Reading the Green: Understanding RGB Vegetation Indices, Histograms and Statistics in Drone Mapping

How a handful of pixel values can help us investigate vegetation patterns, and why the numbers behind a colourful drone map matter.

Table of Contents

Reading the Green: Understanding RGB Vegetation Indices, Histograms and Statistics in Drone Mapping 1

1. A green map is only the beginning 2

2. Before the index: understanding the pixels 3

What changes when we use multispectral imagery? 4

3. Turning pixel values into a vegetation index 4

MPRI: Modified Photochemical Reflectance Index 5

Why divide by the sum? 5

VARI: Visible Atmospherically Resistant Index 6

4. What do negative values, zero and positive values mean? 6

What about −1 and +1? 7

VARI needs a little more care 7

5. From individual pixels to a histogram 8

What the shape tells us 9

6. The mean, the standard deviation and the normal distribution 9

The mean: the average index value 9

The standard deviation: how much the values vary 10

What is a Gaussian, or normal, distribution? 10

7. What happens when vegetation changes over time? 11

Why a histogram can reveal more than the mean 12

8. Why multispectral indices can tell us more about plants 12

NDVI: Normalized Difference Vegetation Index 12

NDRE: Normalized Difference Red Edge Index 13

Why the difference matters 13

9. Histograms work the same way for RGB and multispectral data 14

10. Practical considerations when analysing drone vegetation indices 14

11. The real value is understanding the data 15

12. Putting vegetation indices to work in a real survey 15

Further reading 16

1. A green map is only the beginning

Imagine that you have flown a drone over a patch of indigenous forest, a recovering fynbos site or an area where invasive trees are being cleared. You process the photographs into an orthomosaic and generate a vegetation index in QGIS.

The result is a map of different colours. Some areas appear bright green, others are pale green or yellow, and some are dark or even negative.

What does this actually tell you?

It is tempting to assume that the brightest green areas contain the healthiest plants and that darker areas represent stressed vegetation. Sometimes the pattern will broadly reflect differences in vegetation, but the colours alone do not tell the whole story.

To understand what the map is showing, we need to look at how the numbers were produced, what those numbers mean and how they are distributed across the landscape.

Three concepts help us do this:

- **Vegetation indices** turn measurements of light into numerical indicators.
- **Histograms** show how those numerical values are distributed across an area.
- **Statistics** summarise the distribution so that we can compare areas, identify variation and investigate changes over time.

The same principles apply whether we are examining vegetation recovery after clearing, monitoring tree canopy development, mapping agricultural crops or investigating changes in wildlife habitat.

Let's start with the information that the drone actually captures.

2. Before the index: understanding the pixels

A drone photograph is made up of thousands or millions of tiny squares called pixels. Each pixel represents a small part of the scene.

In an ordinary RGB image, every pixel contains three colour-channel values:

- **R — Red**: the recorded intensity of red light.
- **G — Green**: the recorded intensity of green light.
- **B — Blue**: the recorded intensity of blue light.

For example, a pixel might contain these values:

Channel	Pixel value
Red (R)	80
Green (G)	160
Blue (B)	40

These are illustrative digital numbers, not measurements from a real survey.

In a conventional 8-bit image, each channel commonly ranges from 0 to 255. A value of 0 represents the lowest recorded channel value, while 255 represents the highest.

A pixel with values of R = 80, G = 160 and B = 40 has a stronger green-channel value than either its red or blue channel. It would generally appear greenish.

But there is an important distinction: the pixel values are not direct measurements of plant health.

They are digital representations of the light recorded by the camera. Their values can be affected by the amount and wavelength of light reflected from the surface, the illumination, camera exposure, white balance, sensor characteristics and image processing.

A bright green pixel is not automatically a healthy leaf. It might represent vigorous vegetation, but the recorded colour can also be influenced by lighting, shadows, soil, dry vegetation and other objects.

This is why we use vegetation indices: mathematical calculations that combine selected bands to highlight particular differences in the recorded light.

What changes when we use multispectral imagery?

A multispectral camera records selected wavelength bands beyond, or in addition to, the visible colours. Depending on the sensor, these can include near-infrared (NIR) and red-edge bands.

Plants absorb much of the visible red light used in photosynthesis, while leafy vegetation generally reflects strongly in the near-infrared. The red-edge region is also useful for investigating changes associated with leaf pigments and canopy condition.

These differences give multispectral indices additional information that an ordinary RGB camera cannot record. An RGB photograph does not contain genuine NIR or red-edge measurements simply because software can calculate a vegetation index from it.

The distinction becomes important when we start examining the mathematics.

