2020209

Investigating the environmental impact of the January 2020 Kangaroo Island bushfire using Landsat 8 imagery.

2520 words

Introduction

The 2020 bushfire season was environmentally devastating across huge regions of Australia. Kangaroo Island, a tourism hotspot for nature reserves and wildlife, experienced its most devastating bushfire on record. By utilising the freely accessible multispectral data from the Landsat8 satellite, vegetation health will be mapped, between October 2019 and January 2020 to investigate the impact of the bush fire.

This report will:

- Visualise changes the fire caused to the island.
- Estimate the area of the island burnt by the January bush fire.
- Contextualise the changes in vegetation on the island.
- Quantify the area with severe vegetation loss as a result of the bush fire.

Methods

The Landsat 8 satellite collects 9 bands of data as it orbits the earth, it passes each location every 16 days, creating a detailed spatial and temporal record of our planet's surface (*Landsat 8 | Landsat Science*, 2021). The three satellite scans used in this analysis were collected on the 21st of October 2019, the 8th of December 2019 and the 9th of January 2020. They were selected due to low cloud cover and close relation to the bushfire. The scans of Kangaroo Island are on WRS path 098 and row 085. (see appendix for more information on data used).

In this analysis, the visible range (Red, Green and Blue), shortwave Infrared (SWIR) and near infrared (NIR) bands are used. These are collected at a resolution where each pixel resolves a 30m by 30m area (*Landsat 8 | Landsat Science*, 2021). These bands are used to create colour composite images and calculate a range of indices:

- Normalised Difference in Water Index (NDWI)
- Normalised Difference in Vegetation Index (NDVI)
- Normalised Burn Ratio (NBR).

These indices allow for mapping of the area burnt and quantify losses to vegetation health as a result of the bush fire.

The analysis is run with Python through Anaconda Navigator and Jupyter notebook. The base Jupyter environment is used with Rasterio pip installed though terminal. The satellite imagery data is accessible here and requires saving in the same folder as the notebook file.

To keep this analysis succinct, only the most significant code and results will be described in this report (where similar techniques are used it will only be explained once). Once the appropriate packages are installed, Landsat 8 bands two to six are loaded in for each month.

```python
#Jan_bands
Jan_bands = [
    "LC08_L2SP_098085_20200109_20200823_02_T1_SR_B2.TIF",  # Blue
    "LC08_L2SP_098085_20200109_20200823_02_T1_SR_B3.TIF",  # Green
    "LC08_L2SP_098085_20200109_20200823_02_T1_SR_B4.TIF",  # Red
    "LC08_L2SP_098085_20200109_20200823_02_T1_SR_B5.TIF",  # NIR
    "LC08_L2SP_098085_20200109_20200823_02_T1_SR_B6.TIF"   # SWIR
    ]
```

2020209

To read, crop (to area of interest) and convert from digital to reflectance, a for loop is made for each month, with the same crop and conversion rates for each.

```
# AOI (this crops lots of ocean around out)
crop_start_row, crop_end_row = 2000, 4300
crop_start_col, crop_end_col = 2200, 7400

# Converting each band from digital to reflectance, essential for analysis.
# This was found in the metadata file.
def convert_to_reflectance(band):
    """
    Convert a band from digitized to reflectance.
    """
    mult_parameter = 2.75e-05
    add_parameter = -0.2
    return band * mult_parameter + add_parameter
```

```
# Function to apply the above two to each stack of satellite scans.
# Set empty vector for it to go into.
October = []
for i in Oct_bands:
    with rasterio.open(i, 'r') as f:
        # Read the band data and crop it
        cropped_band = f.read(1)[crop_start_row:crop_end_row, crop_start_col:crop_end_col]
        # Convert to reflectance
        converted_band = convert_to_reflectance(cropped_band)
        # Bring it all together
        October.append((cropped_band, converted_band))

December = []
for i in Dec_bands:
    with rasterio.open(i, 'r') as f:
        cropped_band = f.read(1)[crop_start_row:crop_end_row, crop_start_col:crop_end_col]
        converted_band = convert_to_reflectance(cropped_band)
        December.append((cropped_band, converted_band))

January = []
for i in Jan_bands:
    with rasterio.open(i, 'r') as f:
        cropped_band = f.read(1)[crop_start_row:crop_end_row, crop_start_col:crop_end_col]
        converted_band = convert_to_reflectance(cropped_band)
        January.append((cropped_band, converted_band))
```

By setting a specific area to crop by this early on reduces the amount of data processed each time a band is used. The conversion to reflectance from digital is essential for accurate analysis and is based on the metadata within the Landsat 8 MTL file under REFLECTANCE MULT and ADD. The appending of cropped and converted bands requires careful selection as cropped bands are used for colour composite images and converted bands used for indices. It is important to keep them separated as colour composite images require modification to get them closer to what our eyes see however these adjustments will interrupt the accuracy of statistical results from indices.

Results

Colour composite images.

