes being automatically labeled. We wanted to implement remote sensing algorithms to identify two different classes of objects within the given images. As shown in figure 1, each image has significant cloud cover, reducing any information about the surface below. To be able to classify terrain, it was necessary to first develop a cloud removal algorithm.

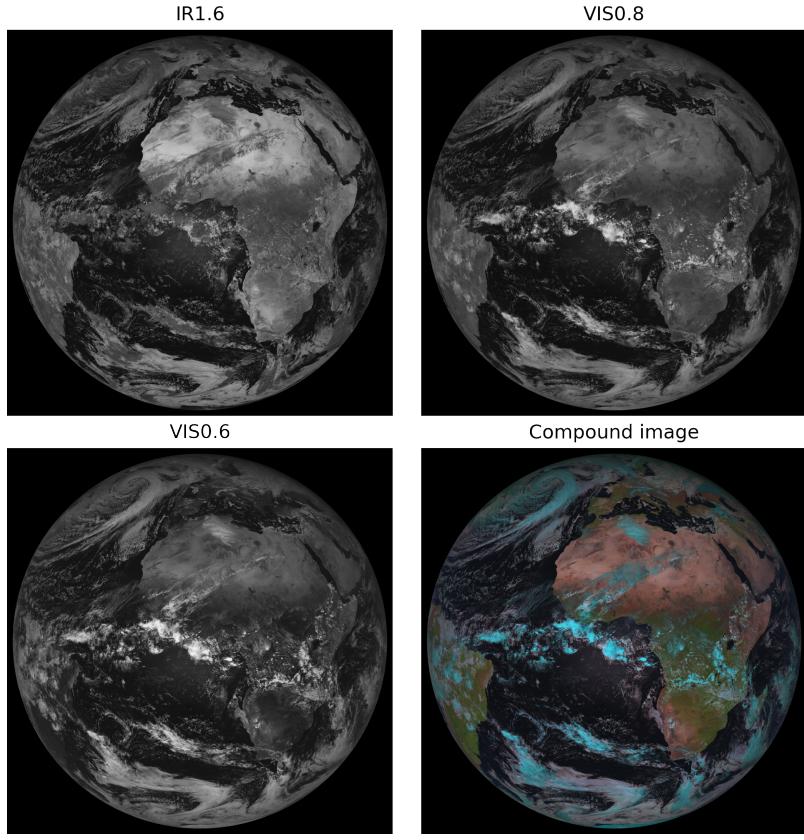


Figure 1: The first images of the dataset, taken January 1st 2019. Top left: Image in IR1.6 filter. Top Right: Image in VIS0.8 filter. Bottom left: Image in VIS0.6 filter. Bottom right: Colour image with R, G, B channels assigned as IR1.6, VIS0.6, VIS0.8.

2 Detection of clouds

2.1 Colour thresholding of images

The initial method for cloud detection involved colour thresholding in RGB colour space. By treating the IR1.6 channel as an effective 'red' channel, VIS0.8 as 'green' and VIS0.6 as 'blue', a colour space threshold can be defined. From inspection of figure 1, it is clear that in the visible channels that some clouds reflect a large amount of light, corresponding to high pixel values. Removing pixel values that lie above a specific threshold of brightness within each channel can differentiate between the much brighter cloud pixels, and the darker land/ocean pixels below. To find suitable thresholds, an investigation into where clouds lie within the colour space is required.

2.1 Colour thresholding of images

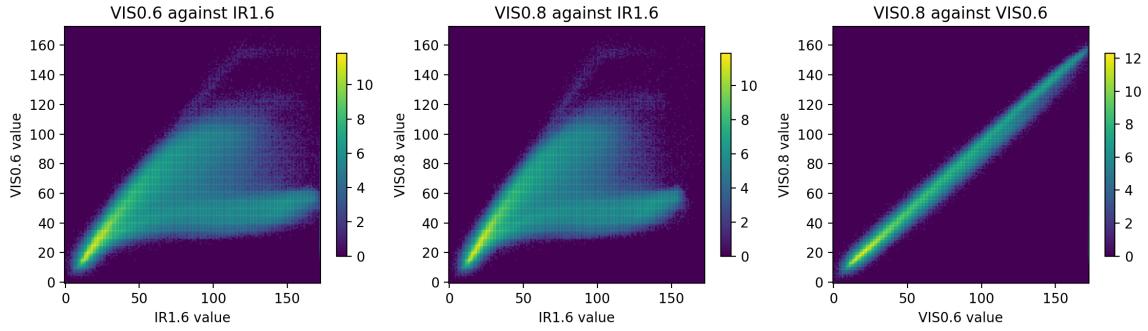


Figure 2: Each of these graphs show a logarithmic histogram of the recorded pixel intensity values in each channel throughout the course of the year (excluding values of [0,0,0]. Due to the high number of pixels in this sample, a 3D colourspace plot was not feasible due to memory constraints. Instead, projections of each respective plane in colourspace have been plotted. These values were taken over a small square region of the image, between pixel (1750, 1750) and (1850, 1850), covering an area in the Gulf of Guinea. The value of each point corresponds to the frequency the given colour value was recorded within the year-long time frame.