3. Turning pixel values into a vegetation index

A vegetation index combines selected pixel values into a new number.

Instead of examining red, green and blue separately, we calculate a relationship between them. The resulting index becomes the value stored in a new raster layer, with a calculated value for each valid pixel.

Let's look at two indices commonly used with drone RGB imagery.

MPRI: Modified Photochemical Reflectance Index

For this article, we will use the following MPRI formula:

MPRI = (G − R) / (G + R)

The calculation compares the difference between the green and red channels with their combined value.

Why do we do this?

Imagine two pixels:

- Pixel A has a green value of 160 and a red value of 80.
- Pixel B has a green value of 80 and a red value of 160.

For Pixel A:

(160 − 80) / (160 + 80) = 80 / 240 ≈ 0.333

For Pixel B:

(80 − 160) / (80 + 160) = −80 / 240 ≈ −0.333

The first pixel has a positive index because its green value exceeds its red value. The second has a negative index because its red value exceeds its green value.

The index has transformed two channel measurements into a single number that expresses their relative difference.

This is useful because a large collection of pixels can now be analysed according to the same measure.

Why divide by the sum?

Suppose we have two pixels with the following values:

Pixel	Red	Green	MPRI
A	80	160	0.333
B	40	80	0.333

Pixel B has half the red and green values of Pixel A, but its MPRI is identical.

This happens because both values have been reduced proportionally. The relationship between green and red remains the same.

Dividing by their sum therefore makes MPRI less sensitive to a simple proportional change in brightness than a calculation using only the difference G − R.

However, it does not remove all the effects of lighting, shadows, camera processing or changes in the surface being photographed.

VARI: Visible Atmospherically Resistant Index

VARI uses the same difference between green and red, but includes blue in the denominator:

VARI = (G − R) / (G + R − B)

Let's use another illustrative pixel:

- Red = 80
- Green = 160
- Blue = 40

The calculation becomes:

VARI = (160 − 80) / (160 + 80 − 40) = 80 / 200 = 0.4

The resulting VARI value is 0.4.

Notice that the same pixel produced an MPRI of approximately 0.333 and a VARI of 0.4. Both calculations used the same original RGB values, but they produced different results because the formulas are different.

VARI was designed to emphasise visible greenness while reducing some sensitivity to atmospheric effects. In drone imagery, its behaviour still depends on the camera, the recorded values and the conditions under which the photographs were captured.

Neither formula measures plant health directly. Each produces an index that can help distinguish spectral patterns in the image.

4. What do negative values, zero and positive values mean?

An index value becomes much easier to understand when we examine the relationship between the numbers in its formula.

For MPRI, the interpretation of the sign is straightforward, provided the denominator is positive.

- Negative values: the red channel is greater than the green channel.
- Zero: the red and green values are equal.
- Positive values: the green channel is greater than the red channel.

Consider these illustrative examples:

Red	Green	MPRI	What the number tells us
180	60	−0.500	Red exceeds green
120	120	0.000	Red and green are equal
100	150	0.200	Green exceeds red
80	160	0.333	Green exceeds red more strongly
40	200	0.667	Green strongly exceeds red

These values describe the relationship between two recorded channels. They do not establish that the pixel represents bare soil, a stressed plant or a healthy canopy.

For example, a negative MPRI value could occur over a surface dominated by red or brown material, but we would need the image and surrounding context to determine what that surface actually is.

Likewise, a positive value may be associated with green vegetation, but it does not prove that the vegetation is healthy.

What about −1 and +1?

For MPRI, if the red and green values are non-negative and their sum is greater than zero, the result is bounded between −1 and +1.

A value of +1 occurs when the red value is zero and the green value is positive. A value of −1 occurs when the green value is zero and the red value is positive.

These are mathematical extremes, not universal thresholds for healthy and unhealthy plants.

A value close to +1 means that green dominates red very strongly within the formula. It does not mean that a plant is 100% healthy, nor does it represent 100% vegetation cover.

Similarly, a value near zero means the two channel values are similar relative to their sum. It does not automatically mean that there is no vegetation.

The index tells us about the recorded spectral relationship, not the biological condition by itself.

VARI needs a little more care

VARI is often displayed using a range of colours that makes negative and positive values easy to distinguish. However, its mathematical behaviour differs from MPRI.

Its denominator is G + R − B, which can approach zero or become negative. Consequently, VARI is not guaranteed to remain between −1 and +1.

If the denominator is close to zero, even a modest difference between green and red can produce an unusually large value. If the denominator is zero, the calculation is undefined.

This matters when interpreting extreme VARI values. A very large positive or negative number is not necessarily an exceptionally strong vegetation signal. It may instead reflect an unstable denominator or unusual pixel values.