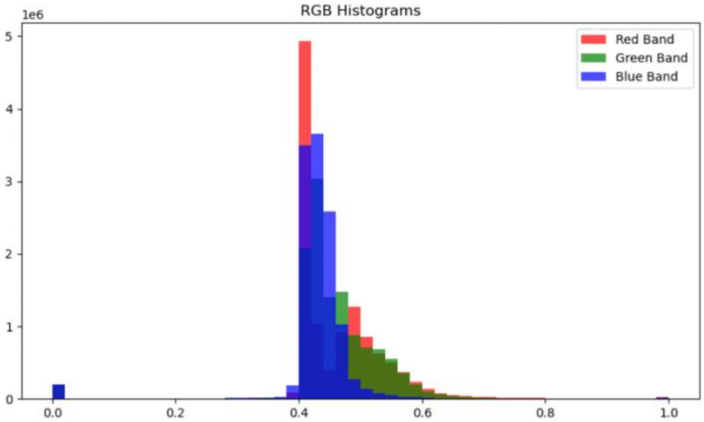
A colour composite for each satellite pass is created by combing the red, green and blue bands into a single image. Image normalization was chosen to create the colour composite maps as the reflectance conversions were producing dark images. A function is used to normalise each band between its minimum and maximum value, then brightness is increased by two to increase the visibility of features within the image. A function is used to create the plots, making the code as efficient as possible. This can be seen below:

```python
# Function to process and plot RGB composites with histograms
def process_and_plot_rgb_with_histograms(cropped_bands, title):
    # Extract the cropped blue, green, and red bands
    b2 = norm(cropped_bands[0][0].astype(np.float32))  # Blue
    b3 = norm(cropped_bands[1][0].astype(np.float32))  # Green
    b4 = norm(cropped_bands[2][0].astype(np.float32))  # Red

    # Create RGB composite
    rgb = np.dstack((b4, b3, b2))  # Stacking bands: Red, Green, Blue

    # Apply brightness adjustment
    rgb = np.clip(rgb * brightness_factor, 0, 1)

    # Plot the RGB image and histograms
    fig, axes = plt.subplots(1, 2, figsize=(16, 5))

    # Plot the RGB image
    axes[0].imshow(rgb)
    axes[0].axis('off')
    axes[0].set_title(f'Kangaroo Island Composite Image - {title}')

    # Plot histograms for the red, green, and blue bands
    colors = ['red', 'green', 'blue']
    for i, color in enumerate(colors):  # Loop through RGB channels
        axes[1].hist(
            rgb[:, :, i].ravel(),
            bins=50,
            color=color,
            alpha=0.7,
            label=f'{color.capitalize()} Band'
        )
    axes[1].set_title("RGB Histograms")
    axes[1].legend()
    plt.tight_layout()
    plt.show()
```

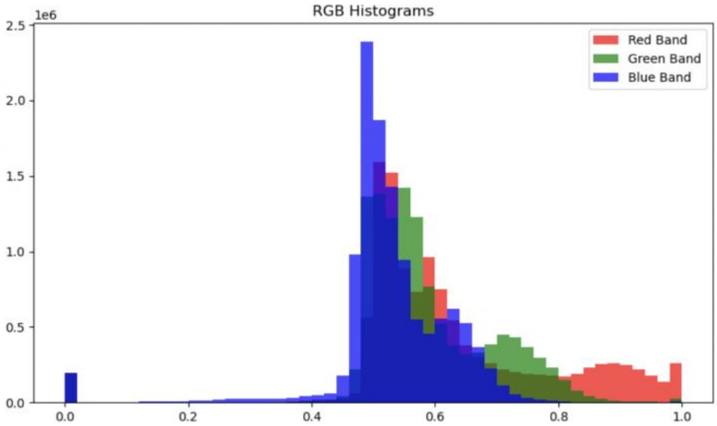
2020209

From the cropped bands, the [][0] indicates the cropped and not converted band is being selected. The bands of interest are selected and the results are plotted. For each scan the bands of interest are selected, and it is run through the function above to create the plots below.

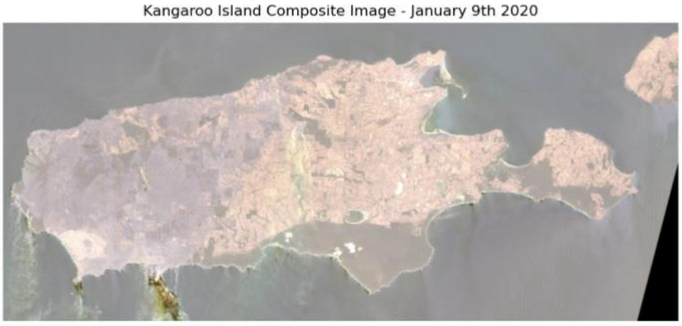
Kangaroo Island Composite Image - October 21st 2019

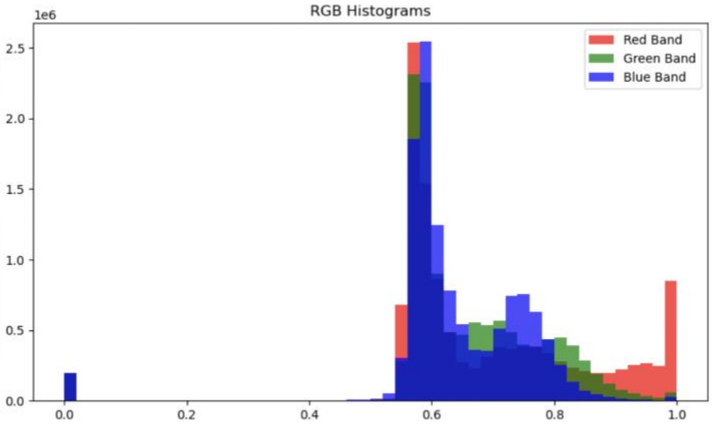


Kangaroo Island Composite Image - December 8th 2019



Kangaroo Island Composite Image - January 9th 2020





The colour composite images show that from October to January there is a visible change in the colour of the island. In October the island is mostly green, by November the forest regions across the west and south of the island are a similar but the rest of the island is looking dry and tan in colour. In the January image these forest regions across the west half of the island look dark and burnt. In both the December and January images plumes of smoke can be seen off the south coast of the island. The histograms for the maps show a shift in the value of pixels to generally increase as time progresses, mostly in the red band, followed by green and then blue. The images show there is some cloud cover across all three scans.

Creating a land mask using NDWI

To reduce the interference of the sea in data analysis, a mask which removes water is created. As shown by the colour composite image, some cloud cover is present in all three scans which causes the NDWI to mis-identify clouds over the sea as land. Additionally, there is some satellite drift between the scans, making a mean of all three images the most suitable method for defining the area the island covers rather than using a single scan as the mask. The NDWI uses the Green and SWIR bands to identify the water cover of a given area. To create the mask a threshold was set at -0.1 and all results larger than this were removed as they are identified to be water, the mask can be seen below.