Figure 2 shows the colour space distribution of a central segment of the image. Upon inspection of the visible channel distribution (right in the figure), lower pixel values within this segment appear somewhat more frequent than in the higher end. This is to be expected, as the ocean pixel intensities correspond to 'darker', lower values. Higher intensity values appear less frequently than that of low intensities, which could indicate the presence of clouds in this case. However, there are no clear distinctions between where exactly clouds lie within this plot, meaning an effective threshold can not be determined from this data alone. The other two graphs (plots left & centre) show the projection of the IR1.6 channel on both visible channels. Again, no clear trends are shown here, with some high intensity values in visible channels combining with low values in the IR channel. The reason for this is that the IR channel detects fluctuations in temperature as well as reflected light, so colour proves even more unreliable within this red channel.

2.2 Temporal Analysis of Colour Space

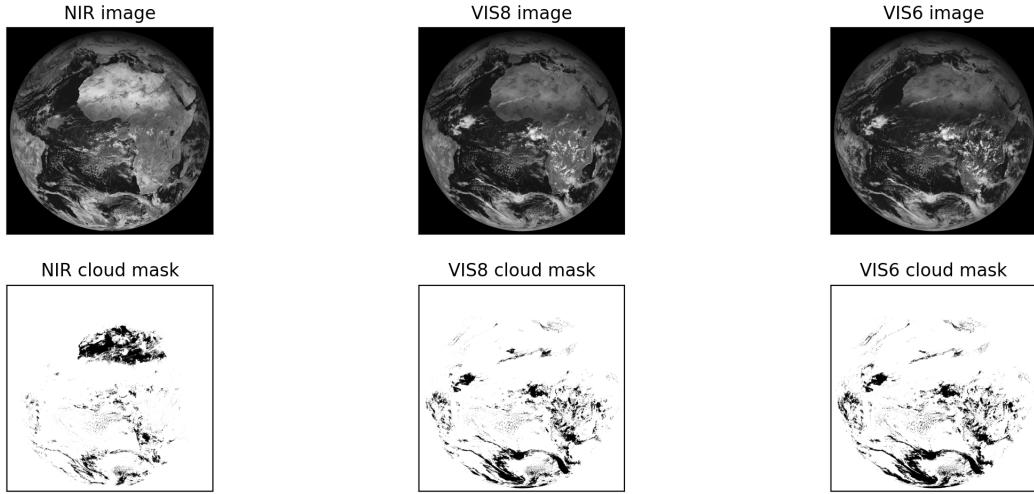


Figure 3: Test images are shown along the top row, with the detected clouds being shown on the bottom. Black pixels within the mask show where clouds have been detected in each test image. Visible channels manage to detect some cloud however

An attempt at implementing this thresholding technique on a single image is shown in figure 3. These represent a manual threshold classing any pixel intensity above 100 as a cloud. Within the visible channel, this simple technique picks up some of the brighter clouds, missing out the darker clouds up toward the poles. When applying this across different images, it becomes apparent that thresholds would need to be varied depending on the time of acquisition due to seasonal changes. In the infrared channel, a threshold of this nature will pick up terrain alongside missing out many clouds. Having found manually setting thresholds in this way to be unsuccessful, a revised method for cloud detection was needed. This is to be discussed in the next section on temporal analysis.

2.2 Temporal Analysis of Colour Space

Analysing how a pixel changes over time gives insight into how a cloud changes a pixels colour. A cloud produces a spike in the brightness in VIS0.8 and VIS0.6 channels but can either cause a spike or a dip in the IR1.6 channel, depending on the cloud type. Water clouds produce a spike due to reflection in the NIR channel, whereas ice and snow clouds produce a dip due to absorption. [2]. These snow and ice clouds appear invisible in the IR1.6 image and blue in the compound image from Figure 1.