Before interpreting a VARI histogram, check its actual value range, look for extreme values and confirm that the input bands and NoData pixels have been handled correctly.

5. From individual pixels to a histogram

So far, we have examined individual pixels. But a drone survey may contain hundreds of thousands or millions of them.

Looking at each number individually would be impractical. A histogram gives us a way to understand the distribution of all those values at once.

A histogram groups pixel values into intervals, called bins. The horizontal axis shows the index values, and the vertical axis shows how many pixels fall into each interval.

Imagine calculating MPRI for every valid pixel in a small patch of forest.

Some pixels might have values near −0.2, others near 0.1, and many might fall between 0.3 and 0.5.

The histogram reveals where the values are concentrated and how much variation exists across the area.

For example, suppose we analyse 1,000 pixels and obtain this simplified distribution:

MPRI interval	Number of pixels
−0.50 to −0.25	50
−0.25 to 0.00	100
0.00 to 0.25	200
0.25 to 0.50	450
0.50 to 0.75	200
Total	1,000

The largest group lies between 0.25 and 0.50. Most of the pixels have positive MPRI values, and the distribution is concentrated in the middle-positive intervals.

If these pixels come from a forest canopy, this pattern may be consistent with a large proportion of pixels having green values greater than red values.

However, the histogram alone cannot tell us which pixels represent tree leaves, which represent shadows, and which contain a mixture of vegetation and background. We must interpret the distribution alongside the orthomosaic and the location being analysed.

What the shape tells us

A histogram can reveal patterns that are hidden when we look only at a single average.

- A narrow distribution: many pixels have similar index values.
- A wide distribution: the index values vary considerably.
- One prominent peak: many pixels cluster around one range of values.
- Two or more peaks: the area may contain distinct groups of pixels, such as green canopy and exposed ground.
- A long tail: a smaller number of pixels have values substantially different from the main group.

These patterns are clues, not automatic classifications. A bimodal histogram, for example, does not prove that there are exactly two vegetation types. Different surfaces can produce similar index values, and one type of vegetation can produce a wide range of values.

The value of the histogram is that it helps us ask better questions about the landscape.

6. The mean, the standard deviation and the normal distribution

Histograms become even more useful when we connect them to statistics.

Three concepts are particularly important: the mean, the standard deviation and the normal distribution.

The mean: the average index value

The mean is calculated by adding all the pixel index values together and dividing by the number of valid pixels.

For example, imagine that five pixels have MPRI values of:

0.10, 0.20, 0.30, 0.30, 0.60

Their mean is:

Mean ($\bar{x}$) = (0.10 + 0.20 + 0.30 + 0.30 + 0.60) / 5 = 0.30

The mean MPRI is 0.30.

This provides a single summary of the average index value across the five pixels. It does not mean that every pixel has an MPRI of 0.30. Some are lower, and one is considerably higher.

When comparing two surveys, the mean can help us identify a shift in the overall distribution. But a change in mean MPRI does not automatically establish that vegetation health has improved or declined. Changes in illumination, exposure, canopy composition or background can also shift the mean.

The standard deviation: how much the values vary

The standard deviation describes how spread out the values are around their mean.

Imagine two areas with the same mean MPRI of 0.30.

- In Area A, nearly every pixel has a value close to 0.30.
- In Area B, some pixels have values near 0.05, while others are near 0.55.

The means may be identical, but the distributions are very different.

Area A has a relatively small spread. Area B has a larger spread.

In vegetation monitoring, that difference may be worth investigating. The wider distribution could reflect a more varied canopy, mixtures of vegetation and bare ground, differences in illumination or other spatial patterns.

The standard deviation tells us that variability exists. It does not explain its cause.

What is a Gaussian, or normal, distribution?

A Gaussian distribution, also called a normal distribution, is the familiar bell-shaped curve.

In a normal distribution, most values cluster around the mean, and progressively fewer values occur farther away from it. The distribution is symmetrical around its centre.

For a perfectly normal distribution:

- Approximately 68% of values lie within one standard deviation of the mean.
- Approximately 95% lie within two standard deviations.
- Approximately 99.7% lie within three standard deviations.

These are properties of the normal distribution, not rules that every vegetation histogram must follow.

A real vegetation-index histogram may be skewed, contain several peaks or be constrained by the index formula. It may not resemble a bell curve at all.

For example, a recovering site containing dense vegetation, bare ground and scattered shrubs may produce several distinct groups of pixel values. Forcing that distribution into a normal model could obscure meaningful differences.

The mean is the centre of the data, the standard deviation describes its spread, and the histogram shows the actual distribution. The Gaussian curve is a useful reference when the data are reasonably close to normally distributed, not a shape that the data are required to follow.