Land Mask (NDWI < -0.1)

```python
Estimate_island_area= np.sum(land_mask>=-1)
# conversion from pixels to km²
area_km = round(total_island_nbr * 30 ** 2 * 1e-6,1)
print("Estimated Area of Kangaroo Island:", area_km, "km²")
print("Actual island extent 4,405 km²")
mask_accuracy = round((area_km/4405)*100,1)
print("Area of island accounted for with mask",mask_accuracy, "%" )
```

```
Estimated Area of Kangaroo Island: 4252.6 km²
Actual island extent 4,405 km²
Area of island accounted for with mask 96.5 %
```

Generally, the mask is effective at estimating the outline of Kangaroo Island. The mask includes some of mainland Australia in the north-east corner of the image, and inland lakes show more than once (as a result of the satellite drift). When compared to the actual area of the island 96.5% is accounted for meaning these issues present minimal impact on the analysis.

2020209

Estimating bushfire extent (NBR)

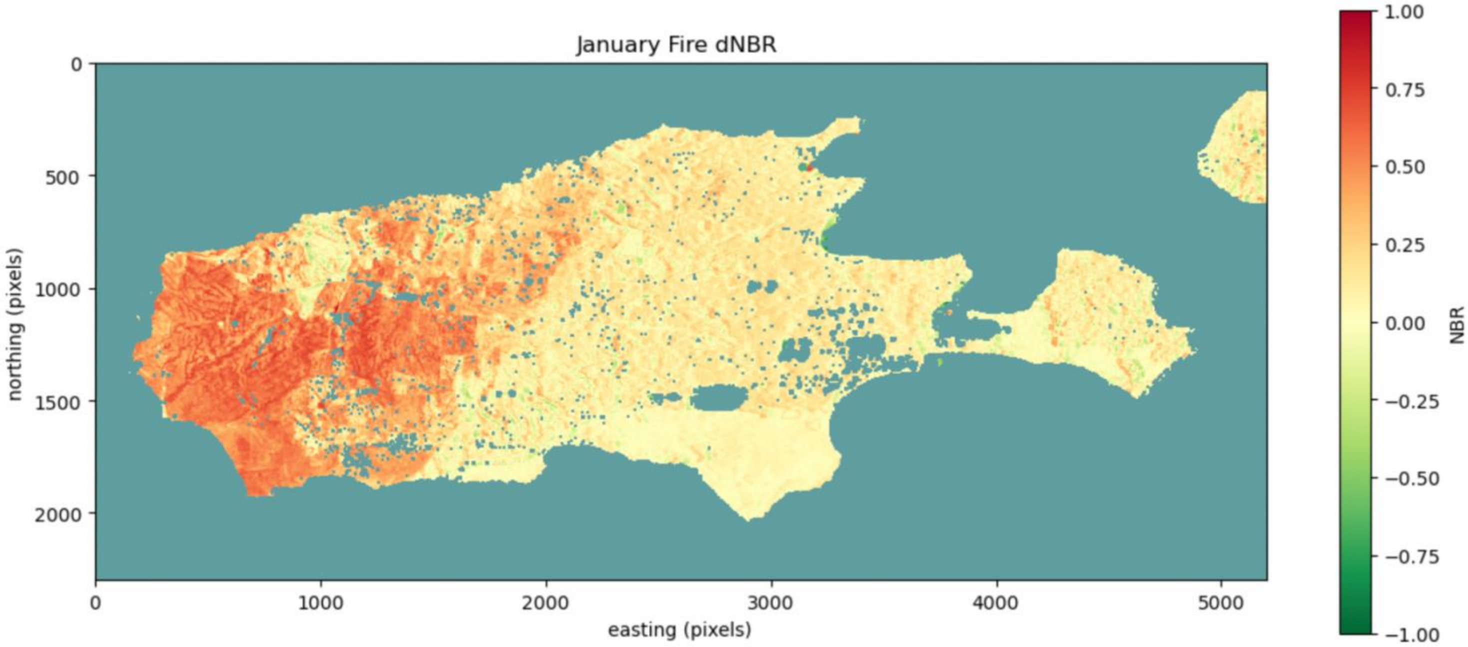
The normalised burn ratio is an index which identifies the severity of burn by comparing an area before and after a fire. In this case the comparison is made with satellite scans between December (pre-fire) and January (post-fire) by SWIR and NIR bands. A NBR function is defined which calculates NBR with the following equation:

```
(nir - swir) / (nir + swir)
```

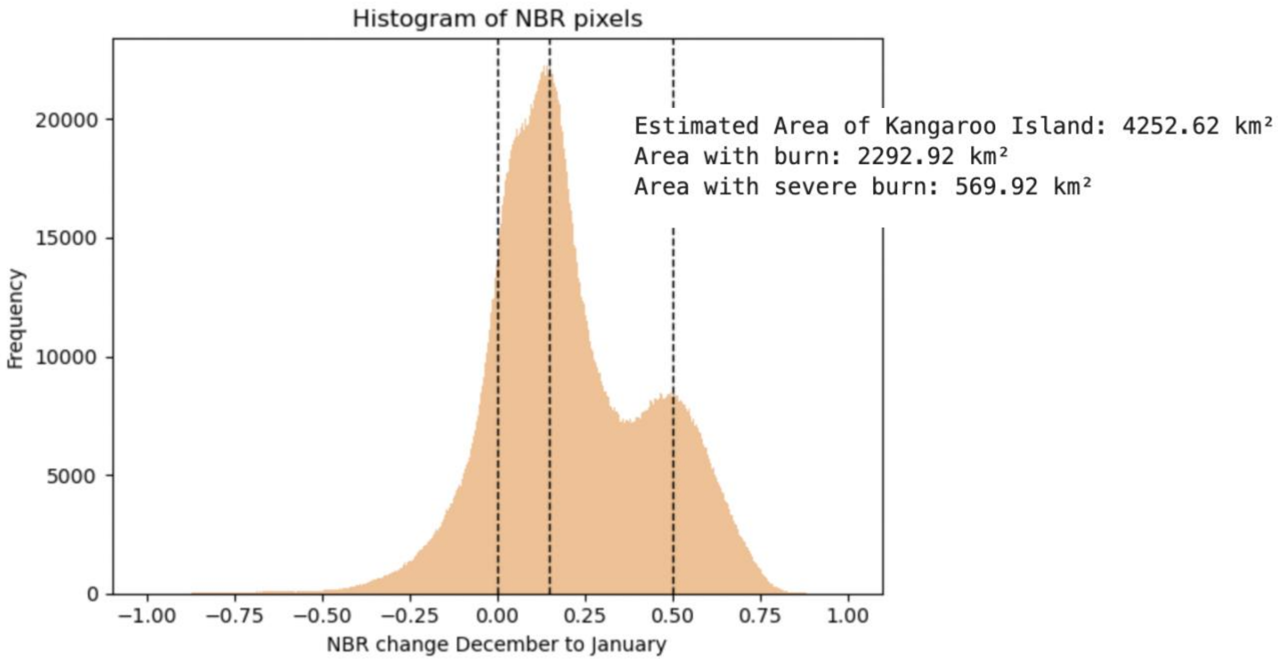
The appropriate bands are loaded and have the sea removed with the NDWI land mask. The difference between each month is calculated with a function. The map is plotted with Na values set to be blue to represent water.

```
# NBR for January
Jan_nir= January[3][1]
Jan_swir = January[4][1]
Jan_NBR = NBR(Jan_swir, Jan_nir)

# land mask
Jan_NBR_mask = Jan_NBR.copy()
Jan_NBR_mask[~land_mask] = float('nan')

# Calculating the change from Dec to Jan
def Dnbr(Jan_NBR_mask, Dec_NBR_mask):
    """
    Calculate the difference in burn rate
    """
    result = (Jan_NBR_mask)-(Dec_NBR_mask)
    return result

Fire_dnbr = Dnbr(Jan_NBR_mask, Dec_NBR_mask)

# Mask NaN pixels in the array
Fire_dnbr_masked = np.ma.masked_invalid(Fire_dnbr)

# Create a colormap and set NaN pixels to a specific color
cmap = plt.cm.RdYlGn_r  # reversed so that red shows more burnt
cmap.set_bad(color="Cadetblue")  # Set NaN pixels to Cadetblue


plt.figure(figsize=(14, 6))
plt.imshow(Fire_dnbr_masked, cmap=cmap, vmin=-1, vmax=1)
plt.colorbar(label="NBR")
plt.title("January Fire dNBR ")
plt.xlabel("easting (pixels)")
plt.ylabel("northing (pixels)")
```