2.2 Temporal Analysis of Colour Space

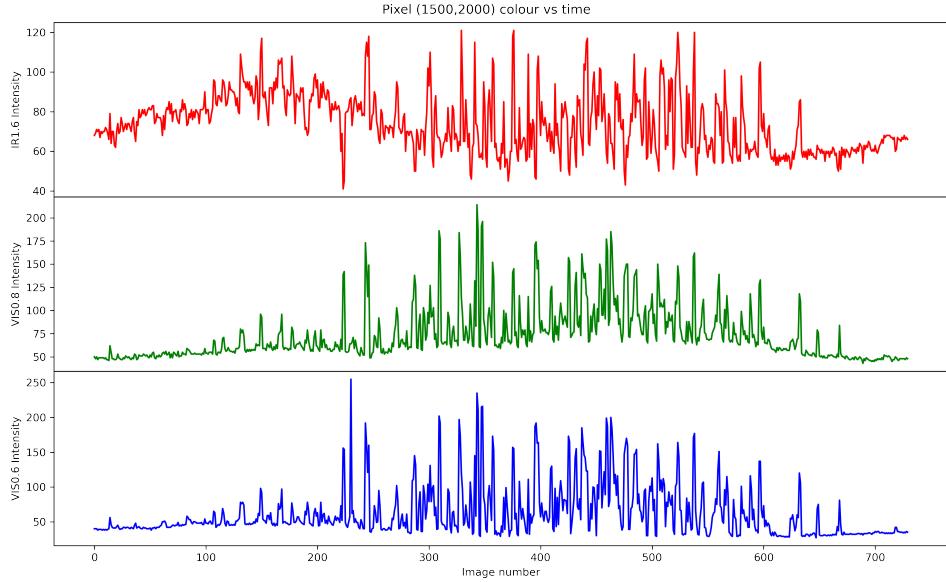


Figure 4: Plot of a single pixels brightness value across the 730 images in each colour filter.

Figure 4 shows how a pixel located at (1500,2000) in the image (which corresponds to eastern Nigeria), varies in colour throughout the year. There is a distinct change in the IR1.6 channel on a long timescale, along with the short timescale spikes which are also present in VIS0.8 and VIS0.6.

The spikes in brightness in the VIS0.8 and VIS0.6 match each other in time, due to both channels returning similar information at different intensities. This can be seen most prominently in the 3 spikes between image 600 and 700 in figure 4. This information can be used to define a cloud detection method based on the colour change of a given pixel throughout time.

To do this, a mean image representing the average value of each pixel is generated. By summing over all the images across each colour band as they were loaded, then dividing the resulting array by the number of images (730), one image containing non-integer colour values was produced. Ideally, a calculation of median would be used instead of the mean as it would reduce the skew from clouds. However, calculation of the median was unfeasible due to memory, processing, and time limitations as it would require storing far too many data points.

The mean image was then used to calculate the standard deviation for each pixel, using equation 1:

$$\sigma = \sqrt{\frac{\sum_k^n (i_k - \bar{i})^2}{n}} \quad (1)$$

Where i_k is the k^{th} image, \bar{i} is the mean image, n is the number of images, and σ is the standard deviation. The process had to also use a cumulative, image-by-image approach due to constraints on memory. The resulting standard deviation can be interpreted as an image as it shares the dimensions of a colour image.

Below both the mean and standard deviation are displayed as images:

2.2 Temporal Analysis of Colour Space

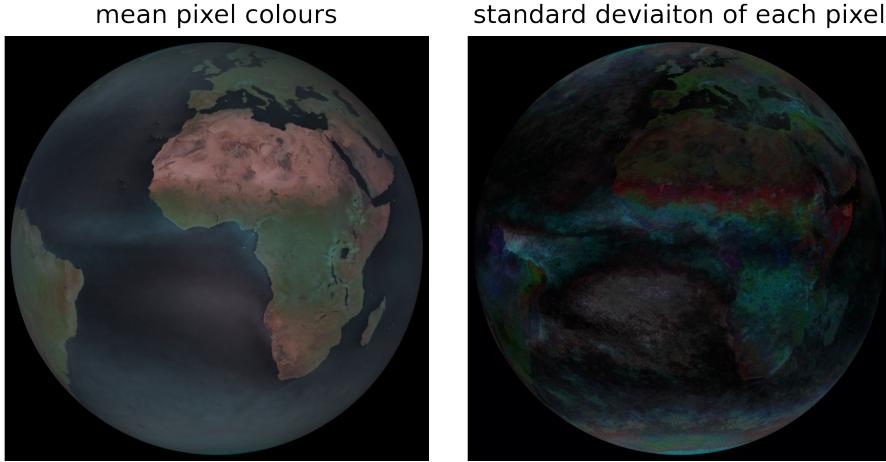


Figure 5: Left: Image of the mean pixel values. Right: Image of the standard deviation, colour indicates the size of the variance in that corresponding colour channel. e.g. A dark pixel has little variance and a bright blue pixel has high variance in the VIS0.6 colour channel.

The mean image in figure 5 is an image of the Earth mostly without clouds. Some regions where there are a large amount of clouds throughout the year such as the Gulf of Guinea, have a blue tint where clouds have skewed the average. This effect will be present throughout the whole image as through the year an effective thin layer of the cloud is present. Despite this, the mean image is a suitable reference value for what a pixel without clouds would be. This is because any skew due to a cloud could be accounted for in the thresholding.

The standard deviation image in figure 5 shows the pixels in the images where there is the most variance from image to image. Notable features are the bright blue region in the Gulf of Guinea and across central Africa especially near the great lakes. These regions are bright blue from large changes in VIS0.6 and VIS0.8 colour bands. This means they see frequent, but not constant snow/ice cloud cover. Europe is a green colour as it varies from the white clouds that cover it for most of the year to the green terrain underneath.