7. What happens when vegetation changes over time?

Now let's apply these concepts to a real-world monitoring question.

Suppose you survey a recovering vegetation site in February 2025 and again in March 2026. You want to understand whether the vegetation has changed following invasive-species clearing or natural regeneration.

You generate an MPRI raster for each survey and extract the pixel values from the same study area.

Your first survey has a mean MPRI of 0.22. Your second survey has a mean MPRI of 0.34.

At face value, the second survey has a higher average MPRI. That tells us the relationship between the green and red channel values has shifted across the analysed pixels.

But what does that shift mean?

It could be consistent with an increase in green vegetation. It could also be influenced by differences in lighting, camera settings, seasonal conditions, exposed ground, vegetation composition or other changes in the scene.

To make a stronger interpretation, we should examine several things together:

1. The histograms: has the distribution shifted, narrowed or broadened?
2. The mean and standard deviation: has the average changed, and has variability increased or decreased?
3. The mapped locations: where are the changes occurring?
4. The original imagery: do the changes correspond to visible vegetation, soil, shadows or other surfaces?
5. Survey consistency: were the images captured and processed in a sufficiently comparable way?

A higher mean combined with a shift of much of the histogram towards higher MPRI values provides stronger evidence of a widespread change in the RGB spectral pattern than a mean change alone. It still does not prove improved plant health without supporting evidence.

Why a histogram can reveal more than the mean

Imagine that the mean stays at 0.30 between two surveys.

In the first survey, most pixels cluster tightly around 0.30. In the second, the histogram develops two peaks, one around 0.10 and another around 0.50.

The mean has not changed, but the distribution has.

This could indicate that the study area now contains more distinct pixel groups. Perhaps a portion of the site has developed denser vegetation while another portion remains sparse, or the area contains different surface types.

The histogram alerts us to a change in variability or structure that the mean would miss.

This is especially useful in ecological monitoring, where recovery rarely occurs uniformly across an entire site.

8. Why multispectral indices can tell us more about plants

RGB indices are useful because ordinary drone cameras are relatively accessible, affordable and widely available. They can help map visible greenness, distinguish some vegetation patterns and track changes in the appearance of a site.

However, RGB imagery records only visible red, green and blue channels. It cannot directly measure the near-infrared response of vegetation.

Multispectral sensors can record this additional information, allowing us to calculate indices designed around plant spectral properties.

NDVI: Normalized Difference Vegetation Index

NDVI uses the red and near-infrared bands:

NDVI = (NIR − R) / (NIR + R)

Green vegetation typically absorbs a substantial proportion of visible red light and reflects strongly in the near-infrared. NDVI uses the contrast between these two responses to highlight vegetation patterns.

Consider a simplified example using illustrative band values:

- Red = 40
- NIR = 180

The calculation is:

NDVI = (180 − 40) / (180 + 40) = 140 / 220 ≈ 0.636

The positive NDVI value reflects a stronger NIR response than red response.

NDVI is mathematically bounded between −1 and +1 when the inputs are non-negative and the denominator is positive. In many vegetation-monitoring settings, higher NDVI values are associated with denser or more vigorous green vegetation. But the relationship depends on the vegetation, sensor, environmental conditions and processing.

NDVI is not a direct measurement of plant health, and it can become less sensitive to differences in dense canopies when the index approaches saturation.

NDRE: Normalized Difference Red Edge Index

NDRE uses near-infrared and red-edge measurements:

NDRE = (NIR − RE) / (NIR + RE)

Here, RE represents the red-edge band.

The red-edge region is sensitive to changes associated with vegetation pigments and canopy condition. NDRE can therefore provide useful information in some situations where NDVI is less sensitive, particularly when examining differences in more developed or dense vegetation.

NDRE requires a sensor that records the appropriate red-edge and near-infrared bands. An ordinary RGB camera cannot produce a genuine NDRE measurement.

Why the difference matters

The important advantage of multispectral imagery is not that its indices always produce larger numbers or more attractive maps. It is that the sensor records additional wavelengths that respond differently to vegetation.

RGB indices primarily describe relationships among visible colour channels. NDVI and NDRE use spectral information that is more directly connected to characteristic vegetation reflectance.

This can make multispectral data more informative for particular plant-monitoring tasks, including investigating canopy vigour, chlorophyll-related changes and vegetation stress.

But multispectral imagery is not automatically accurate simply because it contains more bands. Results still depend on sensor quality, calibration, acquisition conditions, the index selected and the question being investigated. Field observations or other independent measurements may be needed to validate the interpretation.