2020209



The NBR map between January and December shows the extent of the island the bush fire covered. The more red, the more severely burnt that 30 x 30m pixel was. The majority of the west side of the island is red.

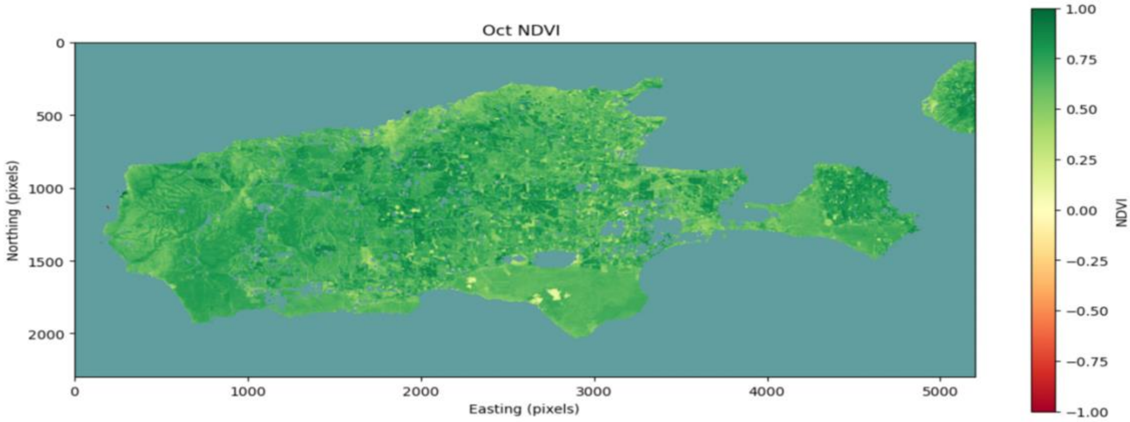


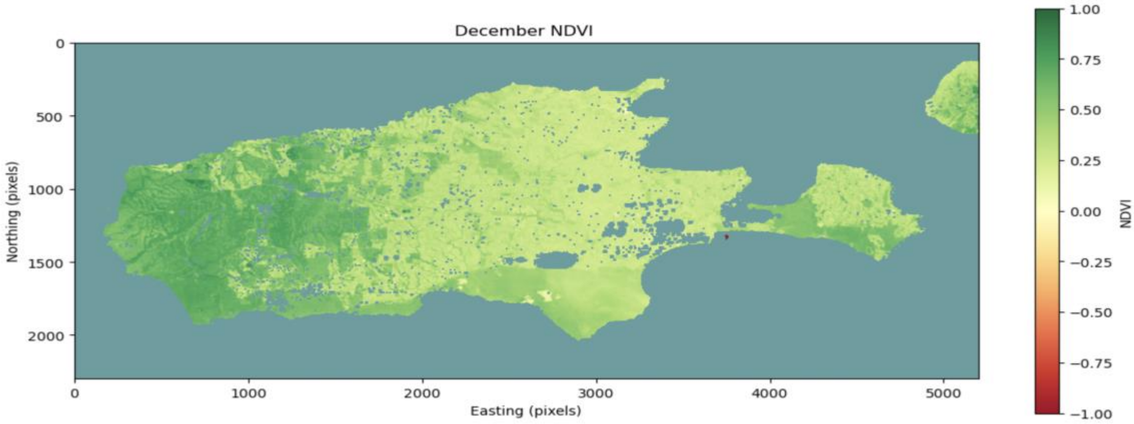
Of the 4250 km^2 area of the island, 2292 km^2 is identified as being burnt in some capacity, with 560 km^2 being severely burnt in terms of NBR. The largest frequency peak is at an NBR increase of 0.15, the second peak is at 0.5, The smaller peak is more likely to be showing the actual area burnt than the large peak as this may be a result of satellite paths not lining up.