These 2 images give information about what a pixel's (nearly) standard colour is and how much it varies in any given colour band. This information was utilised to formulate a pixel specific threshold. Rather than thresholding in the colour space of a single image (spatial dependence) it effectively acted in the colour space of a single pixel across multiple images (temporal dependence).

To do this the mean was subtracted from each image, with the resulting differential image then being compared to the standard deviation image. Pixels could then be classified as a cloud according to the following rules:

$$i_r - \bar{i}_r \geq 2\sigma_r \cup i_r - \bar{i}_r \leq -\sigma_r \quad (2)$$

$$i_g - \bar{i}_g \geq 0.5\sigma_g \cup i_g - \bar{i}_g \leq -2\sigma_g \quad (3)$$

$$i_b - \bar{i}_b \geq 0.5\sigma_b \cup i_b - \bar{i}_b \leq -2\sigma_b \quad (4)$$

Where symbols have their previous described meanings, subscripts indicate the colour band (i.e. i_r indicates the IR1.6 band image). If any of equations 2,3, or 4 were true the pixel was classified as a cloud. A Gaussian blur with a kernel size of 3 was applied to the standard deviation image before the thresholding process. This was to ensure regions of low variance would not have false positive clouds identified due to noise. Using an un-blurred standard

2.2 Temporal Analysis of Colour Space

deviation resulted in disparate pixels across the Sahara desert being identified as clouds. Using this method worked well for identifying snow or ice clouds but performed poorly at identifying water clouds: when analysing figure 4 it is apparent why. Pixels in the VIS0.8 and VIS0.6 channels do not have a large variance on a long timescale, whereas the IR1.6 band fluctuates with the seasons. A bright cloud in the winter may appear the same as the warm ground in the summer, or a cold cloud in the summer appearing like the cold ground in the winter. To counteract this effect the moving average for the IR1.6 channel was used instead of the yearly average. The moving average window used was 60 images (i.e. a month long). This is long enough such that the average would not be greatly skewed by clouds but short enough so that seasonal changes would not be ignored. The moving averages for VIS0.8 and VIS0.6 bands were not calculated due to computational limitations (calculating just the moving average in IR1.6 increased the computation time from ≈ 20 minutes to > 1 hour).

With this moving average data the IR1.6 channel threshold shown in equation 2 was refined to be:

$$i_r - \bar{i}_r \geq \sigma_r \cup i_r - \bar{i}_r \leq -0.45\sigma_r \quad (5)$$

When manually comparing the results from using both this new threshold and equation 2, the new threshold was qualitatively more accurate than the old. When the threshold used is displayed on figure 4, it becomes apparent why these equalities are suitable:

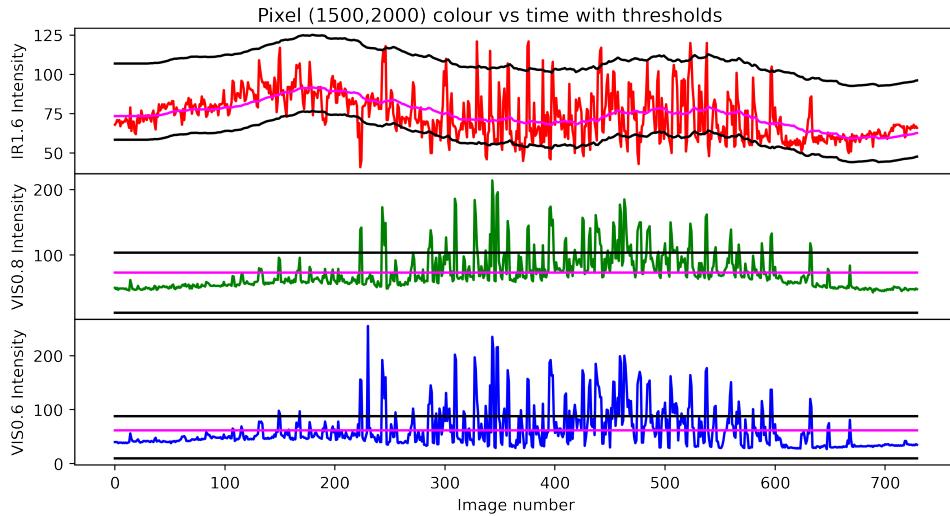


Figure 6: The plot of figure 4 but with the mean value shown in magenta (moving average for the top IR1.6 channel) and threshold bounds shown in black. An object within the black lines is considered not to be a cloud whereas an object outside them is.

Figure 6 shows why the thresholds applied work; nearly all spikes in VIS0.8 and VIS0.6 bands have been identified as clouds and little to no terrain has been falsely identified as a cloud. In the IR1.6 band it is apparent that not many peaks have been identified as a cloud. This did not impact the total threshold however, as many of these spikes are mirrored either of the VIS0.8 or VIS0.6 bands. In contrast, most of the clouds which cause a drop in the IR1.6 band have been identified correctly. Some smaller clouds notable in the early and late time portions of figure 6 cause only a minor change in temperature and remain undetected by this method. However, we believe that the degree of precision used here was sufficient for the

scope of this project.