9. Histograms work the same way for RGB and multispectral data

There is no fundamental difference between analysing an MPRI histogram and an NDVI histogram.

In both cases, each valid pixel has an index value. The histogram groups those values into bins, and statistics summarise the distribution.

The difference lies in what the index measures and how its values should be interpreted.

An MPRI histogram describes the distribution of green-versus-red relationships across the analysed pixels. An NDVI histogram describes the distribution of near-infrared-versus-red relationships.

Their numerical ranges may overlap, but their values are not directly interchangeable. An MPRI of 0.4 does not mean the same thing as an NDVI of 0.4.

Likewise, a higher mean in one index cannot be compared directly with a higher mean in another index to decide which area has healthier vegetation.

When comparing different surveys, it is best to compare like with like: the same index, comparable processing, the same study area and suitable acquisition conditions.

When comparing different indices, interpret each according to its formula and purpose.

10. Practical considerations when analysing drone vegetation indices

A few practical checks make a substantial difference to the reliability of your results.

Check the input bands. Confirm that the red, green, blue, NIR or red-edge layers contain the correct data. Do not assume that a software-generated index is using the bands you intended.

Check how the pixel values are stored. Digital numbers, normalised channel values and calibrated reflectance are not necessarily equivalent. Some indices are invariant to proportional scaling under ideal conditions, but camera processing and inconsistent scaling can still affect comparisons.

Watch out for shadows and exposure differences. A shadow can change the recorded pixel values even when the vegetation itself has not changed. Automatic exposure and white balance can also make two RGB surveys difficult to compare directly.

Inspect extreme values. Especially with VARI, very large values may arise when the denominator approaches zero. Check NoData handling, masks and invalid pixels before calculating statistics.

Use a consistent study area. If one survey includes more bare ground around the edges, its histogram may change even if the canopy itself has not changed. Comparing the same spatial area helps reduce this problem.

Do not rely on a single statistic. A mean is useful, but the histogram, standard deviation and spatial distribution can reveal important differences.

Validate important conclusions. If the objective is to estimate biomass, quantify ecological recovery or assess plant stress, use field observations or independent measurements where possible. A vegetation index is evidence that can support an interpretation, not a substitute for every other form of evidence.

11. The real value is understanding the data

A vegetation index is more than a colourful raster. It is a mathematical transformation of the light recorded by the sensor.

Every valid pixel contributes a value. The histogram reveals how those values are distributed. The mean summarises their average, and the standard deviation describes their spread. Examining these together helps us understand whether a landscape has a consistent spectral pattern, contains distinct groups of pixels or has changed between surveys.

With RGB imagery, MPRI and VARI provide useful ways of investigating visible-channel relationships. With multispectral imagery, indices such as NDVI and NDRE introduce additional spectral information that can be valuable for vegetation monitoring.

Neither approach should be interpreted without considering the sensor, the landscape and the purpose of the analysis.

The aim is not simply to produce a green map. It is to understand what the underlying measurements show, what they cannot tell us, and how they can be used alongside other evidence to answer a practical ecological question.

That is where drone mapping becomes much more than image processing: it becomes a way to investigate and measure change in the landscape.

12. Putting vegetation indices to work in a real survey

Vegetation indices become useful when we use them to investigate a specific question about the landscape. For example, an MPRI or VARI map can help identify areas with different visible vegetation patterns, highlight patches of bare ground

within a recovering site, or show where green vegetation appears more or less prominent. By comparing the spatial distribution of index values and their histograms between surveys, we can investigate whether changes are widespread or concentrated in particular parts of a site.

NDVI and NDRE can provide additional information from near-infrared and red-edge wavelengths, helping us investigate differences in vegetation cover, canopy development and responses associated with plant condition. These outputs can guide field visits, help identify areas that deserve closer investigation, support the monitoring of restoration or invasive-species clearing, and provide measurable indicators to compare alongside photographs, field observations and other survey data.

The important point is that an index does not need to identify a plant species or diagnose plant health on its own to be useful. It can help us find patterns, measure differences, map where those differences occur and track how they change over time. **Think of a vegetation index as an investigative tool: it helps you work out where something may be happening, how extensive it is and what you should examine next.** The more carefully we understand the data and the limitations of the index, the more confidently we can use it to answer practical questions about the landscape.

Further reading

- GeoWing Academy Understanding Histograms
- Interpreting Different Types of Images — Exploring Image Processing with OpenDroneMap
- NDVI: The Foundation for Remote Sensing Phenology — US Geological Survey
- Applying RGB-Based Vegetation Indices Obtained from UAS Imagery for Monitoring the Rice Crop at the Field Scale