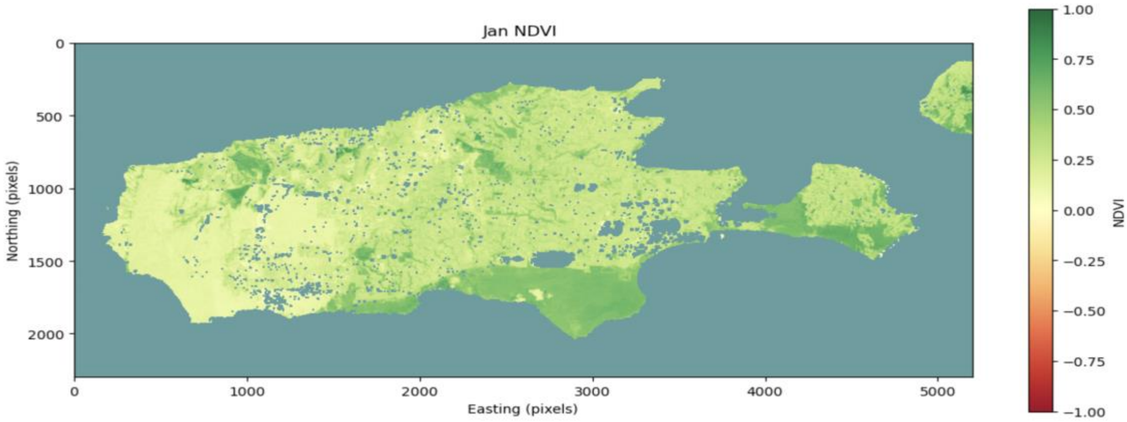
2020209

Contextualising vegetation change (NDVI).

NDVI is calculated in a similar way to NBR with SWIR being replaced by the red band. It is used to plot the following maps.







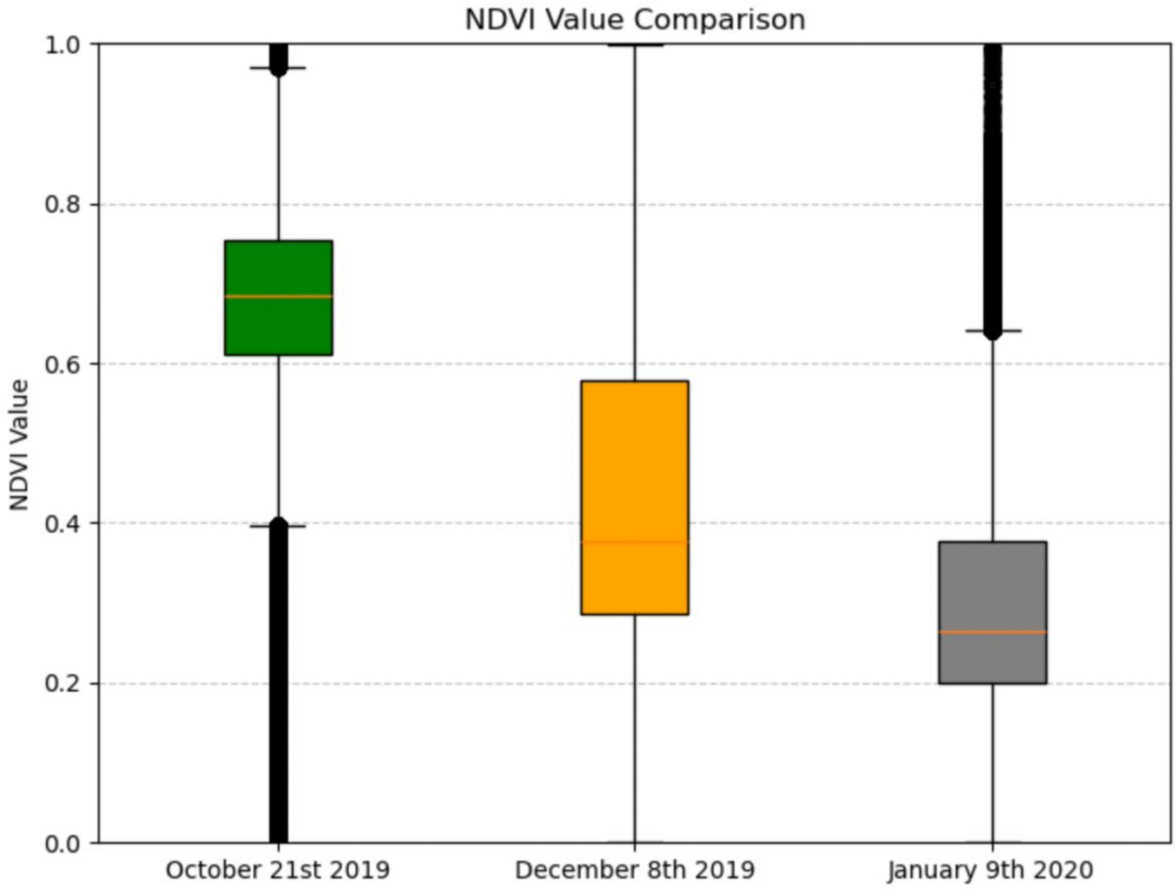
These maps show how vegetation changed through the time period. The NDVI maps show where vegetation is denser and healthier in green. Where values are whiter it represents areas with less healthy vegetation. In October, the island has healthy vegetation across its whole extent. This recedes to only the west coast of the island in December, and almost all of the high NDVI values have gone by January.

2020209

The dramatic change between October and December is unrelated to the fire in terms of direct impact but shows that the island may have been dry when the fire started. To investigate the scale of the loss of vegetation from the bush fire relative to the seasonal change in vegetation cover, the change between October and December can be compared to the change between December and January. This is done through a box plot, one-way Anova testing and Tukey HSD tests.

```
# Filter values between 0 and 1
jan_values = Jan_ndvi_mask[(Jan_ndvi_mask >= 0) & (Jan_ndvi_mask <= 1)]
oct_values = Oct_ndvi_mask[(Oct_ndvi_mask >= 0) & (Oct_ndvi_mask <= 1)]
dec_values = Dec_ndvi_mask[(Dec_ndvi_mask >= 0) & (Dec_ndvi_mask <= 1)]

# Making the lables nice
data = [oct_values, dec_values, jan_values]
labels = ['October 21st 2019', 'December 8th 2019', 'January 9th 2020']

# Colors for each box plot
colors = ['green', 'orange', 'grey']

# Create the box plot
plt.figure(figsize=(8, 6))
box = plt.boxplot(data, labels=labels, patch_artist=True)

# Apply colors to the box plots
for patch, color in zip(box['boxes'], colors):
    patch.set_facecolor(color)

# Add title, labels, and grid
plt.title('NDVI Value Comparison')
plt.ylabel('NDVI Value')
plt.ylim(0, 1)  # Restrict y-axis to NDVI range
plt.grid(axis='y', linestyle='--', alpha=0.7)
```

This code selects NDVI values between -1 and 1 as this is the range they should be between. 1 representing the densest healthy vegetation and -1 the least. I colour the box plot by month.