3 Cloud Removal

If pixels identified as cloud are assigned values: [0,0,0] the resulting image looks like such:

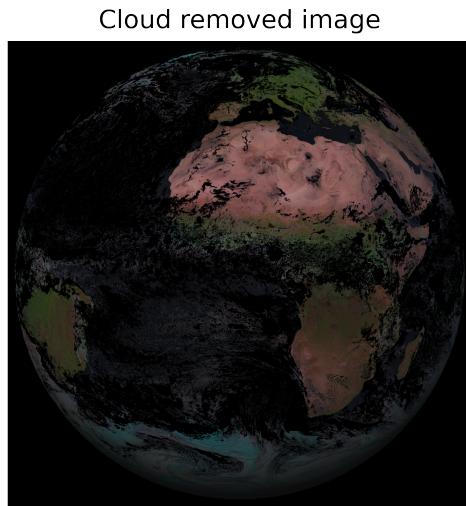


Figure 7: A compound image from May 30th 2019, pixels identified as cloud by the method in the previous section have been replaced with 0. This is the same compound image shown in figure 1

Figure 7 is the "output" of the cloud detection technique. It's clear that the method cannot remove 100% of the clouds, however, nearly all clouds have been removed from the African continent and most of the northern hemisphere. There are many clouds in the south as it is winter there. This makes it difficult to remove them as the moving average IR1.6 value at this time is that of cloud due to the prevalence of them throughout this time of year. The cloud detection technique allows us to assign clouds a colour of [0,0,0] (like shown in figure 7,) which means they are easy to count across images. A pixel with colour [0,0,0] is either in space or a cloud. To remove clouds, 60 consecutive images were averaged - this is the same period as the IR1.6 moving average because of similar reasons. Thirty days is short enough for seasonal changes to not have a large effect but long enough for a reasonable amount of data; meaning a pixel being a cloud in all 60 images is very unlikely. To average across the 60 images the following equation was used:

$$p_o = \sum_k^n \frac{p_k}{n - c_k}, \text{if } c_p = n \rightarrow p_o = 0 \quad (6)$$

Where p_o is an output pixel, n is the number of images in this case $n = 60$, c_p is the number of clouds in pixel p_k across all n images, and p_k is a pixel within the image. To avoid dividing by 0, if there is a cloud in all p_k across the n images, the output pixel takes a value of 0. This process is done for every pixel across the 60 images. This was repeated for ranges of 60

images starting from every 30th image. i.e. images 0-60, 30-90, 60-120... until 660-720. This created 23 cloudless images, the 10th of these is shown below:

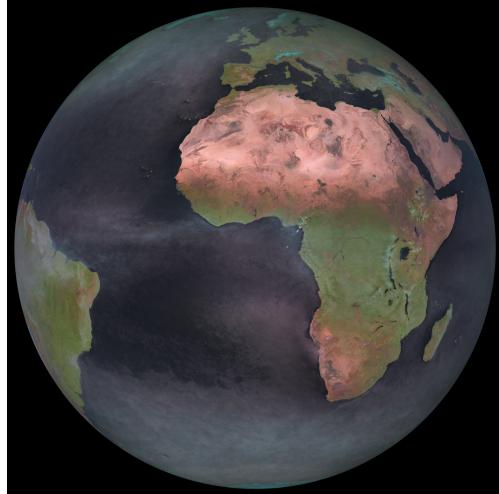


Figure 8: An image of the Earth with the clouds removed by compiling 60 images from late May 2019

Figure 8 shows the cloud removal method applied on the images from late May. Note the high mountains in Switzerland and Norway being visible due to snow and ice. Using simple thresholding for cloud removal may lead to seasonal information about these regions lost. The ocean has regions that appear to have thin clouds, however this did not impact the terrain analysis due to the shift in colour being small from this phenomena. Theoretically, 728 images could be made via this method, however this would prove very computationally expensive. For the purpose of biome analysis, 23 images sufficed in providing enough information. Any remaining pixels with a value of [0,0,0] were replaced with the corresponding pixel from the mean image. These images were then used for biome detection.

4 Biome detection

4.1 Method

Having removed clouds from the satellite images, the terrain below can now be investigated. The aim is to measure the variation in vegetation coverage of Africa throughout the year. To analyse this, a method was developed to classify terrain as either being vegetation, arid, or desert. This relied on implementing an image segmentation algorithm called simple linear iterative clustering (SLIC).