The box plot shows that from October to January there is a decrease in the overall NDVI of the island as the means and IQRs fall. There is the greatest variation in vegetation health in December (when the west side looks green, and the rest of the island doesn't). The difference between the first two months appears larger than the one between the last two months.

2020209

To quantify if the variation between the 3 months is greater than the variation within each month (if there is a statistically significant difference between the vegetation change), I compare ANOVA results using Tukey HSD. The code below runs the ANOVA test as well as plots the Tukey HSD results in a table.

```
# Run ANOVA test
anova_result = stats.f_oneway(oct_values, dec_values, jan_values)

# Print the result
print("ANOVA Test Result:")
print(f"F-statistic: {anova_result.statistic:.2f}")
print(f"P-value: {anova_result.pvalue:.5f}")

# Interpret the result
alpha = 0.05  # significance level
if anova_result.pvalue < alpha:
    print("The p-value is less than 0.05, indicating a statistically significant difference between the groups.")
else:
    print("The p-value is greater than 0.05, indicating no statistically significant difference between the groups.")

from statsmodels.stats.multicomp import pairwise_tukeyhsd

# Combine data into a single array and create corresponding group labels
all_values = np.concatenate([oct_values, dec_values, jan_values])
labels = (['October 21st'] * len(oct_values) +
          ['December 8th'] * len(dec_values) +
          ['January 9th'] * len(jan_values))

# Run Tukey's HSD test
tukey_result = pairwise_tukeyhsd(endog=all_values, groups=labels, alpha=0.05)

# Print the result
print(tukey_result)

# Plot the result
tukey_result.plot_simultaneous(figsize=(10, 6))
plt.title('Tukey HSD Test Results')
plt.show()
```

2020209

```
ANOVA Test Result:
F-statistic: 7773103.52
P-value: 0.00000
The p-value is less than 0.05, indicating a statistically significant difference between the groups.
      Multiple Comparison of Means - Tukey HSD, FWER=0.05
===============================================================
   group1       group2    meandiff p-adj  lower   upper  reject
---------------------------------------------------------------
December 8th  January 9th  -0.1269   0.0 -0.1271 -0.1267   True
December 8th October 21st   0.2436   0.0  0.2434  0.2439   True
 January 9th October 21st   0.3705   0.0  0.3703  0.3707   True
---------------------------------------------------------------
```

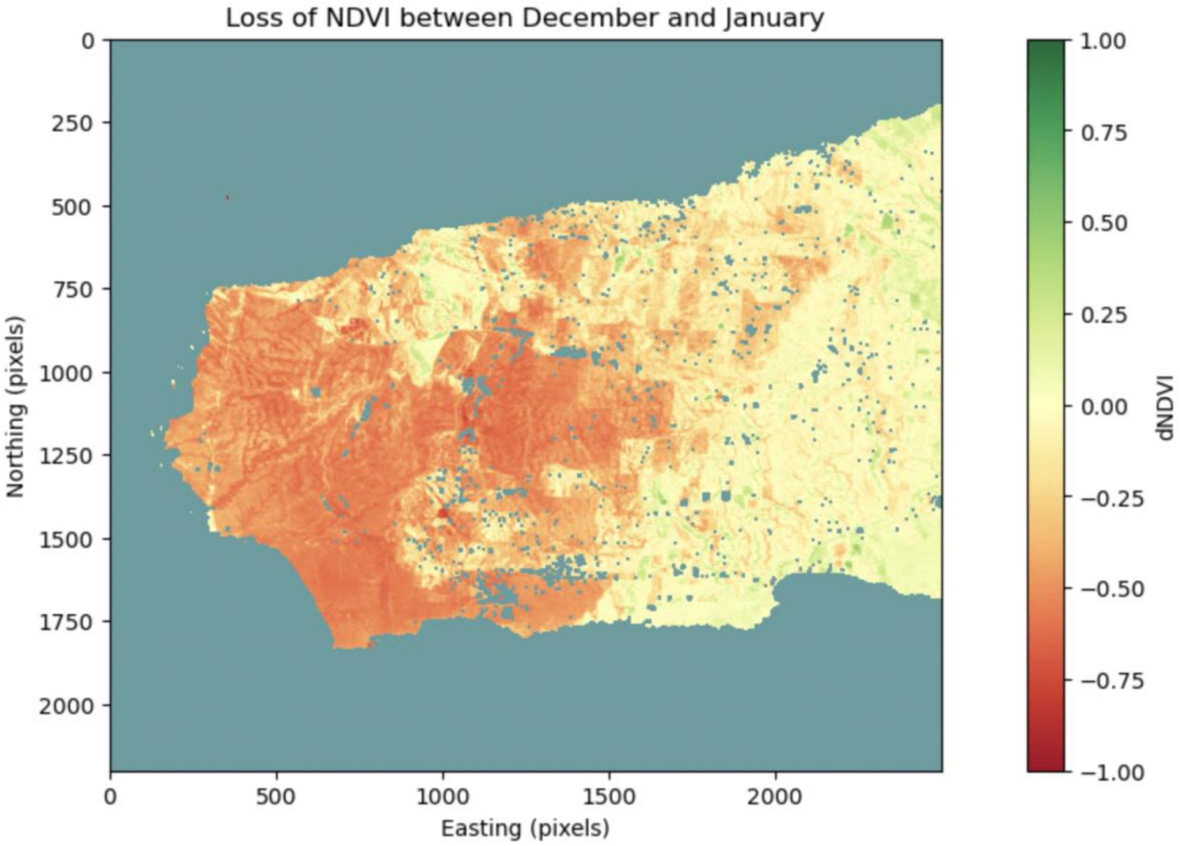
This output shows that the difference between all three months are statistically significant at a 95% confidence interval (as all p values are below 0.05). Unsurprisingly the largest difference is between October and January as it is the longest time period. However, it is notable that there is a greater difference between NDVI values between October and December than there is between December and January, indicating that there is more vegetation health loss before the fire than caused directly by it.

Assessing Vegetation loss as a result of the fire

As previously shown by the NBR map, the fire occurred on the west side of the island, and as the differences between satellite paths are adding some systematic error to the results, minimising the non-fire area assessed is crucial. This analysis will only assess the vegetation loss in this area on the west side of the island, so it is cropped to a smaller area of interest while the land mask is added.

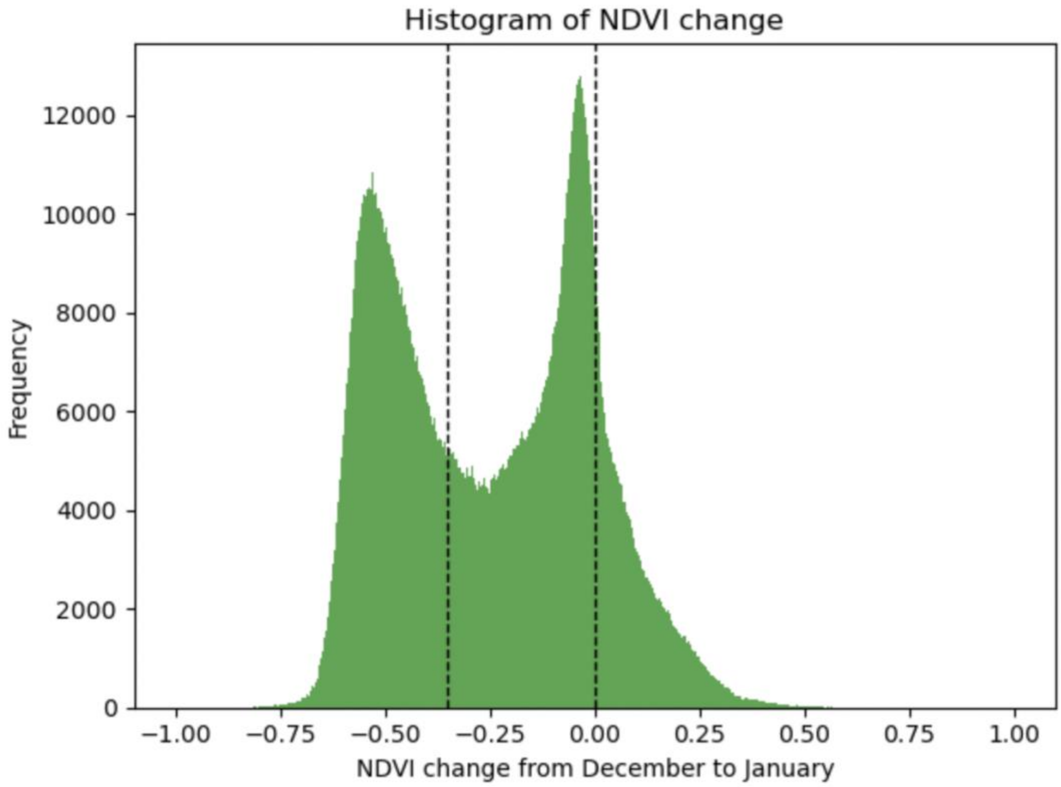
```python
# Crop boundaries
crop_start_row = 100  # This is the y-axis
crop_end_row = 2300
crop_start_col = 0  # This is the x-axis
crop_end_col = 2500

# Apply the new crop
Jan_ndvi_mask_cropped = Jan_ndvi_mask[crop_start_row:crop_end_row, crop_start_col:crop_end_col]
Dec_ndvi_mask_cropped = Dec_ndvi_mask[crop_start_row:crop_end_row, crop_start_col:crop_end_col]

# Difference in NDVI from December to January
def Dndvi2(Dec_ndvi_mask_cropped, Jan_ndvi_mask_cropped):
    """
    Calculate the difference in vegetation rate (dNDVI).
    Handles NaN values.
    """
    # Ensure valid values before subtraction
    result = np.where(
        np.isnan(Dec_ndvi_mask_cropped) | np.isnan(Jan_ndvi_mask_cropped),
        np.nan,
        Jan_ndvi_mask_cropped - Dec_ndvi_mask_cropped
    )
    return result
```