SLIC was developed in 2010 at the École polytechnique fédérale de Lausanne, as a method to group pixels based on spatial and colour similarities. For a full breakdown of the functionality of the algorithm used, please refer to [3]. Implementation of SLIC produces clusters of pixels based on their colour vector and spatial position. Once formed, clusters are then assigned a label corresponding to a superpixel, representative of their collective characteristics. Superpixel colours are determined by the mean colour of their composite pixels. For the purpose of creating regions based on their terrain type, this algorithm proves very useful. To reduce interference from other regions in the image, a mask was created to isolate the African continent. This was created using a basic image editing tool by manually drawing around the

4.1 Method

area of interest, removing any areas outside of this. After adding the mask, some water was still present within the image edges of the african region. Due to the water superpixel values being lower valued than those of terrain, these were quite easily removed using thresholding as the colour vectors are now quantised. Figure 9 shows the output of applying both the mask and SLIC, setting maximum superpixel number as 35 and a relatively low compactness of 1. This gives a reasonably high dependence on colour over spatial location. This can allow the superpixels to form more irregular, disconnected and concave shapes. A more compact representation of an image is output, with 8 different superpixels representing the African continent in this specific case. As superpixels still retain a form of colour information, different regions can then be classified as either vegetation, ocean or arid land using colour thresholds.

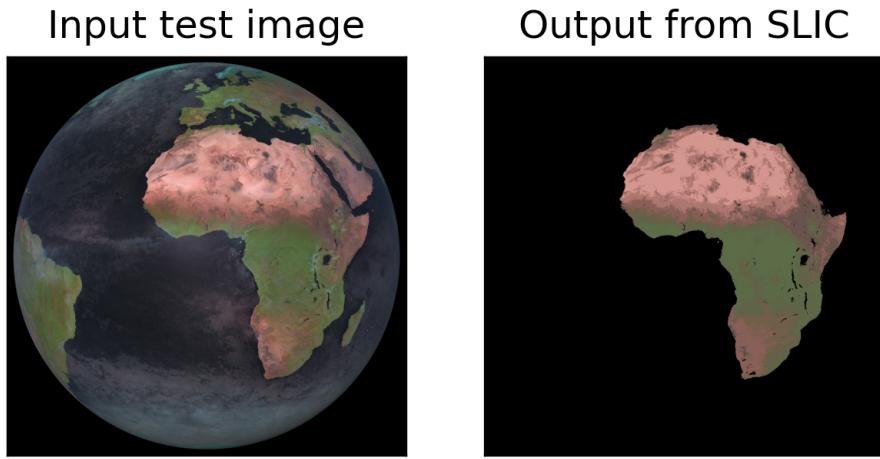


Figure 9: Shows the output having applied both the mask and SLIC function to the input image. This output image is made up of superpixels with 8 unique colours. We can see qualitatively that regions of different terrain have been successfully segmented, with the green superpixel values corresponding to vegetation, light brown areas correspond to desert and darker brown showing arid areas.

Finding thresholds was done by analysing the position of super pixels within colour space relative to each other. Figure 10 shows the Red/Green plane within colour space with respective pixel classifications. Differentiating between vegetation and the other two types of terrain was done using thresholds in this plane. From inspecting the colour space plots the following thresholds were derived:

$$P_c = \begin{cases} P_V & \text{if } G + 15 > R \\ P_{A,D} & \text{if } G + 15 \leq R \end{cases} \quad (7)$$

with P_c being the pixel classification, P_V being vegetation, $P_{A,D}$ being arid or desert, G being green value and R corresponding to red. The +15 term is a scaling factor determined by visual inspection of figure 10. Once pixels were classed to be either P_V or $P_{A,D}$, distinctions between arid and desert terrain had to be made. Classifying between these was done within the blue colour channel, using the following threshold:

4.1 Method

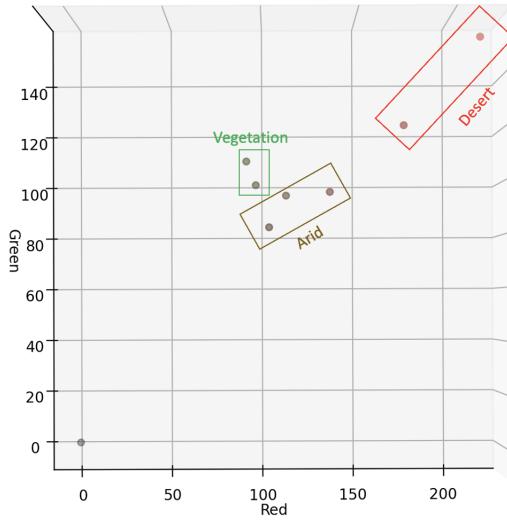


Figure 10: This figure shows a slice of the Red/Green plane, taken from a 3D colour space plot of the superpixels within the segmented image from Figure 9. The bounding boxes show the classification of these superpixels according to the rules outlined in equations 7 & 8.

$$P_c = \begin{cases} P_D & \text{if } B > 100 \\ P_{A,V} & \text{Otherwise} \end{cases} \quad (8)$$

with pixels falling within the otherwise category remaining as their original classification. Figure 11 shows where the desert pixels lie along the blue colour axis, distinctly higher than the other classifications. Across multiple images in this specific dataset, superpixel values vary a small amount, therefore in theory these thresholds shown in equations 7 & 8 can be applied to different images to return consistent classifications.

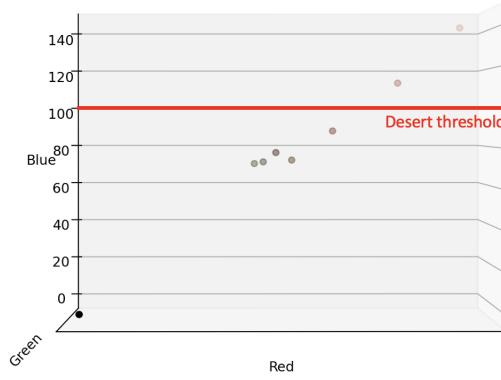


Figure 11: Slice from the same plot as in figure 10, this time displaying the variation on the blue channel. The red line labeled as the desert threshold shows pixels above this line are classified as desert, below as arid or vegetation.

4.2 Results

4.2 Results

Implementation of this method on an image returns an image with four different labels, shown in figure 12. This was applied across all 23 cloud removed images (detailed in section 3) throughout the time window, with arid and desert terrain being labeled within the same class, in order to better track the change in vegetated areas. An animation compiling all 23 images can be found here (<https://github.com/quasur/EUMESAT-proc/blob/main/Yearvid.mp4>). This shows the variation in biome size, with yellow indicating vegetation and blue indicating desert and arid regions. It is worth noting that some anomalous results appeared which can be seen in this animation, due to vegetation being mis-classified as arid terrain. Reasons for these outliers are to be looked at within the discussion. Figure 13 shows the change in pixels along the vertical axis (analogous to latitude). Again, these anomalous images show up between images (7-9), where there is a sharp but significant change in labeled values. The general trend here along latitudes in the region ≈ 1500 is from arid terrain to vegetation throughout the course of the year. Some values towards latitudes ≈ 2500 show a slight trend of increase in vegetation increase and then start reverting back to desert later in the year. Figure 14 shows another perspective, displaying the standard deviation of each latitude label. This is on an arbitrary scale, but offers insight into how frequently each latitude varied throughout the year.

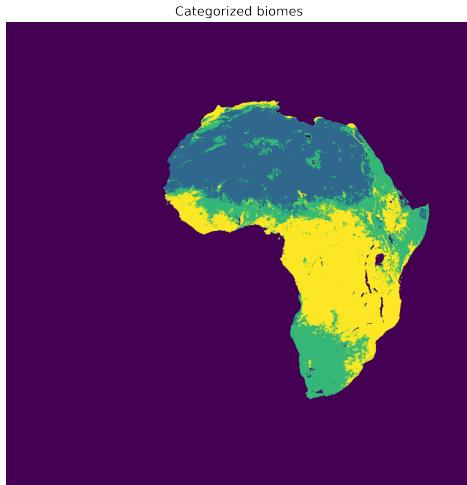


Figure 12: This is the result of applying the biome detection algorithm to the test image shown in figure 9. The purple label defines outside of region or water, the dark blue labels desert, green labels show arid and yellow defines vegetation.

4.2 Results

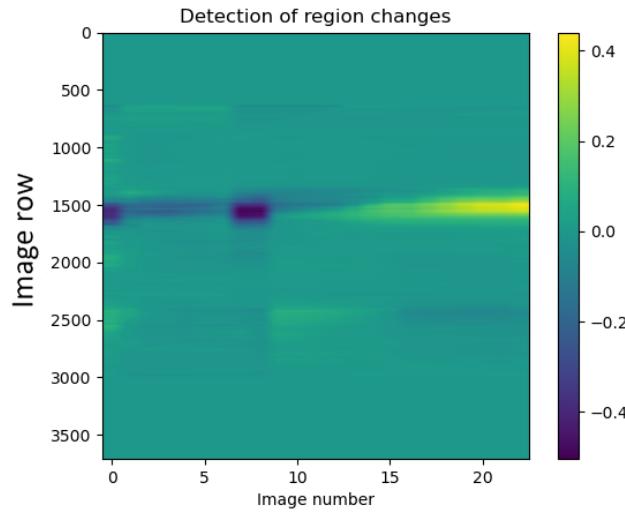


Figure 13: This plot shows detected changes in region classification for latitudes in the set of images. This is done by comparing the average label for each image latitude. If a label is different from the average, it is shown up here. A bright value corresponds to an increase in vegetation in that latitude, with the darker value indicating an increase in arid coverage. Values here are arbitrary, just necessary for scale.