2020209



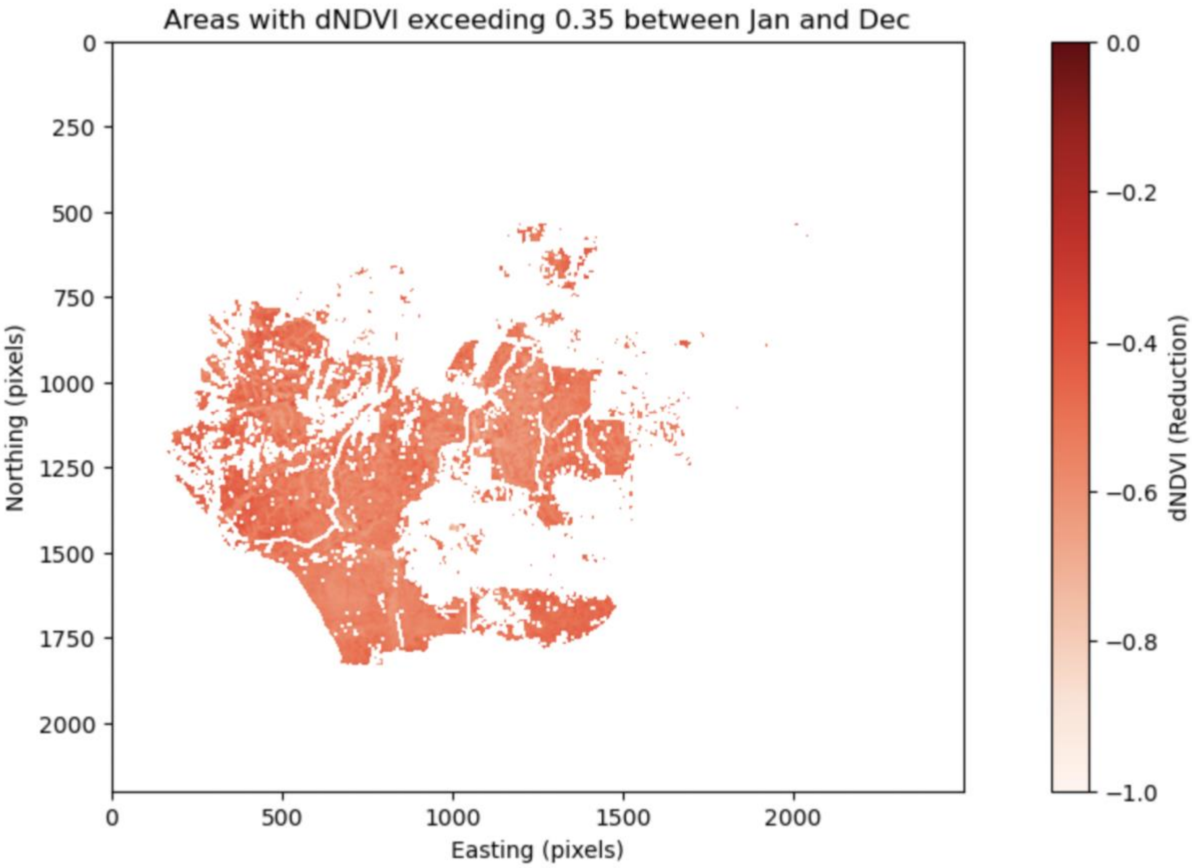
The map of bush fire area and its NDVI loss between December and January shows that most of the west of kangaroo island lost vegetation health. There are some na values, some of these will be areas with water, others a result of the scans not lining up.

2020209



When analysed further with a histogram of the index values, there are two peaks in the NDVI loss values. The largest peak is very close to 0, which could represent the error caused by the scan from January and December not aligning properly. There is a larger peak around -0.5 showing that there is a group of pixels in this west area of the island which has lost around half of its vegetation density.

2020209




```
Estimated Area of Kangaroo Island: 2440.7 km²
Area with lost vegetation: 1570.6 km²
Area with severe loss in vegetation: 983.8 km²
Approximate area of island with burn related Vegetation loss 36.9 %
Approximate area of island with severe burn related Vegetation loss 23.1 %
```

Further focus can be made by mapping all pixels which lost more than 0.35 in NDVI between the two months, representing areas with severe loss in vegetation health. Proportionally this area of severe vegetation loss is 23.1% of the whole island and 36.9% of the islands area is shown to have some loss in vegetation.

Discussion

The use of Landsat 8's multispectral imagery allows for the effective analysis of the 2020 bushfire on Kangaroo Island. The colour composite images and NDVI results show that leading up to the fire, the island lost much of its green vegetation. The NBR is used to visualise and quantify the large extent of the island the fire engulfed and when combined with NDVI data, identify the scope of vegetation lost. It was found that 36.9% of the island's vegetation lost health and density, accounting for approximately 1570km^2 of land, of this 983km^2 was severely impacted.

The strength of multispectral imagery technology is the ability to analyse and visualise changes at a reasonable resolution through time, without having to access remote or dangerous areas. The combination of NBR and NDVI is recognised as an effective method for identifying burnt areas and extracting useful information on vegetation (Liu, Freudenberger and Lim, 2022) (Gibson *et al.*, 2020). The analysis is improved by the removal of data from the sea through the NDWI based sea mask. This reduced interference of NDVI and NBR noise to increase the accuracy of the data analysed. The threshold at which vegetation burn severity was decided on based on the histogram of the data. The level of change caused by the fire is contextualised by the change in vegetation in the months leading up to the fire.

The main limitation and source of uncertainty within this study comes as a result of the drifting path of the satellite between scans. A shift in this path means the pixels compared between passes do not directly line up. This is of particular note around inland lake regions, which due to the NDWI derived sea mask have been removed, removing some land data. Subsequently some index pixels are worked out between a value and a NA value skewing the data. Further uncertainty is introduced by the use of the NDWI as a mask for the islands bounds. Non-island data is included as could cover and mainland Australia is difficult to distinguish from the target island land.

Future studies could focus on reducing uncertainty by working on a solution for the satellite drift issue. This could be done by finding a different way to mask the island such as lining up a shape file of the island with the raster data. Further exploration into the change in land cover as a result of the fire would require more post fire scans but would provide useful insight into the resilience and ability to recover of vegetation in the area.

Conclusion

This report uses satellite imagery from Landsat 8 to visualise and quantify the impact of a bush fire on its vegetation cover. The combination of NDVI and NBR indices allow the impact of the fire on vegetation to be contextualised within the broader timeframe. Colour composite images visualise the change between the 21st of October 2019 and the 9th of January 2020, with the island transforming from green and healthy to dry and burnt. Spectral indies quantify that 1570km^2 or 36.9% of the islands vegetation saw a loss in health and density as direct result of the fire, with 983km^2 of this seeing severe losses in vegetation cover. These areas of loss quantified are supported by maps, histograms and statistical testing. There were some challenges maintaining accuracy in the analysis, as a result of drifting satellite paths and cloud cover introducing potential uncertainty into the results, particularly around coastal regions. Overall, this work highlights the utility of remote sensing for environmental monitoring, with the Landsat 8 data providing reliable and accessible data for assessing the severity and environmental impact of a bushfire.

Appendix

Landsat 8
WRS Path 098
WRS Row 085
https://earthexplorer.usgs.gov/

November 21st 2019 - Green
https://earthexplorer.usgs.gov/scene/metadata/full/5e83d14f2fc39685/LC80980852019294LGN00/

December 8th 2019- Dry
https://earthexplorer.usgs.gov/scene/metadata/full/5e83d14f2fc39685/LC80980852019342LGN00/

January 9th 2020- Burnt
https://earthexplorer.usgs.gov/scene/metadata/full/5e83d14f2fc39685/LC80980852020009LGN00/