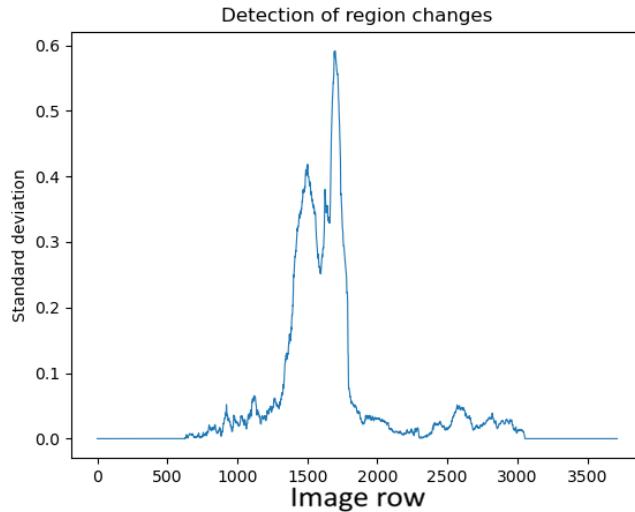


Figure 14: This figure showing the standard deviation in pixel row which corresponds to latitude. This is another perspective on figure 14, showing that image rows from ≈ 1400 to ≈ 1800 are where the greatest changes occurred over the course of the year. Anomalous images have been removed from this calculation.

5 Discussion

5.1 Cloud detection and removal

Potential improvements to the cloud detection method have been mentioned briefly before. Using the median instead of the mean as a standard pixel colour would avoid having a cloud-biased average for regions which do not have predominant cloud cover. We believe doing so would require either an entirely different approach to the coding method and data handling, or significantly more computing power. Another potential improvement would be to utilise more colour channels. Different wavelengths interact with clouds differently, utilising more of these wavelengths could allow greater precision of cloud detection, especially in polar regions. Adding a spatial dependence based on the region of the Earth a pixel lies within may improve the accuracy of cloud identification. This could greatly improve the sensitivity in the polar regions where there is a large amount of cloud cover. If the accuracy of cloud detection in the polar regions could be improved, the algorithm could be successfully applied to the tracking of polar ice. A different dataset with the polar regions as the predominant features would be required for this however. Additionally calculating the moving average of all colour bands may produce more precise or reliable results, though the impact of implementing this may not be as great as the previously discussed potential improvements.

These improvements are limited to modifications to the methods used within this paper, a statistical comparison to other cloud detection methods would be useful to quantify the accuracy of this method.

5.2 Biome detection

From viewing of the animation of biome variation, some seasonal change is detected, with vegetation coverage increasing over regions along the equator. Clearly, through application of SLIC we have been able to segment into distinct regions, albeit with some anomalous values. Anomalous images within the results show some shortcomings of SLIC, as there are some random classifications. Specifically, small regions in north Africa. We would have liked to have spent more time in refining the implementation of the algorithm, as it appears SLIC returned some superpixels that jump above the set thresholds, leading to mis-classification of regions.

By purely qualitative comparison to a particular study looking into terrain distribution across Africa [4], the classifications picked up are a rough match. Most of the previous studies in to this topic use data from many years, so it is not possible to get a direct comparison of previous work. Applying the model over larger data sets could return more meaningful results (i.e overall change in different biome size due to desertification). Having more terrain labels also would provide useful, possibly preventing issues with mis-classification seen in the results. However, these are only qualitative observations, with quantifying seasonal change proving a challenge. With more time allowing, it would have been useful to calculate the percentage coverage of each biome throughout the year, comparing these to values derived from the established NDVI method. This would provide a good numerical measure of how accurate this model is at detecting these regions.

6 Conclusion

It is important to note the environment in which this project was completed. The project was undertaken over the course of a single week using limited consumer grade laptop hardware. This greatly limited the scope of the methods we were able to employ. Despite this the work represents a successful proof of concept for the methods used, returning valuable information that could be refined with more time and greater computing power. We have managed to achieve and exceed the goals we started the project with: To produce an animation of the Earths' surface over the course of a year and to categorize different regions within its terrain. Further work could allow these methods to be used to track the extent of polar ice, an important measurement to track the progress of climate change.

References

- [1] Donny M. A. Aminou, Bernard Jacquet, and Frederick Pasternak. Characteristics of the Meteosat Second Generation (MSG) radiometer/imager: SEVIRI. pages 19–31, December 1997.
- [2] Thomas D. Potter and Bradley Roy Colman. *Handbook of weather, climate, and water : dynamics, climate, physical meteorology, weather systems, and measurements*. P.273–276. Wiley-Interscience, Hoboken, N.J., 2003.
- [3] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurélien Lucchi, Pascal Fua, and Sabine Süsstrunk. Slic superpixels. *Technical report, EPFL*, 06 2010.
- [4] Jeanine Umuhiza, Guli Jiapaer, Yu Tao, Liangliang Jiang, Liancheng Zhang, Aboubakar Gasirabo, Edovia Dufatanye Umwali, and Adeline Umugwaneza. Analysis of fluctuations in vegetation dynamic over Africa using satellite data of solar-induced chlorophyll fluorescence. *Ecological Indicators*, 146:109846, 2 2023.